

Snap DSA Report

Risk Assessment Results and Mitigations



25 August 2025



Foreword.....	16
What is New?.....	18
Previous reports.....	18
What is new in this Report?.....	19
Section 1 - Introduction.....	19
Section 2 - DSA Risk Assessment Scope.....	19
Section 3 - DSA Risk Assessment Methodology.....	20
Section 4 - DSA Risk Assessment Results.....	20
Section 5 - Specific Mitigations.....	22
Section 6 - Ongoing Risk Detection & Management.....	23
Annex - Explainer Series.....	24
Conclusion.....	24
1. Introduction.....	25
1.1 Snapchat 1.01.....	25
1.2 Critical Changes in Functionality.....	28
1.3 Snapchat Community.....	28
2. DSA Risk Assessment Scope.....	32
2.1 Approach.....	32
Scope in our previous Reports.....	33
Scope for this Report.....	33
2.2 Spotlight.....	34
2.3 Discover.....	35
2.4 Public Profiles.....	36
2.5 Snap Map.....	39
2.6 Lenses.....	44
2.7 Advertising.....	46
3. DSA Risk Assessment Methodology.....	47
3.1 Identification of Risks.....	47
3.2 Likelihood Analysis.....	48
3.3 Severity Analysis.....	49
3.4 DSA Risk Factors.....	50
3.5 Overall Potential Risk Prioritization Assessment.....	50
3.6 Snap's Mitigations.....	51
3.7 Conclusions.....	52
3.8 Supporting Documentation.....	53
4. DSA Risk Assessment Results.....	54
4.1 Category 1 - Dissemination of content that is illegal or violates our terms and conditions..	54



4.1.1 Dissemination of Child Sexual Abuse Material.....	56
Harm Description.....	56
Likelihood.....	57
Severity.....	59
DSA Risk Factors.....	60
Overall potential risk prioritization.....	63
Snap's Mitigations.....	63
Conclusion.....	66
4.1.2 Dissemination of illegal hate speech.....	67
Harm Description.....	67
Likelihood.....	68
Severity.....	69
DSA Risk Factors.....	70
Overall potential risk prioritization.....	72
Snap's Mitigations.....	73
Conclusion.....	76
4.1.3 Dissemination of information related to the sale of prohibited products or services..	76
Harm Description.....	76
Likelihood.....	77
Severity.....	78
DSA Risk Factors.....	79
Overall potential risk prioritization.....	82
Snap's Mitigations.....	83
Conclusion.....	86
4.1.4 Dissemination of Terrorist Content.....	86
Harm Description.....	86
Likelihood.....	87
Severity.....	89
DSA Risk Factors.....	89
Overall potential risk prioritization.....	92
Snap's Mitigations.....	92
Conclusion.....	96
4.1.5 Dissemination of content that infringes on intellectual property rights.....	97
Likelihood.....	97
Severity.....	98
DSA Risk Factors.....	98
Overall potential risk prioritization.....	100
Snap's Mitigations.....	101
Conclusion.....	103



4.1.6 Dissemination of Adult Sexual Content.....	104
Harm Description.....	104
Likelihood.....	104
Severity.....	106
DSA Risk Factors.....	108
Overall potential risk prioritization.....	110
Snap's Mitigations.....	110
Conclusion.....	114
4.1.7 Dissemination of Content regarding Harassment and Bullying.....	115
Likelihood.....	115
Severity.....	117
DSA Risk Factors.....	118
Overall potential risk prioritization.....	120
Snap's Mitigations.....	121
Conclusion.....	124
4.1.8 Dissemination of content that glorifies Self-Harm, including Suicide.....	124
Harm Description.....	124
Likelihood.....	125
Severity.....	127
DSA Risk Factors.....	127
Overall potential risk prioritization.....	130
Snap's Mitigations.....	131
Conclusion.....	135
4.1.9 Dissemination of content relating to violent or dangerous behavior.....	136
Harm Description.....	136
Likelihood.....	136
Severity.....	138
DSA Risk Factors.....	139
Overall potential risk prioritization.....	142
Snap's Mitigations.....	142
Conclusion.....	147
4.1.10 Dissemination of Harmful False Information.....	147
Harm Description.....	147
Likelihood.....	148
Severity.....	149
DSA Risk Factors.....	149
Overall potential risk prioritization.....	152
Snap's Mitigations.....	152
Conclusion.....	155



4.1.11 Dissemination of Fraud and Spam.....	155
Harm Description.....	155
Likelihood.....	156
Severity.....	159
Overall potential risk prioritization.....	160
Snap's Mitigations.....	160
DSA Risk Factors.....	167
Conclusion.....	169
4.1.12 Dissemination of information related to Other Illegal Activities.....	169
Harm Description.....	170
Likelihood.....	170
Severity.....	171
DSA Risk Factors.....	171
Overall potential risk prioritization.....	173
Snap's Mitigations.....	174
Conclusion.....	176
4.2 Category 2: Negative Effects on Fundamental EU Rights.....	177
4.2.1 Right to Human Dignity.....	178
Likelihood.....	179
Severity.....	179
DSA Risk Factors.....	180
Overall potential risk prioritization.....	180
Snap's Mitigations.....	181
Conclusion.....	183
4.2.2 Right to Freedom of Expression and Assembly.....	184
Likelihood.....	184
Severity.....	185
DSA Risk Factors.....	185
Overall potential risk prioritization.....	188
Snap's Mitigations.....	188
Conclusion.....	192
4.2.3 Right to Private Life.....	193
Likelihood.....	193
Severity.....	194
DSA Risk Factors.....	194
Overall potential risk prioritization.....	197
Snap's Mitigations.....	197
Conclusion.....	200
4.2.4 Right to Data Protection.....	200



Likelihood.....	201
Severity.....	202
DSA Risk Factors.....	203
Overall potential risk prioritization.....	203
Snap's Mitigations.....	203
Conclusion.....	207
4.2.5 Right to Non-Discrimination and Freedom of Religion.....	207
Likelihood.....	207
Severity.....	208
DSA Risk Factors.....	209
Overall potential risk prioritization.....	209
Snap's Mitigations.....	210
Conclusion.....	216
4.2.6 Children's Rights.....	216
Likelihood.....	217
Severity.....	217
DSA Risk Factors.....	218
Overall potential risk.....	218
Snap's Mitigations.....	218
Conclusion.....	222
4.2.7 Right to Consumer Protection.....	223
Likelihood.....	223
Severity.....	225
DSA Risk Factors.....	226
Overall potential risk prioritization.....	226
Snap's Mitigations.....	227
Conclusion.....	230
4.2.8 Right to Property.....	230
4.3 Category 3: Negative effects on Public Security.....	231
4.3.1 Negative Effects on Democratic and Electoral Processes.....	231
Likelihood.....	232
Severity.....	233
DSA Risk Factors.....	233
Overall potential risk prioritization.....	236
Snap's Mitigations.....	236
Conclusion.....	251
4.3.2 Negative Effect on Civic Discourse.....	251
Likelihood.....	251
Severity.....	252



DSA Risk Factors.....	253
Overall potential risk prioritization.....	253
Snap's Mitigations.....	254
Conclusion.....	258
4.3.3. Negative Effect on Public Security.....	258
Likelihood.....	259
Severity.....	259
DSA Risk Factors.....	261
Overall potential risk prioritization.....	261
Snap's Mitigations.....	262
Conclusion.....	267
4.4 Category 4: Negative Effects on Public Health.....	267
4.4.1 Negative Effects on Public Health.....	269
Likelihood.....	269
Severity.....	270
DSA Risk Factors.....	270
Overall potential risk prioritization.....	271
Snap's Mitigations.....	271
Conclusion.....	275
4.4.2 Negative Effects on Gender-Based Violence.....	275
Likelihood.....	275
Severity.....	276
DSA Risk Factors.....	277
Overall potential risk prioritization.....	277
Snap's Mitigations.....	278
Conclusion.....	284
4.4.3 Negative Effects on Minors.....	284
Likelihood.....	286
Severity.....	288
DSA Risk Factors.....	292
Overall potential risk prioritization.....	294
Snap's Mitigations.....	295
Conclusion.....	304
4.4.4 Serious Negative Consequences on Physical and Mental Well-Being.....	304
Likelihood.....	304
Severity.....	311
DSA Risk Factors.....	312
Overall potential risk prioritization.....	312
Snap's Mitigations.....	313



Conclusion.....	318
5. Specific Mitigations.....	320
5.1 Snapchat Design and Function.....	320
5.1.1 Introduction.....	320
5.1.2 Oversight and Administration.....	320
Roles and Responsibilities.....	320
5.1.3 Adaptations and Mitigations.....	321
Friending, Chat and Private Stories.....	321
Spotlight and Discover.....	322
Spotlight.....	322
Discover.....	323
Public Profile.....	323
Snap Map.....	325
Lenses.....	328
Advertising.....	328
5.1.4 Integrations with other mitigations.....	328
Terms.....	328
Content Moderation.....	328
Content Distribution.....	328
5.1.5 Online Interface Design Process.....	329
5.1.6 Conclusion.....	331
5.2 Terms.....	332
5.2.1 Introduction.....	332
5.2.2 Oversight and Administration.....	332
Change Management.....	332
Roles and Responsibilities.....	334
5.2.3 Terms and Conditions.....	335
Terms of Service.....	335
Community Guidelines.....	339
Privacy Policy.....	344
Product Specific Terms.....	345
Advertising.....	345
5.2.4 Accessing Terms and Conditions.....	346
5.2.5 Support Site.....	347
5.2.6 Languages.....	348
5.2.7 Readability.....	348
5.2.8 Conclusion.....	349
5.3 Transparency.....	349
5.3.1 Information we provide on our website.....	349



Privacy, Safety, and Policy Hub.....	350
Policy Center.....	351
Privacy Center.....	351
Safety Center.....	352
Parents.....	355
Transparency Center.....	356
News Page.....	357
5.3.2 Information provided in app stores.....	358
5.3.3 Information we provide in our application.....	359
Onboarding process.....	359
Just-in-time notifications.....	363
Thematic awareness and notices.....	366
5.3.4 Languages.....	369
5.3.5 Conclusion.....	369
5.4 Content Moderation.....	369
5.4.1 Approach.....	369
Snapchat Design and Function.....	370
Community Guidelines and Terms of Service.....	370
Content Moderation.....	370
Enforcement.....	371
5.4.2 Oversight and Administration.....	371
Roles and Responsibilities.....	371
Human Content Moderators.....	372
Content Moderator Teams.....	372
Selection of Moderators.....	374
Recruitment of Moderators.....	374
Content Moderator Processes.....	374
Moderator Wellness as a Priority.....	377
Automated Moderation.....	378
Training and Testing.....	378
Quality Assurance.....	378
Continuous Improvement.....	379
5.4.3 Broadcast Content - Proactive Moderation.....	379
Detection.....	379
Proactive Safety Detection System.....	379
CSEAI.....	380
Abusive Language Detection.....	380
Content Reviews.....	381
5.4.4 Product-Specific Moderation.....	381



Spotlight.....	381
Discover.....	382
Media Partnered Content.....	383
Public Profiles.....	384
Snap Map.....	385
Lenses.....	386
Advertising.....	387
5.4.5 Reactive Moderation (Reporting).....	389
Content-level in app reporting.....	389
Discover.....	390
Snap Map.....	390
Lenses.....	390
Messages.....	391
Reporting on the web.....	393
Account level in-app reporting.....	394
Illegal Content Notice (Art. 16).....	398
5.4.6 Conclusion.....	402
5.5 Enforcement.....	402
5.5.1 Introduction.....	402
5.5.2 Review & Enforcement.....	402
Severe Harms.....	403
Strike System.....	403
Transparency.....	404
5.5.3 Notification of Criminal Offenses (Art. 18).....	405
Proactive referrals to law enforcement and governmental agencies.....	405
Law enforcement takedown and information requests (Articles 9 and 10).....	405
Law enforcement orders to provide information (Article 10).....	405
5.5.4 Notice and Appeals System.....	406
Notice to Reporter.....	406
Account-Level Notice and Appeals.....	408
Content-Level Notice and Appeals.....	415
5.5.5 Effectiveness of Enforcement.....	419
5.5.6 Protections against Misuse (Art. 23).....	420
Suspending the Processing of Notices and Complaints.....	420
5.5.7 Conclusion.....	420
5.6 Algorithmic Systems.....	421
5.6.1 Introduction.....	421
5.6.2 Content Recommendation Systems.....	421
5.6.3 How do our Content Recommender Systems work?.....	422



Benefits.....	424
5.6.4 Adaptation of Snap Algorithmic Systems to Mitigate Systemic Risk.....	424
Enabling user choice in content prioritization.....	424
5.6.5 Oversight and Administration.....	425
Roles and Responsibilities.....	425
Algorithmic System Review.....	425
Documentation Standards.....	426
5.6.6 Model Development and Deployment.....	426
Development Guidelines.....	426
Common Infrastructure.....	426
Pre-Launch Testing.....	427
Pre-Launch Legal Review.....	427
5.6.7 Adaption and Testing.....	427
Summary.....	427
Illegal or violating content.....	429
Lack of user understanding.....	431
Intrusive personalized recommendations.....	431
Discrimination.....	432
Rapid and Widespread illegal or false content & crisis exposure.....	432
Filter bubbles.....	432
Erroneously excluding content.....	433
Viewers could be watching our content but not enjoying content.....	433
5.6.8 Change Management.....	435
5.6.9 Monitoring and Quality Assurance.....	435
Performance Monitoring.....	435
Quality Assurance.....	435
5.6.10 Conclusion.....	435
5.7 Advertising Systems.....	436
5.7.1 Introduction.....	436
5.7.2 How do our Advertising Systems Work?.....	436
5.7.3 Benefits.....	437
5.7.4 Adaptation and Testing.....	439
Summary.....	439
Invasion of Privacy – Reasonable and Proportionate Targeting.....	441
Special category data – No sensitive data use.....	442
Discrimination – Special Targeting Models.....	442
Harmful or illegal content – Advertising policies.....	442
Policy-violating or illegal content – Advertising Review.....	443
Bypassing Moderation Controls – Advertising Reporting.....	444



Unclear Commercial Intent – Ad Markers.....	444
Political ads – Transparency Safeguards.....	446
Personal Data Use for Targeting – User Choice.....	447
Lack of visibility – Ads Gallery.....	449
Freedom of Expression.....	452
5.7.5 Conclusion.....	452
5.8 Protection of Minors.....	453
5.8.1 Introduction.....	453
5.8.2 Administration and Oversight.....	453
Roles and Responsibilities.....	453
5.8.3 Overview and Approach.....	454
Age Appropriate Design Code.....	456
Privacy, Safety, and Security of Minors on Snapchat.....	457
Advertisements for Minors.....	458
Identifying Minors.....	458
Registration and access to Snapchat.....	458
Access to certain content / features.....	461
Oversight.....	464
Ongoing evaluation.....	465
European Commission Art 28 DSA Guidelines.....	466
Transparency to Minors.....	470
5.8.4 Safeguards.....	471
App Store Level Safeguards.....	471
Device-Level Safeguards.....	471
Platform-Level Safeguards.....	473
Product-Level Safeguards.....	474
Lenses.....	474
Public Content.....	475
Public Profiles.....	475
Creating Public Content.....	475
Viewing Public Content.....	475
Discover.....	475
Spotlight.....	476
Snap Map.....	476
Advertisements.....	477
Reporting and Blocking.....	479
Private Messaging.....	480
Friending.....	480
Family Center / Parent Tools.....	481



5.8.5 Conclusion.....	485
5.9 Content Authenticity.....	486
5.9.1 Introduction.....	486
5.9.2 Risk Assessment Results.....	486
5.9.3 Mitigations.....	487
Guidelines, policies, and practices.....	487
Creation.....	488
Dissemination.....	489
User Guidance on Generative AI features.....	491
Enforcement.....	493
Partnerships.....	493
5.9.4 Conclusion.....	495
5.10 Trusted Flaggers.....	495
5.10.1 Trusted Flagger Program.....	495
5.10.2 Onboarding a new trusted flagger.....	496
5.10.3 DSA Trusted Flaggers.....	496
5.10.4 Trusted Flagger Program Trends.....	496
5.10.5 Conclusion.....	503
5.11 Dispute Settlement Bodies.....	503
5.11.1 Overview and Approach.....	503
5.11.2 Enquiries.....	504
5.11.3 Further considerations.....	504
5.11.4 Conclusion.....	505
5.12 Codes and Crisis Protocols.....	505
5.12.1 Cooperation.....	505
5.12.2 Codes of Practice.....	507
EU hate speech Code.....	508
FSM Code of Conduct.....	509
EU disinformation code.....	509
Article 28 DSA Guidance.....	510
5.12.3 Crisis Protocols.....	510
Crisis Protocol Case Study – October 7 and Israel-Hamas Conflict.....	511
6. Ongoing Risk Detection and Management.....	514
6.1 Platform Principles-based Framework.....	514
6.2 DSA Compliance Team and Cross-Functional Working Groups.....	515
6.2.1 Introduction.....	515
6.2.2 Roles and Responsibilities.....	516
6.2.3 DSA Independent Compliance Function.....	516
6.2.4 Independent Compliance Function Collaboration.....	517



6.2.5 Compliance Officer Designation.....	518
6.2.6 Compliance Officer Qualifications.....	518
6.2.7 Operation of the Independent Compliance Function.....	519
Responsibilities of the Independent Compliance Function.....	519
Oversight and Monitoring of Snap’s DSA Compliance.....	519
DSA Management Body.....	520
Organization of the Independent Compliance Function.....	520
Independence Requirements.....	520
Communication of DSA Obligations to Snap Employees.....	521
6.2.8 DSA Cross-Functional Governance Team.....	521
Responsibilities of the DSA Governance Team.....	521
6.2.9 Points of Contact.....	522
Designation, Publication, and Change Management.....	522
Point of Contact for the Authorities.....	523
Point of Contact for Users.....	523
Legal Representative.....	524
6.2.10 DSA Supervisory Fee.....	524
6.3 Privacy and Safety by Design.....	524
6.3.1 DSA Risk Management.....	524
6.3.2 Privacy and Safety by Design review process.....	525
6.3.3 Holistic Digital Risk Management.....	527
6.3.4 Digital and Data Impact Assessment (DDIA) Template.....	528
6.3.5 DSA Critical Impact Check.....	530
6.4 Prevalence Testing.....	531
6.5 External Request Monitoring and Review.....	540
6.6 Digital Well-Being Index (DWBI) Initiative.....	541
6.7 Snap Advisory Groups.....	542
6.7.1 Safety Advisory Board.....	542
6.7.2 Snap Council for Digital Well-Being.....	543
6.7.3 Regular External Engagement.....	545
6.7.4 Dedicated DSA Risk Assessment Workshop.....	545
6.8 Audit.....	546
7. Conclusion.....	547
8. Final Words.....	551
Annex.....	552
Community Guidelines.....	553
Overview.....	553
Community Guidelines: Explainer Series.....	555
Sexual Content.....	555



Threats, Violence & Harm.....	557
Hateful Content, Terrorism and Violent Extremism.....	559
Harassment and Bullying.....	561
Illegal or Regulated Activities.....	563
Harmful False or Deceptive Practices.....	565
Severe Harm.....	568
Snapchat Moderation, Enforcement, and Appeals.....	568



Foreword

This Risk Assessment Results and Mitigations Report (**Report**) has been prepared to comply with Snap’s obligations under Article 42.4.(a), (b) and (e) of Regulation (**EU**) 2022/2065 on a Single Market For Digital Services and amending Directive 2000/31/EC (the “Digital Services Act” or “DSA”).

This Report is divided into eight sections: (1) Introduction; (2) DSA Risk Assessment Scope; (3) DSA Risk Assessment Methodology; (4) DSA Risk Assessment Results; (5) Specific Mitigations; (6) Ongoing Risk Detection and Management; (7) Conclusion; and (8) Final Words.

Since our 2024 Report, we have made a number of updates and these are summarized in a “What is New?” section that we have added to this Report. At a high level, these updates reflect three themes that we have observed during the last year:

- **European Commission’s Risk Assessment Recommendations:** On 7 May, the Commission held a DSA Multi-Stakeholder Workshop on Systemic Risks. At the end of this Workshop, the Commission explained what it expected to see in the next reports from Very Large Online Platforms and Very Large Search Engines. We have assessed our Report against these recommendations as follows:

Commission Request	Covered in Snap’s Report?	Enhancements in this Report?
1. Transparency: The reports need to provide methodologies and cover all risks in the DSA.	Yes	None. We already include our methodology and cover all risks in the DSA.
2. Design: The reports need to explain how the design of the service impacts risk, in addition to the impact from the content itself.	Yes	Yes. Our reports already include extensive analysis of the design of our service and how this impacts our risk assessment. To make this clearer, this year we have included explicit references to the impact of the DSA’s specific risk factors, including service design, on each harm in Section 4 .
3. Evidence: The reports must explain the data and other evidence relied on and explain how effectiveness is assessed.	Yes	None. Our reports already include data and other evidence to support their conclusions and explain how



		effectiveness is assessed.
4. Mitigations: The reports must provide sufficient descriptions of the mitigations deployed and how they impact the assessment.	Yes	None. We already include extensive descriptions of our mitigations in Section 5 (Mitigations) and explain how these mitigations impact our risk assessment in Section 4 (Risk Assessment Results).
5. AI: The reports must assess the impact and mitigations for AI content.	Yes	None. We already assess the impact of AI content in Section 4 (Risk Assessment Results), particularly with respect to the dissemination of harmful false information in Section 4.1.10 . We provide detailed descriptions of our mitigations for AI content in Section 5.9 (Content Authenticity).
6. CSO Engagement: The reports must rely on available evidence. Must explain expert evidence relied on and engagement with CSOs is expected. Must take into account feedback on risk assessments provided by CSOs.	Yes	Yes. Our previous reports already explained the research and other evidence we rely on. This includes our regular engagement with Snap's Safety Advisory Board, external Civil Society Organisations and other experts. This year, we also organised a dedicated risk assessment workshop with CSOs to obtain specific feedback on our risk assessment. In response to their feedback, amongst other things, we have also included additional violative view rate data to further evidence our conclusions especially with respect to minors. See Section 3 (Methodology) and Section 4 (Risk Assessment Results). See Section 6.7 (Snap Advisory Groups).

- **Minors** - This year has seen the publication of the European Commission's Guidance on Article 28 to ensure a high level of privacy, safety and security for minors. This guidance is the result of extensive work from the Commission and other stakeholders. We have been working proactively to support the Commission's efforts to introduce this guidance and strongly support its goals. We are currently working to assess Snapchat's in-scope services against the recommendations in the guidance. We have included a summary of our initial assessment in Section 5.8 (Protection of Minors), and also referred to this



assessment in Sections 4.2.6 (Children's Rights), 4.4.3 (Negative Effects on Minors) and 5.12 (Codes and Crisis Protocols).

- Out of Court Dispute Settlement - This year has seen a significant increase in the number of out of court settlements received by Snap. While the numbers are still very low, reflecting Snap's effective terms, moderation and enforcement measures, responding to these cases requires extensive resources. This is despite the fact that the majority of cases are decided in Snap's favour. We feel strongly that improvements must be made to how out of court settlement works to ensure only valid and justified cases are brought forward and costs are minimised. We consider this further in Section 5.11 (Dispute Settlement Bodies).

As with last year, we look forward to continuing our constructive dialogue with Commission, our Digital Service Coordinator and other stakeholders to further the objectives of the DSA, as well as encouraging continued close collaboration with regulators and other organisations covering the General Data Protection Regulation, Audio Visual Media Services Directive and the Digital Markets Act to ensure a holistic approach to privacy, safety and security across these digital platform laws.

What is New?

Snap is required to complete a report every year setting out the results of its risk assessment and details of its mitigations pursuant to Article 42(4).

Previous reports

Snap has prepared two previous risk assessment and mitigation reports pursuant to Article 42(4):

- Our first report was completed in August 2023 (the "2023 Report") and was published on Snap's website [here](#) within 3 months of Snap having received its final auditors report for the corresponding period (the "2023 Public Report").
- Our second report was completed in August 2024 (the "2024 Report") and was sent to the Commission without undue delay as required by Article 42(4). Snap's reports are published on its European Union Transparency page within 3 months of Snap having received its final auditors report (as this includes an audit of Snap's compliance with its risk and mitigation assessment and reporting obligations) pursuant to Article 42(4) i.e. approximately 1 year and 3 months after each report is completed. We are working to publish our second report, taking account feedback from civil society organisations on the extent of our redactions following our first dedicated DSA Risk Assessment Workshop earlier this year.



What is new in this Report?

Section 1 - Introduction

- [Snapchat 101](#) -
 - In all material respects, Snapchat's in-scope services remain as described in our previous reports. Our data shows that the vast majority of our users are still primarily using the messaging aspects of our platform, and we continue to believe this is an important lens through which to view Snapchat.
 - **We deployed one functionality that was considered likely to have a critical impact on the risks identified pursuant to Article 34 of the DSA.** This was the launch of additional posting options for 16-17 years olds. This new functionality underwent extensive testing and evaluation, and presentation to the Commission and other regulators, prior to launch. Our monitoring of key safety metrics showed that the EU launch performed in line with expectations.
 - We chose not to launch the Simple Snapchat change that was flagged in our 2024 Report. This was primarily a cosmetic change and was not expected to have any critical impact on our risk assessment.
- [Snapchat Community](#) - We continue to observe positive growth in our user base globally. In the European Union ("EU") we grew to 93.7 million average monthly active recipients of our Snapchat app (as at 1 January 2025). Our community demographics have not seen any significant changes since our previous reports.

Section 2 - DSA Risk Assessment Scope

- [Scope Assessment](#) - Since the 2024 Report: (i) our Snapchat designation has not changed; (ii) the Commission has not issued any new guidance relating to scope and (iii) the functionality of Snapchat has not significantly changed. We have therefore confirmed that Snap still considers the Spotlight, Discover, Public Profiles, Snap Map, Lenses, and Advertising services of Snapchat to fall within the scope of our risk assessment and mitigation obligations in Articles 34 and 35. We have confirmed that Snap continues to consider My AI and other similar generative AI tools made available by Snap to be out of scope of Snapchat's designation except for one advertising case identified below.
- In our short descriptions of the in-scope aspects of [Spotlight](#), [Discover](#), [Public Profiles](#), [Snap Map](#), [Lenses](#), and [Advertising](#) services of Snapchat, we have noted the following changes since our previous report:



- As previously communicated to the Commission, we have decided not to proceed with the potential simplification of Spotlight and Discover tabs known as Simple Snapchat.
- We have updated our summary of Public Profiles features to incorporate the new public posting options available for 16-17 year olds. This new feature was subject to extensive testing and evaluation, and was presented to the Commission and other regulators, prior to launch. Our monitoring of key safety metrics showed that the EU launch performed in line with expectations. We have provided more details on the specific mitigations for this feature in Section 5.

Section 3 - DSA Risk Assessment Methodology

- [DSA Risk Assessment Methodology](#) -
 - We have not made any material changes to our risk assessment methodology since our 2024 Report.
 - We have added an additional paragraph to explain the inclusion of 'violative view rate' (VVR) data as additional evidence of the likelihood of adult and minor recipients of Snapchat's in-scope services being exposed to illegal or otherwise violating content. This has been included to address civil society feedback in our first Snapchat risk assessment workshop that organisations would like us to include more specific information on the extent to which users, especially minors, may be exposed to illegal or violative content disseminated on Snapchat.
 - We have also added an additional paragraph to explain how we have considered the specific risk factors set out in Article 34(2) of the DSA. Our previous reports already explained how these factors, such as the design of our recommender systems, impact each harm and how increased risks, if any, have been addressed. In this Report, we have more explicitly set this out and included a dedicated sub-section on DSA risk factors for each harm in Section 4. This has been included to ensure that we meet a Commission recommendation to clearly demonstrate the DSA's specific risk factors had been considered.

Section 4 - DSA Risk Assessment Results

- We have updated the likelihood, severity, overall potential risk prioritization assessments, mitigation assessments. We have confirmed that there have been no changes to the conclusions we reached in our 2024 Report that we have reasonable, proportionate and effective mitigation measures for each systemic risk identified in Article 34 of the DSA:
 - Category 1 - Dissemination of content that is illegal or violates our terms and conditions i.e. dissemination of [CSAM](#), [Illegal Hate Speech](#), [Sale of Prohibited Goods and Services](#), [Terrorist Content](#), [IP Infringement](#), [Adult Sexual Content](#),



[Harassment and Bullying](#), [Self-Harm and Suicide](#), [Violent and Dangerous Behaviour](#), [Harmful False Information](#), [Fraud and Spam](#), and [Other Illegal Activities](#).

- Category 2 - Negative Effects on Fundamental EU Rights i.e. [Human Dignity](#), [Freedom of Expression and Information](#), [Private Life](#), [Data Protection](#), [Non-Discrimination and Freedom of Religion](#), [Children's Rights](#), [Consumer Protection](#) and [Property](#).
 - Category 3 - Negative Effects on [Democratic and Electoral Processes](#), [Civic Discourse](#) and [Public Security](#).
 - Category 4 - Negative Effects on [Public Health](#), [Gender-Based Violence](#), [Minors](#) and [Physical and Mental Well-Being](#).
- In line with updates to our Methodology (Section 3 above), we have included additional internal evidence (i.e. violative view rate data minors and adults) in response to feedback from civil society organisations and more explicitly identified how we have assessed the specific risk factors identified in Article 34 of the DSA in response to a Commission recommendation.
 - We have seen a substantial reduction in prevalence rates across all of the illegal and other violating content categories that we monitor (see our update on prevalence testing below) and observed low violative view rates across all harm categories. Evidence of a low number of reports regarding violative ads also demonstrates the effectiveness of our pre-publication ad review and rejection processes. As a result, we have been able to lower the relative likelihood of three of our risk categories from medium relative likelihood to our lowest relative likelihood category:
 - [Adult Sexual Content](#)
 - [Bullying & Harassment](#)
 - [Fraud and Spam](#)
- We had identified all three categories as a focus for our ongoing monitoring and management of risk in the conclusion of our 2024 Report. We have noted that we will continue to monitor these categories, but are very pleased with the progress.
- We have added to our assessment concerning the dissemination of information relating to the [Sale of Prohibited Goods and Services](#) to include information relating to vaping. We have been made aware of a particular issue in the Netherlands regarding the sale of prohibited vaping products. This related to the private functionality of Snapchat which is out of scope of this Report. Nevertheless, we note that we had existing measures in place to mitigate this risk and have made further improvements to better detect and take enforcement action against accounts and content involved in this illegal activity.
 - We have added a specific paragraph to reference to the new Article 28 Guidelines in our assessment of the risk of [Negative Effects on Minors](#). We cross-referred to [Section 5.8](#) (Protection of Minors) where we have outlined our initial assessment of the



recommendations with respect to the in-scope services of Snapchat. We have continued to conclude that Snapchat's in-scope services have reasonable, proportionate and effective measures to mitigate negative impacts on minors.

- General - We have also made a few more minor enhancements in response to feedback from civil society organisations at our recent risk mitigation workshop:
 - We have updated the harm descriptions to provide additional clarity on the scope of each risk category in [Section 4](#).
 - We have also merged the 'highlights' paragraphs into the specific mitigations table under the 'Snap's mitigations' heading for each harm in [Section 4](#). This was done to streamline the Report and make it easier for readers to locate the corresponding detail we have provided about our specific mitigations in [Section 5](#).

Section 5 - Specific Mitigations

- [Snapchat Design and Function](#) - We have not noted any significant changes to Snapchat's Design and Function mitigations. We have removed reference to Simple Snapchat as we decided not to launch this change. We have also made adjustments to reflect the launch of additional public posting options for 16-17 year old accounts.
- [Terms](#) and [Transparency](#) - We have not noted any significant changes to our terms or transparency mitigations.
- [Moderation](#) and [Enforcement](#) - We have not noted any significant changes to our moderation and enforcement mitigations. We have made some further adjustments to streamline the information for easier presentation and efficiency reasons.
- [Algorithmic Systems](#) and [Targeted Advertising](#) - We have provided some additional information regarding algorithmic systems used for the personalization of Maps and Lenses. We have also provided additional information regarding political advertising.
- [Protection of Minors](#) - We have added a specific section on the new Article 28 Guidelines. In this section, we have provided a summary of our initial analysis comparing the in-scope services of Snapchat with the measures recommended in the Guidelines to ensure a high level of privacy, safety and security. We have continued to conclude that Snapchat's in-scope services have reasonable, proportionate and effective measures to mitigate negative impacts on minors, but continue to assess the Guidelines and monitor this area to determine if any changes are needed.
- [Content Authenticity](#) - We have updated this section with further information relating to our transparency safeguards both in respect to the use of generative AI in content

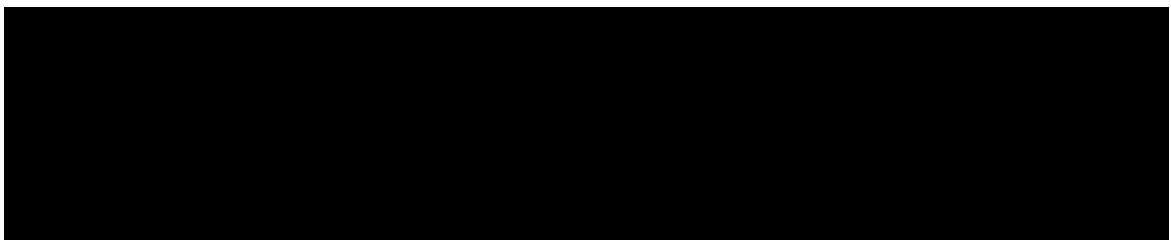


creation (which we continue to consider to be out of scope of Snap's designation (save for certain commonplace ad creation tools) and are providing for context) and in dissemination to the public on Snapchat's inscope services.

- [Trusted flaggers](#) - We have provided additional information regarding Snap's trusted flagger program, as updating our statistics and trend analysis.
- [Dispute Settlement Bodies](#) - We have provided updates regarding our ongoing engagement with out-of-court settlement bodies. Although we are receiving relatively low numbers of dispute settlement body escalations, these are still placing a significant burden on legal and operational teams.
- [Codes and Crisis Protocols](#) - We have added a specific paragraph to refer to the new Article 28 Guidelines. We have cross-referred to [Section 5.8](#) (Protection of Minors) where we have outlined our initial assessment of the recommendations with respect to the in-scope services of Snapchat.

Section 6 - Ongoing Risk Detection & Management

- [Platform Principles Framework](#) and [DSA Compliance Team and Cross-Functional Working Groups](#) and [Privacy and Safety by Design](#): There have been no significant updates to these sections. We have provided some updated information regarding the DSA Compliance Team.
- [Prevalence Testing](#) - We continue to be extremely pleased with the progress we have observed from our prevalence testing over the last year. This demonstrates that the effectiveness of our proactive detection mechanisms, agent training and other content moderation and enforcement efforts has continued to increase significantly since our 2024 Report. In particular:
 - We have observed an additional significant decrease in our overall 'Policy Violating Prevalence' (PVP) rate.
 - All violating content categories in the top ten have significantly reduced PVP rates.
 - All illegal or other violating content categories now fall within our lowest in our likelihood category [REDACTED]





We also observed a moderate increase in the percentage of distinct violating Snaps. This is determined to be caused by Snap expanding its prevalence testing to cover new areas. Feedback from these new areas will result in further adjustments to the moderation systems and we expect to see a gradual decrease in these levels in due course.

- [External Request Monitoring and Review](#) - We have confirmed that we continue to produce transparency reports and monitor content moderation and enforcement data, advertising review rejections, reporting and enforcements, 'privacy, data protection and DSA' requests and general community support requests. We continue to use this data to support the conclusions reached in this Report.
- [Digital Well-Being Index \(DWBI\) Initiative](#) - We have updated this section to report on an event we hosted on March 19, 2025 to share the latest findings from the *Digital Well-Being Index (DWBI)*, followed by an informal discussion with subject-matter experts.
- [Snap Advisory Groups](#) - We have provided further information on the progress of our work with the Snap Safety Advisory Board and Snap Council for Digital Well-Being. We have also provided additional information on our external engagement with experts and civil society organizations. In particular, we have shared a report on our first dedicated DSA Risk Assessment and Mitigation workshop in Brussels and online on July 10, 2024 and the steps we have taken to address their feedback.
- [Audit](#) - We have noted the completion of our second external DSA audit of Snap's compliance with its obligations under Chapter 3 of the Digital Services Act for the audit period between 1 July 2024 and 30 June 2025 pursuant to Article 37. We have noted that the conclusion of the 2025 Audit Report is expected to indicate that Snap complied with all DSA obligations within the audit period.

Annex - Explainer Series

- [Annex - Explainer Series](#) - There have been a number of minor updates to our Explainer Series and the new versions are included.

Conclusion

- [Conclusion](#) - **We note that we have carried out a risk assessment of Snapchat's in-scope services and continued to confirm that we have in place reasonable, proportionate and effective mitigation measures, tailored to the specific systemic risks identified.** We have also reflected on the progress with the areas for improvement identified in our 2024 Report, as well as identifying new areas for improvement over the coming year.



1. Introduction

At Snap, our mission is to contribute to human progress by empowering people to express themselves, live in the moment, learn about the world, and have fun together.

Even as Snap grows and faces new opportunities and challenges, we remain grounded in kindness. Our engineers, designers, product managers, and other team members build our products and services to serve people. The well-being of the community informs our decision making, which in turn creates more value for our business over the long term.¹

1.1 Snapchat 1.01

Snapchat is a communications app designed for people ages 13 and up, who primarily use it to talk with their close friends, similar to the ways they interact in real life. It's similar to how older generations use text messaging or their phone to stay in touch with friends and family. Since our 2024 Report, our data shows that the vast majority of our users are still primarily using the messaging aspects of our platform. While the products detailed in this report and within scope of the DSA primarily revolve around our public content surfaces, our core use is a messaging app, which sets us apart from many other VLOPs. We believe this is an important lens through which to view Snap and our platform.

¹ See our [Citizen Snap Report](#) for more details.



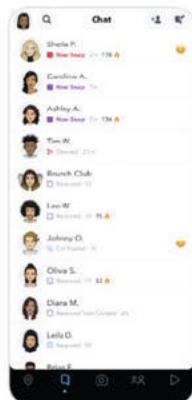
We purposely designed Snapchat differently from traditional social media. It doesn't open to a public news feed powered by an algorithm with likes and comments. Instead, Snapchat opens to a camera and has five tabs: Camera, Chat, Map, Stories, and Spotlight.²

The next section provides a reminder of our platform architecture.

Platform Architecture



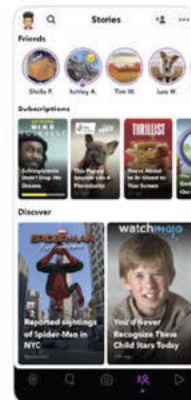
Map



Chat



Camera



Discover



Spotlight

² View our [Snapchat 101 video](#) for more details.



Camera

Snapchat opens into a camera, making it an easy and visual way for people to share what's on their mind with the people that matter most to them. Snapchatters can Snap a quick video or photo with our augmented reality Lenses to put fun and educational layers on the world, and get creative by overlaying text, stickers, and more.

Chat

To the left of the Camera is Chat, where Snapchatters can talk with their friends and family using text and pictures. Chats will show when both friends are there at the same time. They'll also indicate when a friend has opened and viewed a Snap.

Snapshots and Chats delete-by-default to mirror real life conversations, where what one says or does isn't recorded forever and shared with a bunch of strangers. This helps people feel more comfortable expressing themselves, the same way they would if they were just hanging out with friends in person. While Chats and Snaps delete by default, Snapchatters do have the option to save Chats – simply by tapping on the ones they want to save.

In Chat, you can also make voice and video calls and join group conversations and chat with My AI, our chatbot powered by OpenAI's ChatGPT technology.

Map

Swipe to the left of Chat for the Map. Our Map is an interactive way for Snapchatters to share their favorite spots, discover new places, and see what their friends are up to – but only if they choose to share their location with their friends.

Profile

My Profile features a user's Snapchat info, like their [Bitmoji](#) (which is an avatar representation of the user), location on the Map, friend info, and more. My Profile is also where Snapchatters can manage their friendships, and report, block, or remove a friend.

Public Profile

Public Profiles enable Snapchatters to be discovered in the app. If Snapchatters want a Public Profile, they will need to create one first. Once they have created a Public Profile, they can showcase their favorite public Snaps and share Lenses and other information.

Discover

Swipe to the right of the Camera for Stories. Snapchatters can add Snaps to their Stories to share more of their day with friends and family, and scroll down to discover new Stories and content about the world — produced by trusted media publishers and popular creators.



Spotlight

Right next to Stories is our entertainment platform, Spotlight. This is where Snapchatters can submit and watch short, fun, and creative videos for our community.

In [Section 2](#) of this report we provide more details on the products and services that are in scope of the DSA Risk Assessment.

1.2 Critical Changes in Functionality

Since our 2024 Report, **we have not deployed functionalities that were likely to have a critical impact on the risks identified pursuant to Article 34 of the DSA.** However, as envisioned in the prior report, Snap introduced a change allowing older teenagers (ages 16–17) in the EU, EEA, and UK to access Public Profiles and public posting options. This enables them to create or update a publicly accessible profile, share Public Stories, and submit content for Spotlight.

Before launch, Snap conducted a risk assessment and added mitigations. For example, Public Stories from 16–17 year olds are only shown to friends, followers, or mutual connections - not to the wider community through Discover. Teens can receive Story replies but cannot start direct chats from them, and replies are moderated more strictly for this age group. They can also turn off replies or block certain terms, and all users retain easy reporting tools.

Data from outside the European Union, where the feature was already available, suggests these measures are effective. [REDACTED]

[REDACTED]

Consequently, Snap concludes that the residual risk of this change is low, with safeguards proving effective in protecting older teenagers while allowing them access to public posting features.

1.3 Snapchat Community

We reach over 932 million³ monthly active users around the world, and we have over 469 million⁴ daily active users globally.

We provide information on the average monthly active recipients of our Snapchat app, across the EU and per EU Member State, in our [European Union transparency page](#) on our website.

³ Snap Inc. public data Q2 2025, see <https://investor.snap.com>.

⁴ Snap Inc. public data Q2 2025, see <https://investor.snap.com>.



[REDACTED]

[illegible]



Our EU Snapchatter community consists of a diverse range of ages and genders. While Snapchat does have a young demographic, the largest age category is 18-24; the second largest age category is 25-34, 35+ makes up the third place, and 13-17 is the smallest age category.

[REDACTED]



[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

	[REDACTED]		
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

[REDACTED]
associated with their account. The first versions of Snapchat did not require a birthdate on registration.



2. DSA Risk Assessment Scope

2.1 Approach

Articles 34 and 35 apply to Very Large Online Platforms designated by the European Commission. Snapchat was designated as a Very Large Online Platform by the Commission on 25 April 2023 because the Average Monthly Active Recipients of Snapchat exceeds 45 million.

The Commission Decision to designate Snapchat as a Very Large Online Platform states that it only applies to services provided as part of Snapchat that meet the definition of online platform laid down in Article 3, point (i), of Regulation (EU) 2022/2065. The designation does not apply to services that are provided together with Snapchat, such as a private messaging service, and that, based on their technical functionalities, do not in themselves meet the definition of online platform laid down in Article 3, point (i), of Regulation (EU) 2022/2065.

Article 3.(i) of the DSA defines ‘online platform’ as:

“a hosting service that, at the request of a recipient of the service, stores and disseminates information to the public, unless that activity is a minor and purely ancillary feature of another service or a minor functionality of the principal service and, for objective and technical reasons, cannot be used without that other service, and the integration of the feature or functionality into the other service is not a means to circumvent the applicability of this Regulation”.

Recital 14 explains that:

“The concept of ‘dissemination to the public’, as used in this Regulation, should entail the making available of information to a potentially unlimited number of persons, meaning making the information easily accessible to recipients of the service in general without further action by the recipient of the service providing the information being required, irrespective of whether those persons actually access the information in question.

Accordingly, where access to information requires registration or admittance to a group of recipients of the service, that information should be considered to be disseminated to the public only where recipients of the service seeking to access the information are automatically registered or admitted without a human decision or selection of whom to grant access. Interpersonal communication services, as defined in Directive (EU) 2018/1972 of the European Parliament and of the Council ⁽²⁴⁾, such as emails or private messaging services, fall outside the scope of the definition of online platforms as they are used for interpersonal communication between a finite number of persons determined by the sender of the communication.

However, the obligations set out in this Regulation for providers of online platforms may apply to services that allow the making available of information to a potentially unlimited number of recipients, not determined by the sender of the communication, such as through public groups or



open channels. Information should be considered disseminated to the public within the meaning of this Regulation only where that dissemination occurs upon the direct request by the recipient of the service that provided the information.”

Scope in our previous Reports

Taking account of the above DSA definition and guidance, and that fact that Snapchatters are automatically registered without a human decision or selection of whom to grant access, for our previous Reports, Snap considered the Spotlight, Discover, Lenses, Public Profiles, Snap Map, and Advertising services of Snapchat to fall within the scope of risk assessment and mitigation obligations in Articles 34 and 35. These services entailed making information published by recipients of those services easily accessible to other recipients of Snapchat in general without further action by the recipients publishing the information in question.

Scope for this Report

For this Report, (i) the DSA definitions and guidance remain the same and (ii) the functionality of Snapchat has not significantly changed and Snapchatters are still automatically registered without a human decision or selection of whom to grant access. **Therefore, Snap still considers the Spotlight, Discover, Lenses, Public Profiles, Snap Map, and Advertising services of Snapchat to fall within the scope of risk assessment and mitigation obligations in Articles 34 and 35.** These services still entail making information published by recipients of those services easily accessible to other recipients of Snapchat in general without further action by the recipients publishing the information in question.

When we refer to “Snapchat” or “Snapchat’s in-scope services” in this Report, therefore, we are referring to those six services in Snapchat unless the context is clear that it is referring to Snapchat as a whole.

As explained to the Commission in our response to its Gen AI RFI, we continue to assess My AI to be out of scope of Snapchat’s designation as it concerned content creation rather than dissemination and was not system related to Snapchat’s inscope services. In particular, as confirmed to the Commission in subsequent discussions, Snap does not use My AI data relating to European Union users for advertising within Snapchat’s inscope services. Snap continues to consider My AI and other similar generative AI tools made available by Snap to be out of scope of Snapchat’s designation, except where identified otherwise in this Report.

The six in scope services of Snapchat are described in more detail in the following sections.



2.2 Spotlight

What is Spotlight?

Spotlight is the Snapchat community's destination for entertaining short-form video content. Launched in November 2020, Spotlight provides users a simple way to view short-form videos created and submitted by the Snapchat community via a personalized feed. All users can post videos to Spotlight either via the Snapchat app or on the website, and videos on Spotlight are public and visible to users on the Snapchat app, on the web, and a link to the Spotlight video can be shared to other platforms. Users can also add Comments to Spotlight videos, which go through moderation before being shown to the creator to either accept or reject, or auto-approve. If accepted or auto-approved, the Comment is publicly visible on the Spotlight video. Spotlight Comments may be deleted or reported, and viewers can also indicate fondness by clicking on a heart icon.



In addition to compliance with Spotlight Terms, users must also comply with the [Community Guidelines](#) and the [Spotlight Guidelines](#).

How does Spotlight work?

Spotlight provides a content experience that is intended to entertain and delight users in the same app they use to communicate with their friends and family. It offers creators at all stages of their career a variety of opportunities and tools to help them grow their audiences, build sustainable businesses and make content creation a full-time career. Spotlight is an easy entry point to start your creator journey and is a source of relevant cultural trends and credible partner to the industry (media, music, sports, fashion, etc.) that offers meaningful reach, relevance and revenue.

The content shown in Spotlight is personalized to provide the user with a more relevant experience. Spotlight's ranking algorithm is described [here](#). Users may opt out of personalization as described [here](#). Spotlight content is moderated using a combination of auto-moderation and human moderation, and all Spotlight content is human moderated before being widely distributed. Spotlight also uses various engagement and metadata to determine eligibility to receive revenue from their content.

Snapchat+ Features relevant to Spotlight

None. As explained in Section 1 (Snapchat 1.01), we decided not to proceed with our plans to simplify the application from 5 to 3 tabs (known as Simple Snapchat).



Gen AI Features being used by Spotlight

None. Snap recognises that content created by generative AI tools (whether on third party platforms or Snapchat) could be disseminated via Spotlight. Snap has taken this into account in this Report, including explaining how this activity impacts our risk assessment and what measures we have taken to address risks relating to the dissemination of content created by generative AI.

2.3 Discover

What is Discover?

Discover is part of the 4th tab in the Snapchat app, below your friends' Stories. Note that this product has at times been known as "For You" and any reference to For You that remains in this Report or previous Reports should be interpreted as a reference to Discover.

Discover is dedicated to Creator Stories, which includes Media Partner content, and some user generated content ("UGC") created from Snaps by popular users ("Creator Content"). The UGC that appears on Discover includes the Public Stories from Snap Stars and other users who meet a follower count threshold. The videos in Discover are accessible to all users including those between 13-17 years old.

How does Discover work?

Discover displays personalized content to users. Discover achieves this using its ranking algorithm, which is described [here](#). The intended purpose of this processing is to personalize Discover and make it easy for users to discover new content that is relevant to their interests. The intended effect/impact on users is that they enjoy what they are watching and remain engaged users of Snapchat. Users may opt out of personalization as described [here](#).

Discover content is moderated using a combination of auto-moderation and human moderation, and all Discover content is human moderated before being widely distributed.

Discover also generates information about how Snapchatters interact with the content in Discover. It achieves this by generating 'event' metadata each time a user does something noteworthy, like viewing or skipping a video. The intended purpose of this processing is to select content the user is likely to be interested in, in order to further personalize content on Discover and elsewhere in Snapchat (such as other content areas like Spotlight and also Advertising - the revenue from which is used to pay for Snapchat). The intended effect/impact on users is that they enjoy their experience and remain engaged users of Snapchat.

Snapchat+ Features relevant to Discover

None. As explained in Section 1 (Snapchat 1.01), we decided not to proceed with our plans to simplify the application from 5 to 3 tabs (known as Simple Snapchat).

Gen AI Features being used by Discover



None. Snap recognises that content created by generative AI tools (whether on third party platforms or Snapchat) could be disseminated via Discover. Snap has taken this into account in this Report, including explaining how this activity impacts our risk assessment and what measures we have taken to address risks relating to the dissemination of content created by generative AI.

2.4 Public Profiles

What are Public Profiles?

Public Profiles enable Snapchatters to be discovered and followed in the app and showcase their favorite public Snaps, Lenses and other information. Snapchatters (including businesses) can create and access Public Profiles and grow an audience with their public identity. Public Profiles enables Snapchatters to showcase Stories, Spotlights and Lenses. For more information, see [here](#).



How do Public Profiles work?

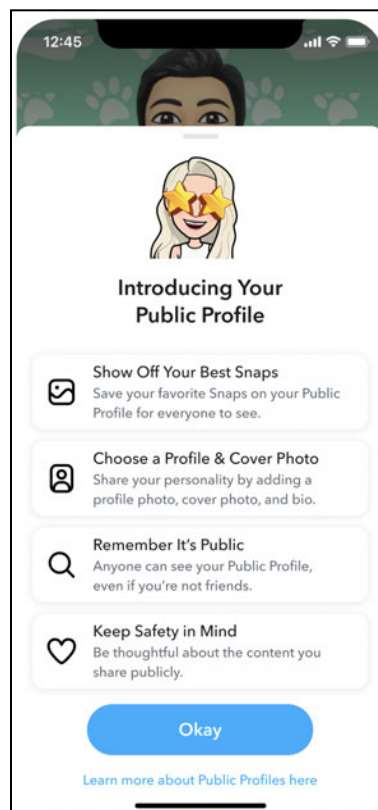


Snapchatter accounts aged 16 and over can choose to use a Public Profile to share a bit more about themselves with a wider audience (beyond their immediate friends) under their Public Profile. To create a public profile an eligible Snapchatter is required to: (i) tap their Bitmoji or Story icon at the top to go to My Profile; (ii) Scroll down to the 'Public Profile' section and Tap 'Create Public Profile' and (iii) then follow the simple instructions to create their Public Profile. We provide a dedicated guide for teen accounts that is displayed before they can create a Public Profile.

A Snapchatter can then choose to add information to their Public Profile, including a Photo, Bio, Description. Users can choose to display their Spotlight and Public Story content on their public profile. Users can also be followed by other Snapchatters. Lenses published by users can also be displayed on their Public Profile. They will be able to show their Follower Count and view Public Story, Lens, and Audience Insights.

Snapchatters with a Public Profile that are particularly active can have their accounts upgraded to a Creator Account. These have advanced features that are designed to enable professional Creators to connect and grow with their audience. Creator Accounts are eligible to have their content shown in the Discover section of Snapchat.

When older teenagers (16-17 accounts) first interact with their Public Profile page, post to Spotlight or Snap Map, or share a Public Story, they are shown a dedicated notice explaining what Public Profiles are and how to use them appropriately. This notice links to the [support pages](#) providing more information on Public Profiles.





As explained in Section 1 (Snapchat 101), before the launch of public profiles for older teens, Snap conducted a risk assessment and added mitigations. For example, Public Stories from 16–17 year olds are only shown to friends, followers, or mutual connections - not to the wider community through Discover. Teens can receive Story replies but cannot start direct chats from them, and replies are moderated more strictly for this age group. They can also turn off replies or block certain terms, and all users retain easy reporting tools.

Data from outside the EU, where the feature was already available, suggests these measures are effective. All changes to Public Profiles undergo moderation, and only 0.01% of profiles failed checks. For Public Stories, just 0.5% were rejected and 6% were subject to restrictions. Moreover, internal analysis shows no significant increase in contact risks for 16–17 year olds. They receive roughly the same share of adult friend requests, show no rise in chat reports, and are not seeing predatory patterns such as aggressive adult adds. Their connections' median age and friendship duration are consistent with those of peers without Public Profiles.

Consequently, Snap concludes that the residual risk of this change is low, with safeguards proving effective in protecting older teenagers while allowing them access to public posting features.

Snapchat+ Features relevant to Public Profiles

None

Gen AI Features being used by Public Profiles

None

2.5 Snap Map

What is Snap Map?

Snap Map is designed to open up a world of possibilities for our community, enabling friends to experience something new in the world every day. Through an interactive map interface, Snap Map shows users what's happening nearby and around the world, anchored by the context of friends' Bitmojis. It's a personal map that starts with the user at the center and reflects the people, places, and activities they care about, and helps users meet up with friends, express themselves, find things to do, and explore places elsewhere. Snap Map has been developed with the privacy and safety of our community of Snapchatters in mind.

How does Snap Map work?

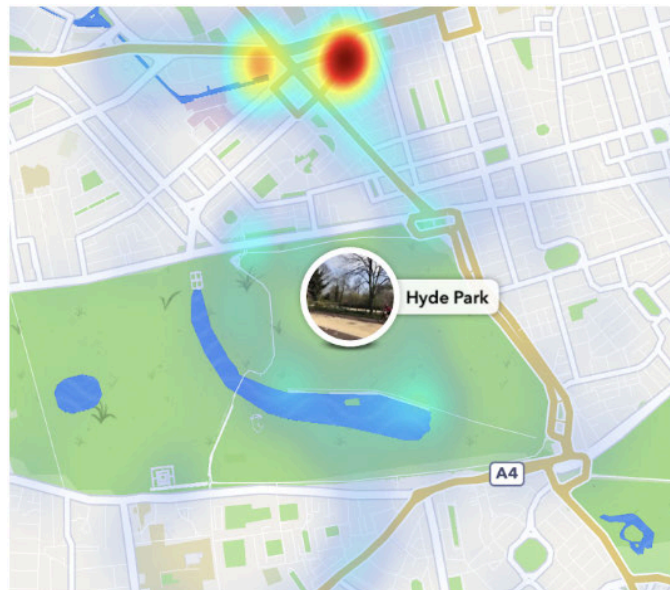


Snapchatters can share their Snaps to the Map by selecting “Snap Map” on the “Send To” page. If the Snapchatter has a Public Profile or is sharing their My Story with everyone, they may also have their Snap shared on Snap Map when it's tagged to a place or venue. Snapchatters can also choose to share their location on the Map with friends while the Snapchat app is actively being used, or share their live location with them even when the app is backgrounded.

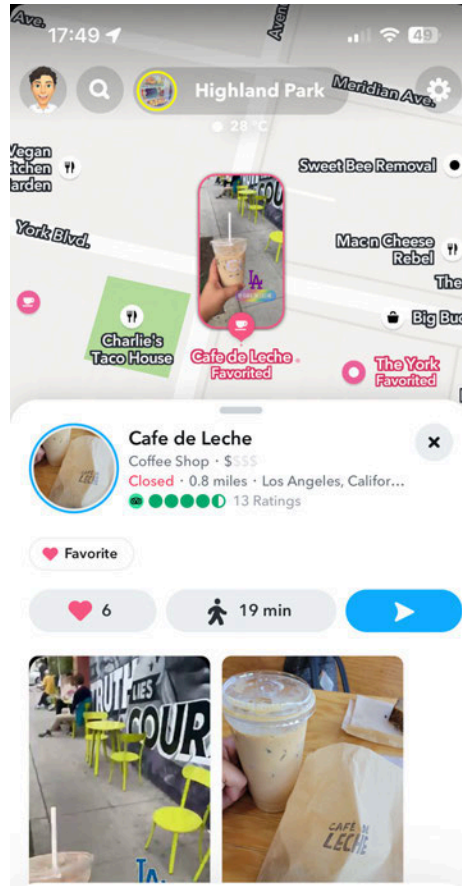
Location sharing is with Friends, rather than the public, and out of scope of Snapchat’s designation and this Report. Nevertheless, we have outlined safeguards we have implemented for location sharing in the Mitigations section below (Sections 5.1 and 5.3 specifically).

Snap Map features five types of user-generated content that can be served:

1. **Map Stories** include thumbnails on the map that highlight interesting events and popular places on the Map.



2. **Place Stories** appear on Place profiles. They contain Public Stories snaps explicitly tagged with the place, using either venue filters or place stickers.

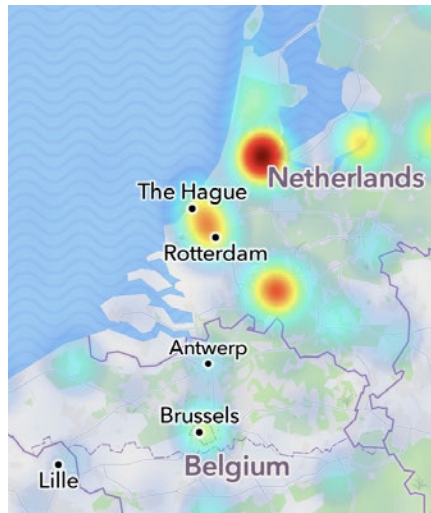


3. **City Stories** appear in the header of the Map and display the best snaps in that locality from the last 7 days. They can appear in cities and neighborhoods.

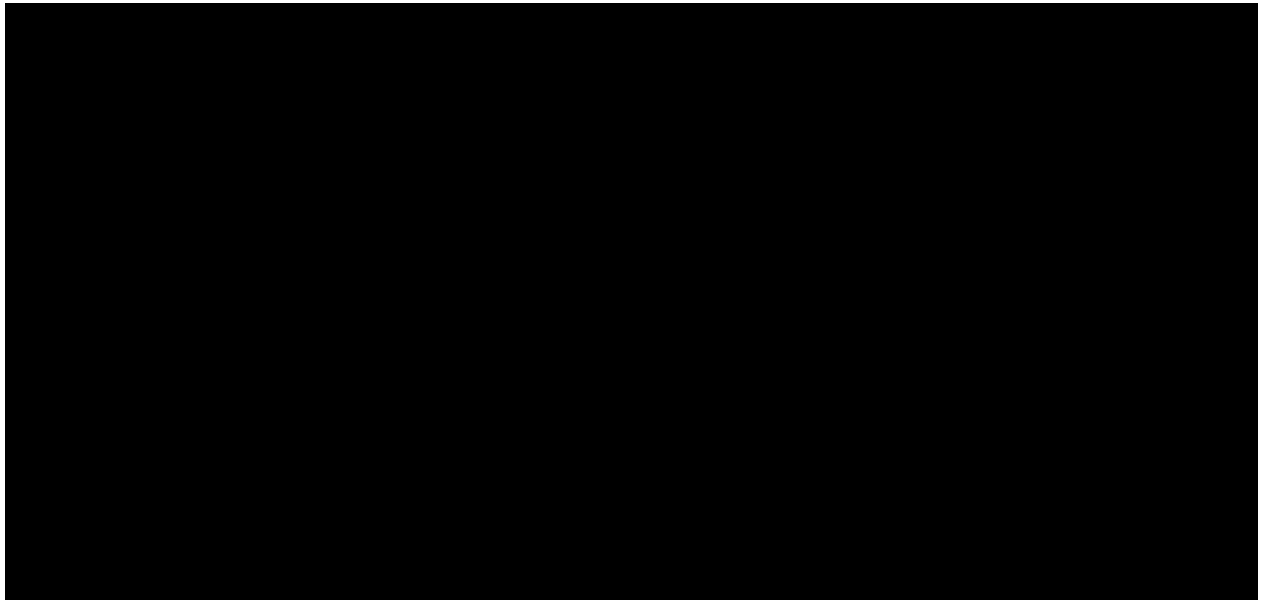




4. The **Heat Map** is used to visualize the volume and recency of content that's submitted to Public Stories. Content up to seven days old can be accessed by the heatmap. Heat spots represent areas where there is recent, high volume content.



5. **Friend Stories tagged with Places** presents a view of snaps that have been tagged with Places by a user's friend, along with the Bitmoji of the friend, that would appear with the Place on the basemap. This helps to personalize places on the user's Map, highlighting the places friends have recently visited.

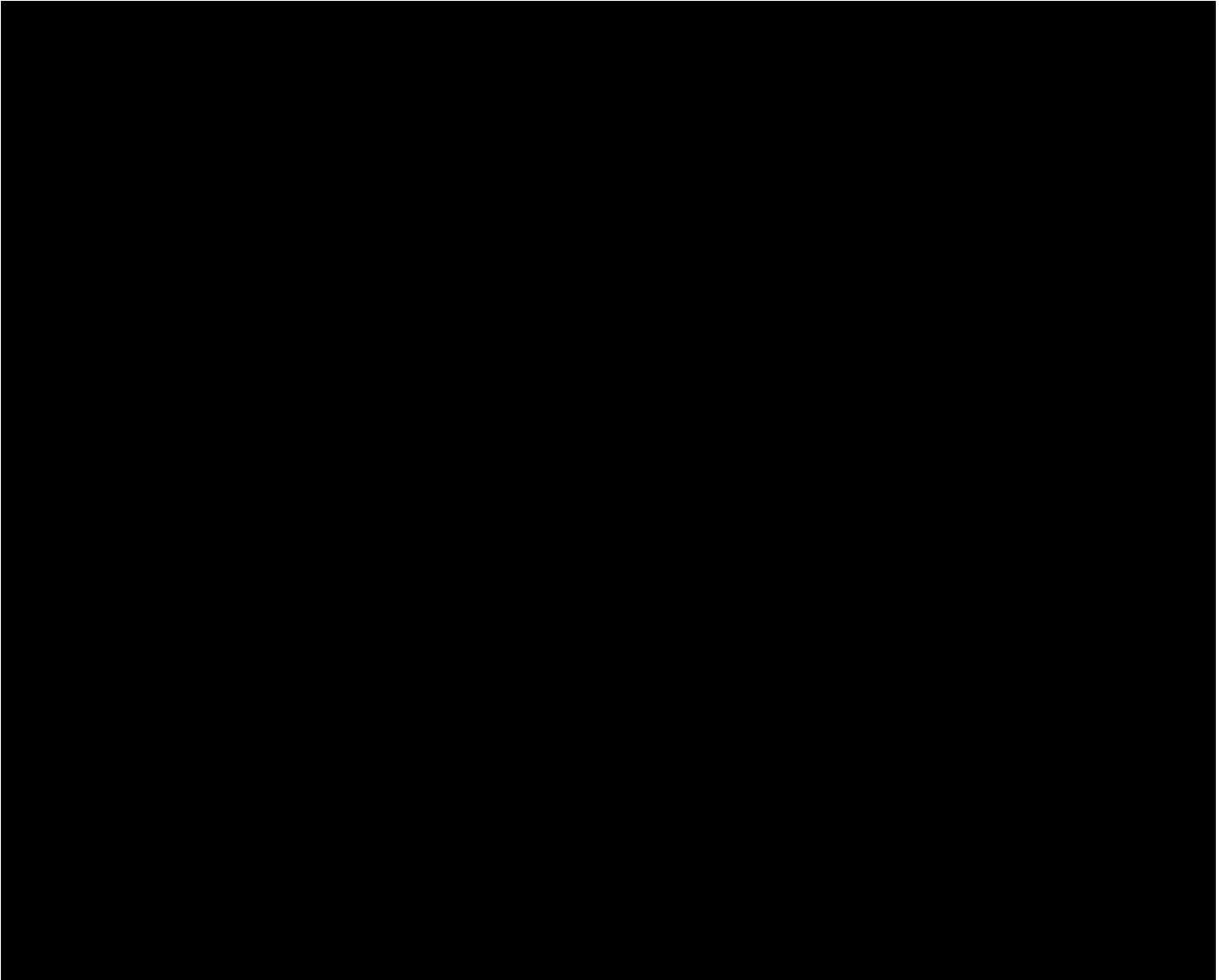


Snap Map submissions may be stored for a while and may be visible on Snapchat for long periods of time as explained to Snapchatters in privacy notices. Snapchatters can remove a Snap they submitted to Snap Map or place-tagged in Spotlight at any time via their profile.

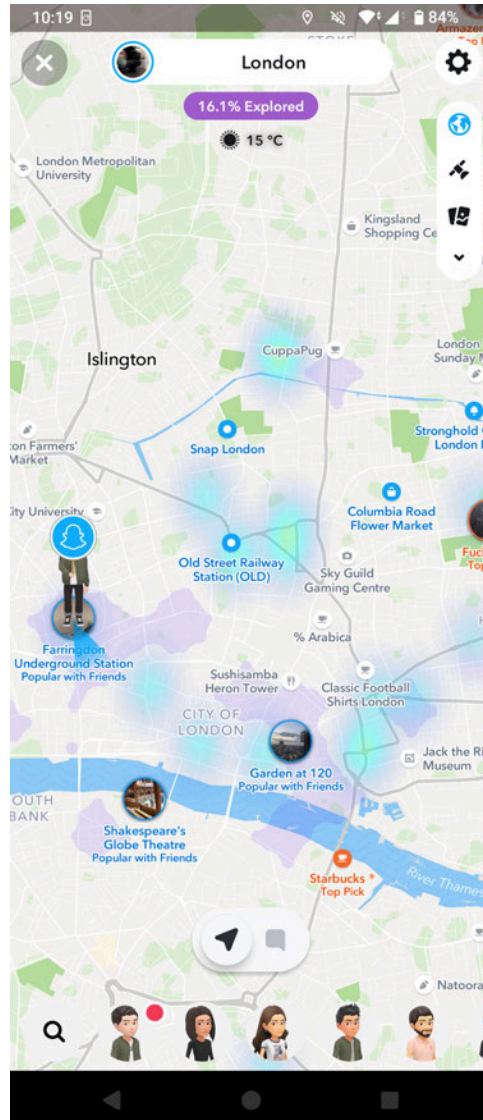


Snapchat+ Features relevant to Snap Map

Customizations - Snap Map offers a couple of customization features for Snapchat+ users including the ability to customize your home on the Map and adding your pet on the Map (either via a preselected group of pets or by creating a generative AI version of your pet).



Footsteps - Snap Map also lets SC+ users see how much of the world they've explored. This content is only made available to the user, and we have assessed it to be out of scope of Snapchat's designation.



Gen AI Features being used by Snap Map

Gen AI Pet - Snap Map offers a Gen AI pet feature as described above. Note however that while this feature is integrated into the Snap Map, the content is only available to Friends with whom the user chooses to share their location. It does not disseminate content to the broader public, and we have assessed it to be out of scope of Snapchat's designation. Nevertheless we have outlined our approach to both Gen AI creation and dissemination in the Mitigations section (in particular [Section 5.9](#) (Content Authenticity)).

2.6 Lenses

What are Lenses?

Snapchat Lenses are [augmented reality \("AR"\)](#) experiences designed to transform the way users look and the world around them. Snapchatters frequently use Lenses for entertainment purposes,



for example by creating Snaps with added 3D effects, digital objects, characters, and transformations to their image and voice. For example, Lenses can be used to add a layer of make-up to the user's face, to distort the user's face, to add a different background or certain elements to the surroundings. The most popular Lenses at the moment can be found [here](#). Snapchatters can interact with Lenses in the Carousel, via Search, and via Lens Explorer. In addition, we offer advertisers the possibility of creating [Sponsored Lenses](#).

How do Lenses work?

Lenses (in popular language often dubbed as 'filters') are created by a relatively limited number of community developers, and Snap's internal Lens Team. The transformational effects of Lenses are often accomplished through object detection, which is an algorithm designed to help a computer generally understand what objects are in an image. For example, it lets us know that a nose is a nose or an eye is an eye. There are numerous AR development tools Snap has made publicly available through Lens Studio, Snaps' Lens development platform and there are also internal tools that only the Lens Team can use to develop Lenses. Snap's AR development tools are reviewed by privacy engineering and legal before being used in Lenses. Some examples of AR development tools are object detection, text to speech, location landmarks and ML models and algorithms to support AR effects like tools for depth and context understanding, all designed to help a computer generally understand what objects are in an image.

We provide provided further information about how Lenses works in product specific support pages:

- [How to use lenses](#)
- [Create Your Own Filters & Lenses • Snapchat](#)

Snapchatters can create or develop Lenses in the desktop application 'Lens Studio'. There is a Public version and an internal Snap version of Lens Studio. Lens developers may publish Lenses through 'My Lenses', a web based portal. Lenses built by Snap's Lens Team are organic Lenses.

Snapchat+ Features relevant to Lenses

Certain Lenses are only available to Snapchat+ subscribers.

Gen AI Features being used by Lenses

Lens Studio features a GenAI Suite which lets developers take advantage of our generative AI technology to create assets (such as text, effects and backgrounds) for Lenses. Note however that Lens studio is only a creation tool. It does not disseminate content to the broader public, and we have assessed it to be out of scope of Snapchat's designation. Nevertheless we have outlined our approach to both Gen AI creation and dissemination in the Mitigations section (in particular Section 5.9 (Content Authenticity)).



2.7 Advertising

What is Snap's Advertising product?

Snap relies on online advertising to support its business. Snap has digital ad products created for advertisers who would like to easily create and manage ads that target relevant audiences on Snapchat ("Snapchat Ads Manager"). We process user information about Snapchatters to serve them with ads within Snapchat that we think they might be interested in.

An overview of Snap's ads services can be found [here](#) and [here](#). Some of Snap's advertising tools allow advertisers to provide Snap with data about their customers to improve their advertising campaigns. These tools are explained here:

- [Custom List Audiences](#)
- [Snap Pixel](#)
- [Conversion API](#)
- [Advanced](#) and [Estimated](#) Conversion

In addition, we offer advertisers the possibility of creating [Sponsored Lenses](#).

How does Advertising Work?

Our ad ranking algorithm determines which ads are displayed to a Snapchatter who is in the selected audience for those ads. The ad ranking algorithm uses various signals, including prior ad interactions and social signals, to determine which ads that user is more likely to interact with and then combines this with the results of advertiser ad action for that Snapchatter, to select an ad to display. Snap analyzes prior ad interactions to target advertisements. For example, we may determine that a user is likely to swipe up on certain types of ads or download certain types of games when they see an ad on Snapchat. We may then use this information to show that user similar ads.

Snapchatter interactions with the ad (i.e. impression data) is then logged to (a) attribute impressions to conversion events (such as a purchase on an advertiser website or download of an advertiser app) to demonstrate the performance of the ad and (b) to further train the ad ranking algorithm.

Snapchat+ Features relevant to Advertising

None

Gen AI Features being used by Advertising

Advertisers can take advantage of generative AI tools during ad creation (such as text translations and background image creation). These are in systems directly related to the dissemination of content (advertising) to the public on Snapchat (although they are common tools and minor in nature). We have taken the use of generative AI tools into account in this Report (in particular Section 4.3.1 (Democracy/Elections) and Section 5.9 (Content Authenticity)).



3. DSA Risk Assessment Methodology

Update: There have been no major changes to our risk assessment methodology since our 2023 and 2024 Reports. In this Report, Snap has enhanced its approach by incorporating additional internal evidence, specifically ‘violative view rate’ data, and more explicitly specifying how Snap has taken into account the specific risk factors references in Article 34 of the DSA.

In order to meet its obligations under Articles 34 and 35 of the DSA, Snap has applied a standard risk methodology adapted from that commonly used to assess risks in other contexts, including the EU general risk assessment methodology for product safety⁷ and the [ICO’s DPIA guidelines](#). Ofcom’s guidance for completion of the illegal harms risk assessment and protection of children risk assessments under the UK Online Safety Act proposed a similar methodology in its risk assessment guidance documents.⁸

The risk assessment methodology used by Snap has several steps:

3.1 Identification of Risks

As a first step, Snap identified potential systemic risks for each of the four categories outlined in the DSA:

- a. **Category 1 (Article 34.1.(a) / DSA Recital 80):** Dissemination of illegal or violating content, particularly rapidly and widely or as a result of intentional / automated manipulation, including:
 - i. Child sexual abuse material
 - ii. Illegal hate speech
 - iii. Criminal offenses and the conduct of illegal activities, such as the sale of prohibited products or services, dangerous or counterfeit products, or illegally-traded animals.
- b. **Category 2 (Article 34.1.(b) / DSA Recital 81):** Impact on fundamental EU rights, including in particular rights for:
 - i. Human dignity
 - ii. Freedom of expression and of information, including media freedom and pluralism
 - iii. Private life
 - iv. Data protection
 - v. Non-discrimination
 - vi. Children
 - vii. Consumer protection
- c. **Category 3 (Article 34.1.(a) / DSA Recital 82):** Negative effects on:
 - i. Democratic and electoral processes

⁷ EU general risk assessment methodology (Action 5 of Multi-Annual Action Plan for the surveillance of products in the EU (COM(2013)76)), [url](#).

⁸ Illegal Harms Risk Assessment Guidance ([url](#)) and Protection of Children Risk Assessment Guidance ([url](#)), Ofcom.



- ii. Civic discourse
- iii. Public security
- d. **Category 4 (Article 34.1.(a) / DSA Recital 83):** Negative effects, in particular from design and use/misuse such as a coordinated disinformation campaign, on:
 - i. Public health
 - ii. Gender-based violence
 - iii. Minors
 - iv. Physical and mental well-being (including addictions)

3.2 Likelihood Analysis

As a second step, Snap analyzed the extent to which the identified risk(s) are likely to occur on Snapchat. **In practice the prevalence of almost all of Snapchat's risks are considered to be very low, in part because of robust mitigations and the inherent design of relevant Snapchat functionality**, so Snap used a measure of relative likelihood between each risk on Snapchat so we can continue to prioritize and improve (as explained in the following table). Note: this is not measuring likelihood relative to other platforms; it is measuring likelihood relative to risks assessed by Snap.

With this in mind, Snap used three levels of relative likelihood:

<i>Relative likelihood of risk occurring on Snapchat</i>	<i>Description</i>
Low Likelihood	This means this risk has the highest chance of occurring on Snapchat vs other risks: • where Prevalence Testing data is available, this risk has a percent of policy-violating prevalence (PVP) of 0.5% or greater; and • where VVR data is available, the violative view rate is 0.005 or greater.
Very Low Likelihood	This means this risk has an average chance of occurring on Snapchat vs other risks: • where Prevalence Testing data is available, this risk has a percent of policy-violating prevalence (PVP) of between 0.05% and 0.49%; and • where VVR data is available, the violative view rate is between 0.005 and 0.0049.
Extremely Low Likelihood	This means this risk has the lowest chance of occurring on Snapchat vs other risks: • where Prevalence Testing data is available, this risk has a percent of policy-violating prevalence (PVP) of 0.049% or less; and • where VVR data is available, the violative view rate is 0.00049 or less



In order to assess likelihood, Snap uses a mix of internal information (such as [Prevalence Testing](#) data or illegal / violating content reporting data or complaint data, input from our safety advisory board and Snap commissioned research as well as VVR data) and external information (such as external research, news reports and government and NGO guidance). Where internal information is required, this was obtained from the relevant teams responsible for maintaining that information (for example, Customer & Review Operations with respect to complaint data).

3.3 Severity Analysis

As a third step, Snap analyzed the severity of the identified risk(s) by considering evidence of the potential harm they have **caused individuals or society in general. In practice the severity of all the identified risks could cause at least significant harm (which is why they have been identified). So we used a measure of relative severity between each risk so it can continue to prioritize and improve.**

With this in mind, Snap used three levels of severity:

<i>Harm classification industry wide</i>	<i>Description</i>
Severe harm industry wide	This means this risk has the highest severity vs other risks. We consider severe harm to include both (1) harms that risk significant damage to the physical or emotional well-being of Snapchatters and society at large e.g. external parties influenced by (other people's use of) Snapchat, and (2) the imminent, credible risk of severe harm, including threats to human life, safety, and well-being.
Serious harm industry wide	This risk has a medium level of severity vs other risks. We consider these risks not to be severe (as defined above) but still have the potential to cause serious harm.
Significant harm industry wide	This means this risk has the lowest severity vs other risks. While not the most severe or serious, these risks still have the potential to cause significant harm.

The safety of Snapchatters is our top priority. We take behavior that threatens the safety of our community very seriously. We collaborate with experts, safety groups, and law enforcement on these topics in order to better educate ourselves and our community, and to ensure we are sufficiently informed to analyze different levels of severity for each risk.



3.4 DSA Risk Factors

Snap has also taken into account whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of each harm on Snapchat's in-scope services:

Service Risk Factor
(a) the design of recommender systems and any other relevant algorithmic system;
(b) content moderation systems;
(c) the applicable terms and conditions and enforcement;
(d) systems for selecting and presenting advertisements; and
(e) our data related practices.

We have also analysed whether and how the risk of each harm category is influenced by the following general factors:

General Risk Factor
Intentional manipulation, including inauthentic use or inauthentic use or automated exploitation of the service.
Amplification and potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.
Specific regional or linguistic aspects, including when specific to a Member State.

3.5 Overall Potential Risk Prioritization Assessment

As a fourth step, Snap confirmed an overall potential risk prioritization for each identified risk taking account of the likelihood and severity analysis outlined above. This prioritization helps us to assess whether the mitigations we have put in place (as described in Snap's Mitigations) are proportionate, reasonable and effective as required by Article 35. As a guide we use the following matrix that is commonly used in risk assessment methodologies to determine the overall potential risk. However, this is only an approximation and we make a decision on the overall potential risk, and therefore the prioritization, of a particular issue on a case by case basis. As a result, there are instances where we deviated from the overall potential risk prioritization matrix below.

Overall Potential Risk Prioritization Matrix



Harm classification on industry wide	Severe harm industry wide	Level 3	Level 1	Level 1
	Serious harm industry wide	Level 3	Level 2	Level 1
	Significant harm industry wide	Level 3	Level 3	Level 2
		Extremely low	Very low	Low
Relative likelihood of risk occurring on Snapchat				

3.6 Snap's Mitigations

As a fifth step, Snap considered the mitigation measures that it has taken to address each of the risks identified in the overall potential risk prioritization assessment. When considering these mitigations, Snap has taken into account in particular the list of possible mitigations set out in Article 35.1. For ease of reference, we have set out a table below that maps the Article 35.1 list of mitigations to the corresponding section of this report where Snap has explained how it is using that mitigation measure on Snapchat.

#	DSA Mitigation	Relevant Report Section
a	Adapting the design, features or functioning of their services, including their online interfaces.	Snapchat Design and Function
b	Adapting their terms and conditions and their enforcement.	Terms and Enforcement
c	Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Moderation
d	Testing and adapting their algorithmic systems, including their recommender systems.	Algorithmic Systems
e	Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Advertising Systems



f	Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Ongoing Risk Detection and Management
g	Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Trusted Flaggers
h	Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Codes and Crisis Protocols
i	Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Transparency
j	Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate;	Protection of Minors
k	Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	Content Authenticity

3.7 Conclusions

As a final step, Snap confirmed whether the mitigation measures it has taken were reasonable, proportionate and effective for each risk identified. To determine this, Snap considered if the mitigations it has in place were effective to address the risk, given its overall potential risk prioritization category, by considering available evidence from its [Prevalence Testing](#) data, VVR data or illegal / violating content reporting data or complaint data, input from our safety advisory board and Snap commissioned research) and external information (such as external research, news reports and government and NGO guidance).

Where there was evidence that the existing measures risks may need some improvement to ensure reasonable, proportionate and effective measures had been taken, Snap identified this in its conclusion and explained what steps it would be taking to achieve the improvement. [REDACTED]

[REDACTED]

[REDACTED]



3.8 Supporting Documentation

The data and documentation supporting the risk and mitigation assessment report is retained for a minimum of 3 years.



4. DSA Risk Assessment Results

In this Section of the Report, we explain the result of the risk assessment of Snapchat’s in-scope services that Snap has carried out pursuant to Article 34 of the DSA. This risk assessment was conducted in accordance with the scope and methodology explained in Section 1 of this Report. One general point to note is that these risks impact a wide range of individuals, including our Snapchatter community, victims of crime, the general public and the moderators that review the content on Snapchat. The results of this risk assessment apply to all such individuals, and where appropriate we have noted impacts that extend beyond Snapchat (including the wellness of our moderators).

It is Snap’s mission to reduce virtually all harmful content on our platform. To that end, we are continually improving our systems every single day, and are investing into (machine learning) technology, human moderation, and other measures to make our platform safer for our community. As described in the [Ongoing Risk Management](#) section below, Snap has reasonable, proportionate and effective measures to detect and manage risks on an ongoing basis.

4.1 Category 1 - Dissemination of content that is illegal or violates our terms and conditions

(Article 34.1.a / DSA Recital 80)

In this first part we report on our assessment of the risk of illegal content or content that is incompatible with our [Terms](#) being disseminated on Snapchat as required by Article 34.1.a (“Category 1”), including in particular the illegal content identified in Recital 80. In our assessment, we have taken account of the extent to which these risks are influenced by intentional manipulation, including by inauthentic use or exploitation of the service, as well as the extent to which Snapchat allows for amplification and potentially rapid and wide dissemination.

The table below provides a summary of the results of our assessment of likelihood, severity and overall potential risk prioritization, together with our conclusions given the mitigations that Snap has put in place for each Category 1 risk. *Note that for all harms, where there is (1) a risk of significant damage to the physical or emotional well-being of Snapchatters, and (2) imminent, credible risk of severe harm, including threats to human life, safety, and well-being, we treat this as a severe harm and an Level 1 overall risk prioritization.*



Category 1 - Dissemination of content that is illegal or violates our terms and conditions (including our Community Guidelines)				
Category	Relative likelihood of risk occurring on Snapchat	Harm classification industry wide	Risk Prioritization	Conclusion
4.1.1 Dissemination of Child Sexual Abuse Material	Extremely low Likelihood	Severe harm industry wide	Level 1	Low Risk / Reasonable, proportionate and effective mitigations
4.1.2 Dissemination of Illegal Hate Speech	Extremely low Likelihood	Significant harm industry wide	Level 3	Low Risk / Reasonable, proportionate and effective mitigations
4.1.3 Dissemination of information related to the Sale of Prohibited Products or Services (such as dangerous products, counterfeit products or illegally-traded animals)	Extremely low Likelihood	Severe harm industry wide (Drugs)	Level 1 (Drugs)	Low Risk / Reasonable, proportionate and effective mitigations
	Extremely low Likelihood	Serious harm industry wide (Weapons)	Level 2 (Weapons)	Low Risk / Reasonable, proportionate and effective mitigations
	Extremely low Likelihood	Significant harm industry wide (Other goods)	Level 3 (Other goods)	Low Risk / Reasonable, proportionate and effective mitigations
4.1.4 Dissemination of Terrorist Content	Extremely low Likelihood	Serious harm industry wide	Level 2	Low Risk / Reasonable, proportionate and effective mitigations, which are being closely monitored due to a further slight increase in prevalence.
4.1.5 Dissemination of Content that infringes on Intellectual Property Rights	Extremely low Likelihood	Significant harm industry wide	Level 3	Low Risk / Reasonable, proportionate and effective mitigations
4.1.6 Dissemination of Adult Sexual Content	Extremely low	Serious harm industry wide	Level 2 (Sexual crimes)	Low Risk / Reasonable,



	Likelihood	(Sexual crimes)		proportionate and effective mitigations
		Significant harm industry wide (Other Adult Sexual Content)	Level 3 (Other Adult Sexual Content)	Reasonable, proportionate and effective mitigations, which are being monitored to confirm prevalence continues to decline.
4.1.7 Dissemination of content regarding Harassment and Bullying	Extremely low Likelihood	Serious harm industry wide	Level 2	Low Risk / Reasonable, proportionate and effective mitigations.
4.1.8 Dissemination of content that glorifies Self-Harm, including the promotion of Self-Injury, Suicide or Eating Disorders	Extremely low Likelihood	Serious harm industry wide	Level 2	Low Risk / Reasonable, proportionate and effective mitigations
4.1.9 Dissemination of content encouraging or engaging in violent or dangerous behavior	Extremely low Likelihood	Significant harm industry wide	Level 3	Low Risk / Reasonable, proportionate and effective mitigations
4.1.10 Dissemination of Harmful False Misinformation	Extremely low Likelihood	Significant harm industry wide	Level 3	Low Risk / Reasonable, proportionate and effective mitigations
4.1.11 Dissemination of Fraud and Spam	Extremely low Likelihood	Significant harm industry wide	Level 3	Low Risk / Reasonable, proportionate and effective mitigations.
4.1.12 Dissemination of information related to Other Illegal Activities	Extremely low Likelihood	Significant harm industry wide	Level 3	Low Risk / Reasonable, proportionate and effective mitigations

4.1.1 Dissemination of Child Sexual Abuse Material

This Section 4.1.1 (CSEA) considers the risk of harm arising from child sexual exploitation and abuse (as described in the Harm Description below) on Snapchat's in-scope services.

Harm Description

CSEA refers to content and activity involving taking advantage of a vulnerability, trust, or an imbalance of power to coerce, manipulate or deceive a minor (i.e. a person under the age of 18)



into sexual activity or any activity for the sexual gratification of another, online or in person, including grooming and dissemination of child sexual abuse material (“CSAM”):

- Grooming refers to building trust, establishing a relationship, and developing an emotional connection with a minor in order to manipulate, exploit, and/or abuse them. This includes any communication or behavior that attempts to persuade or coerce a minor with the intent of sexual abuse or exploitation, or which leverages fear or shame to keep a minor silent.
- CSAM includes still images, videos, and illustrated, computer-generated or other forms of realistic depictions, as well as live streaming broadcasts of a child in a sexually explicit context, or engaging in sexually explicit acts. CSEA may involve the generation, possession, distribution, promotion, or solicitation of such material, including both ‘known CSAM’ (previously hashed and catalogued by organisations like NCMEC), and ‘new CSAM’ (not previously hashed or catalogued), whether authentic or AI-generated. It may also involve the sharing or promotion of hyperlinks that direct users to CSAM hosted off-platform. These URLs may lead to:
 - Known CSAM-hosting sites or darknet marketplaces;
 - Cloud drives, encrypted messaging channels, or file dumps; or
 - Evasion techniques (e.g., obfuscated links, shorteners).
- CSEA may also include:
 - Accounts dedicated to sexually harassing, intimidating or extorting minors;
 - Teens or adults requesting/encouraging/enticing/coercing minors to engage in sexual acts or produce sexual images;
 - Content that facilitates a sex act with a minor;
 - Content that promotes the sale of commercial sex with a minor; and
 - Any threats to share, exploit, or expose a minor’s intimate images or videos without their consent (i.e., sextortion), whether by an adult or a fellow minor.

We provide a summary and explanation of ‘CSEA’ in our explainer on [Sexual Content](#) and our [Transparency Report Glossary](#).

Note that we also prohibit content that, while not reaching the level of CSEA, includes minors in an inappropriate manner. This may include images of naked minors absent any sexually explicit context or sexual suggestiveness, such as nude babies or toddlers playing in bathtubs or swimming pools or minors engaged in sexually suggestive behaviour.

Likelihood

We have considered the likelihood of this harm in line with Section 3 (Methodology) and observed the following for 2025:



- In order to better understand how we might combat this harm, [REDACTED] Snap began efforts to measure Policy Violating Prevalence (PVP) via random sampling of Public Stories to estimate the percent of policy-violating views.⁹
 - [REDACTED]
 - In our 2023 Report, we noted that CSEAI content represented an extremely low percentage of total views found to be violating, [REDACTED] We also noted that in the second half of 2022, the proactive moderation detected and actioned 94% of the total child sexual exploitation and abuse violations reported in our Transparency Reports.
 - By 30 July 2024 and as a result of our continued focus, the percentage of CSEAI violating views saw a further, substantial fall to an extremely low rate: [REDACTED] [REDACTED] The steps Snap took to mitigate this risk diminished the likelihood that Snapchatters would encounter CSEAI on Snapchat’s in-scope services, such that CSEAI has dropped out of the top 10 harm list entirely.
 - As of 30 April 2025, we have successfully maintained this extremely low prevalence rate on Snapchat’s in-scope services in 2025. [REDACTED]
- Where CSEA content was identified (either proactively via automated tools or reactively following a report), our median turnaround time was rapid. Our [latest European Union Transparency Report](#), which covers the second half of 2024, observed a median turnaround time for our Safety teams to take enforcement action in response to proactive or reactive detections of CSEA content of **322 minutes**¹⁰.
- As a result of this rapid turnaround time:
 - the violative view rate for EU users for CSEA content across the following Snapchat in-scope services is very low in the first half of 2025:

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

- the violative view rate for EU minors for CSEA content across the following Snapchat in-scope services is also very low in the first half of 2025:

⁹ See Section 6.4 (Prevalence Testing).

¹⁰ Our CSEA turnaround time is higher than in other policy areas because some CSEA content is subject to a specialized process that includes double-review with a select team of specially trained agents. This does not apply to detected CSEA content matching hashes of known, verified CSEA material, which will be actioned more swiftly.



- In February 2024 by prominent Finnish NGO Protect Children published a study on the use of tech platforms by online child sexual abuse offenders, funded by the Tech Coalition and Safe Online.¹¹ Snapchat ranks last among the social media platforms used to search, view, and share CSAM (10%), compared to Instagram (29%), Twitter/X (26%), Discord (23%), TikTok (21%), Facebook (20%), Youtube (18%), Reddit (17%).
- Although Snap is aware of concerns regarding the dissemination of CSEAI created using generative AI tools in the wider industry, Snap has not identified this as being a material issue on Snapchat’s in-scope services.

As a result, Snap continues to place CSEAI into the **lowest likelihood category**.

Severity

For this Report, as in our previous Reports, Snap still considers all CSEAI to have a risk of **severe harm**. This includes sexual images of minors aged 13-17 and grooming activities (e.g. enticing a Teen to produce sexual images). There is still no doubt that child sexual exploitation and abuse is one of the most serious crimes which violates Children’s Rights and has far-reaching and serious detrimental lifelong consequences for its victims.¹² As well as the obvious severe harm caused by CSEAI, the distribution of images online has been found to cause ongoing harm. A study led by the Canadian Centre for Child Protection in 2017¹³, noted that 67% of CSAM survivors said the distribution of their images impacts them differently than the hands-on abuse they suffered because the distribution never ends and the images are permanent. InHope notes¹⁴ that children who have been victimised and experienced grooming are likely to suffer from serious long-term mental health issues such as anxiety, depression, post-traumatic stress, and suicidal thoughts. Especially children who have been solicited into creating and sharing intimate content of themselves tend to experience shame and blame themselves for the abuse, regardless of whether the abuse took place offline or if all interaction between the perpetrator and child took place online. Children can be seriously impacted by grooming even when no personal contact has occurred, which is why it is crucial to intervene as soon as possible.

¹¹ Tech Platforms Used by Online Child Sexual Abuse Offenders, [url](#), February 2024.
¹² European Commission, *EU strategy for a more effective fight against child sexual abuse* (COM(2020)607 final), [url](#), 24 July 2020; Council of Europe, *European Day on the Protection of Children against Sexual Exploitation and Sexual Abuse*, [url](#), updated October 2016.
¹³ https://protectchildren.ca/pdfs/C3P_SurvivorsSurveyExecutiveSummary2017_en.pdf.
¹⁴ InHope, Grooming, as at June 2025, [URL](#).



Figures released by the Internet Watch Foundation (IWF) show that 91% of assessed reports of CSAM in 2024 were self-generated—often involving children aged 7–13, sometimes as young as 3–10. Between 2020–21, instances for ages 7–10 soared from 8,000 to 27,000—a 235% jump. Further research¹⁵ from the IWF also reveals 14% of UK young adults (18–24) accidentally encountered CSAM, while the 25–34 age group reported 10%. Figures released by the IWF also show that Europe remains a global hub for CSEA. The percentage of child sexual abuse reports were traced to hosting services in Europe (including Russia and Turkey) has remained consistently high (59% of global cases in 2022¹⁶, 64% in 2023¹⁷ and 62% in 2024.¹⁸ The IWF recorded year-on-year increases in both the number of URLs and direct reports which contain CSAM, with a 6% increase between 2023 and 2024. Unfortunately, this shows that the severe harm caused by CSEAI is still growing.

According to updated figures from NCMEC¹⁹, it received reports relating to 29.2 million incidents of suspected child sexual exploitation via its CyberTipline in 2024. While this is a significant number, the overall number of reports declined from 36.2 million reports received in 2023. NCMEC also noted in its report two trends which have continued to rise since its 2023 report: (i) a continued increase in online enticement reports, fueled in part by the crime of sextortion with NCMEC receiving nearly 100 reports of financial sextortion a day; and (ii) an increase in reports relating to generative AI, from 4700 in 2023 to 67,000 in 2024. While these are still relatively small compared to other categories of CSEAI, these are trends that Snap is very conscious of and is monitoring.

DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat’s in-scope services. As set out below, while the following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap’s existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Our recommender systems are designed not to knowingly recommend CSEA – i.e., there are no ‘CSEA’ interest categories. Detailed information regarding the mitigations relating to our recommender systems can be found in Section 5 (see the Algorithmic Systems sub-section).

¹⁵ IWF, More than one in 10 British young people exposed to online child sexual abuse, [url](#).
¹⁶ Internet Watch Foundation (IWF), *Annual report 2022*, [url](#), 2022.
¹⁷ Internet Watch Foundation (IWF), *Annual report 2023*, [url](#), 2023.
¹⁸ Internet Watch Foundation (IWF), *Annual report 2024*, [url](#), 2024
¹⁹ NCMEC, CyberTipline Report 2024, <https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata>



(b) content moderation systems;	Snap has implemented specific proactive and reactive moderation procedures to prevent and remove CSEA. As explained in the Content Moderation Section (specifically, the section on CSEA), Snap deploys a range of automated content moderation (which includes abusive language detection, other keyword-based detection, and machine-learning-based proactive detection) to scan media uploads for CSEA, including: (i) PhotoDNA, CSAI Match, and other hash-based detection (including NCMEC's Take It Down hashes) to detect known CSAM; (ii) Google's Content Safety API to detect novel CSAM and (iii) proprietary signal based tools to detect sextortion and other sexual harms against minors. We provide more information in Section 5 (see Content Moderation sub-section)
(c) the applicable terms and conditions and enforcement;	Our Community Guidelines (which form part of our terms of service) prohibit CSEA for all users (both adult and Teen accounts) and they are strictly enforced given the risk of severe harm. We provide more information in Section 5 (see the Terms and Enforcement sub-sections)
(d) systems for selecting and presenting advertisements; and	Our advertising systems require agreement to advertising policies and guidance that prohibit adverts from displaying information that violates the law or causes certain harms. We check advertisers are complying with their obligations via our advertising review process. Our advertising systems use a mix of automation and human review to review adverts before they are published. We provide easy mechanisms for users to hide and report advertisements that violate our policies or the law. We monitor ad rejection, reporting and enforcement data to monitor the effectiveness of our approach. We provide more information in Section 5 (see the Advertising Systems sub-section)
(e) our data related practices	We have strong data principles, practices and privacy, safety and security by design processes. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).

We have also analysed whether and how the risk of CSEA is influenced by the following general factors:

General Risk Factor	How does it apply to Snapchat and this harm?
---------------------	--



[REDACTED]

[REDACTED]

[REDACTED]

Amplification and potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.

[illegible]

Snapchat’s inscope services have a number of features and design configurations that act to limit the amplification and potentially rapid and wide dissemination of CSEA content, in particular:

1. Snapchat is not an attractive platform for spreading CSEA, in particular because it is difficult to reach a broad audience and content is deleted by default. We provide more information in Section 5 (see Snapchat Design / Function subsection)
2. Snap has implemented specific proactive and reactive moderation procedures to prevent and remove CSEA. As explained in the Content Moderation Section (specifically, the section on CSEA), Snap deploys a range of automated content moderation (which includes abusive language detection, other keyword-based detection, and machine-learning-based proactive detection) to scan media uploads for CSEA, including: (i) PhotoDNA, CSAI Match, and other hash-based detection (including NCMEC's Take It Down hashes) to detect known CSAM; (ii) Google's Content Safety API to detect novel CSAM and (iii) proprietary signal based tools to detect sextortion and other sexual harms against minors. We provide more information in Section 5 (see Content Moderation)
3. Content recommended to users on Spotlight and Discover, our video sharing platforms, is moderated using a combination of auto-moderation and human moderation, and is human moderated before being widely distributed. Lens and Ads are subject to review



	processes before submission. We provide more information in Section 5 (see Content Moderation)
Specific regional or linguistic aspects, including when specific to a Member State.	We recognise that our users may come from different Member States and content may be shared in different languages. To address this, • Snapchat is provided in multiple EU languages and our Terms are available in all EU languages. We provide more information in Section 5 (see Terms). • We deploy content moderation systems and moderators work across multiple EU languages. We provide more information in Section 5 (see Content Moderation).

Overall potential risk prioritization

Although the prevalence of CSEAI on Snapchat has continued to decline and is now at extremely low levels (and is considered to be at the lowest level of all our risks), due to the potential for the most severe harms to be caused by CSEAI and the continued growth in the prevalence of this issue online in general, Snap still considers CSEAI to be a **Level 1 risk priority**. There is no change in this assessment from our past assessments.

Snap's Mitigations

Snapchat enables photo and video sharing and other user interaction across Snapchat's inscope services. While this functionality has many positive benefits for users, including in particular facilitating engagement with friends and family, Snap recognises Snapchat's functionality could also be abused by bad actors for CSEA purposes. Snap has put in place significant measures to substantially diminish the likelihood and impact of CSEA on Snapchat.



In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services. These are organised into Snap's risk assessment mitigation categories. Note that the primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive. As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **[link in the left hand column to a full summary](#)** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, several aspects of Snapchat's design and function reduce the risk that teens will come into contact with strangers. For example, by default users need to accept bi-directional friend requests or already have each other in their contact book to start communicating directly with each other. This is an important mitigation to prevent strangers from contacting users on Snapchat. Friend lists remain private. Snapchat does not disclose the friend lists of users to other users, nor do we expose the total number of friends that a user has. There is also no option to share location with strangers.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes, our Community Guidelines (which form part of our terms of service) prohibit CSEA and they are strictly enforced given the risk of severe harm. When Snap becomes aware that CSEA is present on our platform, we remove the content from the platform or take enforcement action on the user account as appropriate. Upon identifying any CSEA on Snapchat, Snap also reports the content and the user's account to NCMEC.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and	Yes, Snap has implemented specific proactive and reactive moderation procedures to prevent and remove CSEA. As explained in the Content Moderation Section (specifically, the section on CSEA), Snap deploys a range of automated content moderation (which includes abusive language detection, other keyword-based detection, and machine-learning-based proactive detection) to scan media uploads for CSEA, including: (i) PhotoDNA, CSAI Match, and other hash-based detection (including NCMEC's Take It Down hashes) to detect known CSAM; (ii) Google's Content Safety API to detect novel CSAM and (iii) proprietary signal based tools to detect sextortion and other sexual harms against minors. We enforce against accounts found to be engaging in CSEA. Snapchatters can also report CSEA to us via in-app reporting options and anyone can



Mitigation Category	Applies to this risk?
dedicated resources for content moderation.	submit a report through the Snapchat Support Site.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, our algorithmic systems do not knowingly recommend CSEA – i.e., there are no ‘CSEA’ interest categories.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, for example we have specific prevalence testing and monitoring moderation and enforcement data which we use to help detect and manage CSEA-related risk.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Yes, we cooperate with trusted flaggers in relation to CSEA.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various groups – e.g. EUIF, the Technology Coalition, WeProtect Global Alliance, IWF (Report Remove). We are members of StopNCII.org in the UK and we use their hashes to help detect and remove NCII. Snap works with NCMEC and other safety experts to learn about harms relating to CSEA and how they may manifest themselves on our platform, and to report such harms to the proper authorities. Snap also has trusted flaggers to bring these and other types of



Mitigation Category	Applies to this risk?
	harms to the attention of our safety teams. There are also industry wide initiatives such as the Tech Coalition's Lantern Program which was launched on 10 November 2023. ²⁰
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on our Terms, harms, moderation and enforcement practices, and how to get help in our Safety Center . We also have in-app resources on sexual content which educates and empowers users about these harms (including CSEA) and how to report these in the UK, including law enforcement and Report Remove. We have provided significant support resources for our users concerning financial sextortion ²¹ .
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, we have protective measures to limit teen contact with strangers, as well as provide warnings to teens when they first start chatting with someone who may be a stranger; we offer Family Center where parents can see who their teen is friends with, speaking to and report any suspicious accounts for review; we make available robust reporting. Our new parents site provides additional guidance for parents and caregivers on risks and support. ²²
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	We recognise there is growing concern regarding use of generative AI tools for CSEA in general online. Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services. Snap has also requested that NCMEC share its hash list of AI generated CSEA content. While we are alive to the risk, we have not identified any material issue with Snapchat being used to create or disseminate generative CSEA materials.

²⁰ Tech Coalition, Announcing Lantern: The First Child Safety Cross-Platform Signal Sharing Program, url, 2023.

²¹ <https://values.snap.com/safety/financial-sextortion>

²² Snapchat Family Safety Hub, url.



Conclusion

Given the severity of the harm industry-wide, Snap still treats CSEAI as a Level 1 risk priority in response to which it has put in place a range of mitigation measures. This includes in particular our [proactive content moderation](#) which is designed to detect and prevent CSEAI from appearing on each of Snapchat's in-scope services – for example, our automated and human review on Spotlight. Our prevalence testing has continued to help us to improve this proactive content moderation. As a result, we've seen the prevalence of CSEAI on Snapchat fall and be maintained to an extremely low level [REDACTED]

[REDACTED] In addition, while we are alive to the risk, we have not identified any material issue with Snapchat's in-scope services being used for the dissemination of generative CSEAI materials.

We have concluded therefore that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures for the dissemination of CSEAI. There is no change in this conclusion from our 2024 Report.

4.1.2 Dissemination of illegal hate speech

This Section 4.1.2 (Illegal Hate Speech) considers the risk of harm arising from hate content and activity (as described in the Harm Description below) on Snapchat's in-scope services.

Harm Description

Snap describes 'hate' as content that demeans, or promotes discrimination towards, an individual or group of individuals on the basis of their race, color, caste, ethnicity, national origin, religion, sexual orientation, gender, gender identity, disability, veteran status, immigration status, socio-economic status, age, weight, or pregnancy status. Hate Speech may include references to people that are dehumanizing or that compare humans to animals on the basis of these traits and categories. Hate Speech also includes the valorization of perpetrators—or the denigration of the victims—of hateful atrocities (e.g., genocide, apartheid, slavery, etc.), as well as the promotion of hate symbols.

We provide a summary and explanation of 'hate' in our explainer on [Hateful Content Explainer Series](#), and our [Transparency Report Glossary](#).

Likelihood

Snap is sensitive to the issue of hate speech on internet platforms,²³ as well as the damaging effects hate speech can have on a community. Thankfully, hate speech is rarely found on the public surfaces of Snapchat. In our 2024 Report, we highlighted that our prevalence testing showed that hate speech accounted for an extremely low percentage of total views of Snaps in Public Stories. In this Report, we are pleased to confirm that our prevalence testing has shown

²³ Eurostat, *50% of young people encounter hostile messages online*, 1 August 2024, [url](#).



[REDACTED]

that the PVP for Hate Speech has continued to fall dramatically [REDACTED]
[REDACTED]

Moreover, when content related to this specific Harm was detected (whether proactively through automated tools or reactively following a user report) our Safety teams acted swiftly. According to our [latest European Union Transparency Report](#) (covering the second half of 2024), the median turnaround time for enforcement action in response to proactive or reactive detections of this content was **36 minutes**.

As a result of this rapid turnaround time:

- the violative view rate for EU users for this type of content across the following Snapchat in-scope services is very low in the first half of 2025:

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

- the violative view rate for EU minors for this type of content across the following Snapchat in-scope services is also very low in the first half of 2025:

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

Finally, while we engage with Trusted Flaggers on hate speech, out of the 1,424 reports we received from Trusted Flaggers across all harm categories in the EU during the second half of 2024 and the first half of 2025, none came from entities focused on hate speech.

Therefore, the steps Snap has taken to mitigate this harm mean that it is very unlikely that Snapchatters will encounter hate speech on Snapchat’s in-scope services, and Snap continues to place Hate Speech in the **lowest likelihood level of all our risks**.

Severity

In prior reports, we noted that the Council of Europe acknowledged that “hate speech is a deep rooted, complex and multidimensional phenomenon, which takes many dangerous forms and can be disseminated very quickly and widely through the internet, and that the persistent availability of hate speech online exacerbates its impact” and “realising that hate speech negatively affects individuals, groups and societies in a variety of ways and with different degrees of severity, including by instilling fear in and causing humiliation to those it targets and by having a chilling effect on participation in public debate, which is detrimental to democracy”.²⁴ The OSCE High

²⁴ Council of Europe, ‘Recommendation CM/Rec(2022)16 of the Committee of Ministers to member States on combating hate speech’, May 2022, [url](#).



Commissioner on National Minorities also recognised the dangers of hate speech and offered guidance on how to create, nurture, and develop the role of the media and information technologies for conflict prevention in its Tallinn Guidelines.²⁵

Snap’s assessment of the severity of Hate Speech has not changed in this Report. Snap still considers content that qualifies as hate speech to be significant in terms of severity. We continue to believe it is imperative for Snap to combat hate speech in all forms to cultivate a welcoming, positive, non-discriminatory online environment for all Snapchatters. This includes content that demeans, defames, or promotes discrimination towards an individual or group of individuals on the basis of protected characteristics and the promotion of hate groups. Because of the wide scope of the term, and the thin line with expressions that do fall within the scope of freedom of expression, we continue to include hate speech in the “**significant**” harms category.

DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat’s in-scope services. As set out below, while the following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap’s existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Our recommender systems are designed to not knowingly recommend hate speech, i.e. there are no interest categories related to hate speech. Detailed information regarding the mitigations relating to our recommender systems can be found in Section 5 (see the Algorithmic Systems sub-section).
(b) content moderation systems;	Snap has implemented specific proactive and reactive moderation procedures to prevent and remove hate speech, including service-specific adaptations to address violating content, such as hate speech. We provide more information in Section 5 (see Content Moderation sub-section).
(c) the applicable terms and conditions and enforcement;	Our terms prohibit hate speech and are strictly enforced. As explained in Section 5 (see Terms and Enforcement sections), on our potentially high-reach surfaces, like Spotlight and Discover, we take a proactive approach to moderating any content that may violate our rules on hate speech. Our in-app reporting tool also allows users to

²⁵ OSCE, ‘The Tallinn Guidelines on National Minorities and the Media in the Digital Age’, February 2019, [url](#).



	directly report hateful content or activities that support terrorism or violent extremism. When hateful content is reported, our teams will remove any violating content and users who engage in repeated or egregious violations will have their account access locked. Lenses identified with hate speech were rejected when found during submission and disabled in Discover upon review if subsequently identified. As an additional measure, we encourage Snapchatters to block any users who make them feel unsafe or uncomfortable. Snap removes violating hate speech as soon as we become aware of it, and will promptly disable accounts dedicated to hate speech, hate symbols or groups, or the glorification of hate groups or members of a hate group.
(d) systems for selecting and presenting advertisements; and	Our advertising systems require agreement to advertising policies and guidance that prohibit adverts from displaying information that violates the law or causes certain harms. We check advertisers are complying with their obligations via our advertising review process. Our advertising systems use a mix of automation and human review to review adverts before they are published. We provide easy mechanisms for users to hide and report advertisements that violate our policies or the law. We monitor ad rejection, reporting and enforcement data to monitor the effectiveness of our approach. We provide more information in Section 5 (see the Advertising Systems sub-section)
(e) our data related practices	We have strong data principles, practices and privacy, safety and security by design processes. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).

We have also analysed whether and how the risk of this harm is influenced by the following general factors:

General Risk Factor	How does it apply to Snapchat and this harm?
Intentional manipulation, including inauthentic use or inauthentic use or automated exploitation of the service	There are two key ways in which we consider our systems could be manipulated: (1) Users could seek to share novel illegal and violating material that is not detected by our automated systems. We are constantly working to



	adjust our systems and policies to address this. We provide more information on our approach in Section 5 (in particular the Content Moderation subsection) and Section 6 (Ongoing Risk Detection and Management). (2) Users could abuse our content moderation processes and report non-violating content / accounts in bad faith. We have processes to combat misuse and more information can be found in Section 5 (Content Moderation and Enforcement subsections).
Amplification and potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.	Snapchat's inscope services have a number of features and design configurations that act to limit the amplification and potentially rapid and wide dissemination of illegal hate speech, 1. Snapchat is not an attractive platform for spreading hate speech, in particular because it is difficult to reach a broad audience and content is deleted by default. We provide more information in Section 5 (see Snapchat Design / Function subsection) 2. Snap has implemented specific proactive and reactive moderation procedures to prevent and remove hate speech. We provide more information in Section 5 (see Content Moderation). 3. Content recommended to users on Spotlight and Discover, our video sharing platforms, is moderated using a combination of auto-moderation and human moderation, and is human moderated before being widely distributed. Lens and Ads are subject to review processes before submission. We provide more information in Section 5 (see Content Moderation).
Specific regional or linguistic aspects, including when specific to a Member State.	We recognise that our users may come from different Member States and content may be shared in different languages. To address this, • Snapchat is provided in multiple EU languages and our Terms are available in all EU languages. We provide more information in Section 5 (see Terms). • We deploy content moderation systems and moderators work across multiple EU languages. We provide more information in Section 5 (see Content Moderation).



Overall potential risk prioritization

As Snap continues to qualify hate speech as ‘significant’ in terms of severity but ‘lowest’ in likelihood given the relatively low prevalence on the platform, Snap considers hate speech a **Level 3 risk prioritization**. There is no change in this assessment from our 2024 Report.

Snap's Mitigations

Snap is aware of the serious harm posed by hate content online. In line with our commitment to user safety and legal compliance, we have implemented a broad range of measures to significantly reduce the likelihood and impact of hate-related offences occurring on Snapchat’s in-scope services.

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat’s in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive. As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **link in the left hand column to a full summary** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, our in-scope services have been adapted to include proactive moderation for hate speech.



Mitigation Category	Applies to this risk?
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes, our terms prohibit hate speech and they are strictly enforced. As explained in our Moderation and Enforcement sections in Section 5 of our report, on our potentially high-reach surfaces, like Spotlight and Discover, we take a proactive approach to moderating any content that may violate our rules on hate speech. Our in-app reporting tool also allows users to directly report hateful content. When hateful content is reported, our teams will remove any violating content and users who engage in repeated or egregious violations will have their account access locked. Lenses identified with hate speech were rejected when found during submission and disabled in Discover upon review if subsequently identified. As an additional measure, we encourage Snapchatters to block any users who make them feel unsafe or uncomfortable. Snap removes violating hate speech as soon as we become aware of it, and will promptly disable accounts dedicated to hate speech, hate symbols or groups, or the glorification of hate groups or members of a hate group. Specific rules²⁶ apply to our Lenses.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, specific proactive and reactive moderation procedures to prevent and remove hate speech. This includes service-specific adaptations to address violating content such as hate speech. [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] If Lenses are found to be violating our rules and promoting misogynist content, Snap takes enforcement action against such Lenses (for example, we have removed a few Lenses promoting Andrew Tate).
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, our algorithmic systems do not knowingly recommend hate speech – i.e., there are no interest categories relating to hate speech.
Advertising Systems	Yes, other mitigations listed here also apply to our Advertising Systems.

²⁶ <https://businesshelp.snapchat.com/s/article/Lens-Restrictions?language=en>



Mitigation Category	Applies to this risk?
Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, for example we have specific prevalence testing and monitoring moderation and enforcement data which we use to help detect and manage hate speech.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Yes, we cooperate with trusted flaggers in relation to hate speech. In the EU, this includes Licra (France), the German Department for Internet Services & Social Media and Someturva (Finland).
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various groups in relation to hate speech. This includes the Government's Counter Disinformation Unit and more broadly, Snap remains a signatory of the EU Code of Conduct to counter hate speech online and has worked hard to ensure Snap meets the requirements (including with respect to recent revision of that Code).
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on harms and how to get help in our Safety Center . We make available robust reporting; and we provide guidance to parents on the web.
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control	Yes, we have protective measures to limit Teen contact with strangers; we offer Family Center , reporting and guidance. Our new parents site provides additional guidance for parents and caregivers on risks and support. ²⁷

²⁷ Snapchat Family Safety Hub, [url](#).



Mitigation Category	Applies to this risk?
tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	Yes, Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services.

Conclusion

Snap considers illegal hate speech a Level 3 risk prioritization. In response it has put in place a range of mitigation measures. These include in particular our alignment to the EU Hate Speech code of practice) and our proactive content moderation which is designed to detect and prevent illegal hate speech from reaching a broad audience on Snapchat's in-scope services. We monitor the prevalence of hate speech in general via our [Prevalence Testing](#) and external reporting which we publish in our [Transparency Reports](#). As a result of the mitigation measures Snap has taken, hate speech continues to be one of our lowest prevalence risks

We have concluded therefore that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures for the dissemination of illegal hate speech. There is no change in this conclusion from our 2024 Report.

4.1.3 Dissemination of information related to the sale of prohibited products or services

This Section 4.1.3 (Sale of Prohibited and Regulated Products or Services) considers the risk of harm arising from the sale of prohibited products or services (as described in the Description below) on Snapchat's in-scope services.



Harm Description

This harm includes content or activity that promotes, facilitates, or enables the illegal supply, trade, or solicitation of drugs, weapons, and other prohibited or controlled goods and services. It encompasses user-generated content, profiles, messages, and interactions which:

- Advertise or promote the sale of illicit drugs or controlled substances;
- Offer firearms, knives, weapons or promote illegal weapons trafficking;
- Enable or promote the trafficking of individuals, including content relating to the recruitment, transportation, or exploitation of people through force, fraud, or coercion.
- Facilitate the sale or procurement of other goods and services prohibited by law.

Snap prohibits any content that encourages or enables engagement in illegal transactions or activity relating to drugs, weapons or other controlled/prohibited goods. Further definitions on this are set out in our [Transparency Report Glossary](#) and in our explainer on [Illegal or Regulated Activities](#).

Likelihood

We recognize that the sale of illegal products and services via social media is an issue of increasing concern among public health authorities, law enforcement, and regulators. For example, the *WIRED* article “Drug Dealers Have Moved On to Social Media”²⁸ reflects broader findings from European institutions, including the European Union Drugs Agency (EUDA), which has documented the migration of drug markets onto online intermediary services. Similarly, a report commissioned by the EUDA²⁹ highlights that digital drug dealing is active across several EU member states, particularly in Northern Europe. However, these reports typically focus on the private messaging and limited broadcast features of online intermediary services, such as encrypted messaging. Please note that, while Snap takes these issues very seriously and continues to invest significant resources to combat this criminal and violative activity on the private and limited broadcast services of Snapchat, such activity is out of scope of this Report.

With regards to Snapchat’s in-scope services, which concern how our online platforms disseminate information to the general public, our testing shows that the prevalence of information related to the sale of prohibited products or services continues to decrease year on year. [REDACTED]

[REDACTED] This is a further, significant decrease from [REDACTED] that we reported for July 2024 in our 2024 Report.

Moreover, when content related to this specific Harm was detected (whether proactively through automated tools or reactively following a user report) our Safety teams acted swiftly. Our [latest](#)

²⁸ Wired, ‘Drug Dealers Have Moved on to Social Media’, [url](#), December 17, 2024.

²⁹ European Monitoring Centre for Drugs Addiction, ‘An Analysis of Drug Dealing via Social Media’, [url](#), 2022.



[REDACTED]

[European Union Transparency Report](#) (covering the second half of 2024) recorded a median enforcement turnaround time of **58 minutes** for drug-related content, **3 minutes** for content related to weapons, and **10 minutes** for content related to other regulated related goods, following either proactive detection or user reports.

[REDACTED]

- [REDACTED]
[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

- [REDACTED]
[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

The likelihood of encountering such violating content on Snapchat’s in-scope services is now at its lowest level and falls within the **lowest category in terms of likelihood**.

Severity

Snap continues to consider this issue to be **severe** where a credible threat to human life, safety, or well-being existed. Snap also considers the depiction or use of, or attempts at buying, selling, exchanging, or facilitating sales of illegal drugs to be a **severe** harm. As noted in our 2024 Report, Snap has experienced bad actors trying to exploit our product architecture for the illicit sale of drugs, generally presenting acute risks.³⁰ According to the European Monitoring Centre for Drugs and Drug Addiction, “the use of illicit drugs causes a range of acute and chronic harms and is a recognised contributor to the global burden of disease.” The 2023 European Drug Report states

³⁰ New York Times, ‘Fentanyl Tainted Pills Bought on Social Media Cause Youth Drug Deaths to Soar’, [url](#), 19 May 2022. Cf. Snap, ‘How Snap is responding to the fentanyl crisis’, [url](#), 7 October 2021; Snap, ‘Continuing our Efforts to Combat the U.S. Fentanyl Crisis’, [url](#), 12 October 2022 ; Snap, ‘National Fentanyl Awareness Day: Continuing Our Efforts to Combat the U.S. Fentanyl Crisis’, [url](#), 9 May 2023.



chronic problems include “dependence and drug-related infectious disease, while there is a range of acute harms, of which drug overdose is the best documented.”³¹

Snap considers attempts to buy or sell weapons and depicting or brandishing weapons in a threatening or violent context to be a **serious** risk in terms of severity. As several reports show, gun trafficking is a major concern in the context of human security.³² Being instrumental in much violence, guns pose a significant threat to human life and well-being. Illicit firearms may be used for self-harm or domestic violence, or ultimately end up with criminals, supporting operations related to armed conflicts and terrorism.³³ Studies show that firearm manufacturers use social media to attract audiences to their websites, contributing to the spread of gun violence.³⁴ For instance, the European Parliamentary Research Service (EPRS) briefing “Understanding EU policy on firearms trafficking” highlights a concerning trend: illicit firearms are increasingly accessible through online channels, enabling individuals, even without criminal ties, to procure weapons more easily than before. This marks a shift from traditionally closed illicit markets to more open, internet-enabled trafficking networks, facilitated by the availability of easily convertible or illegal components online.³⁵

Snap also tracks the dissemination of illegal or otherwise violating content that relates to other regulated or prohibited goods as it recognises the use of online platforms for selling other illegal goods or services online also places consumers at risk.³⁶ For example, Snapchat has taken into serious consideration the concerns raised by Dutch health experts and youth organizations regarding the alleged illegal sale of vape related products to minors on the platform while further engaging with NGOs and law enforcement authorities to adopt measures aimed at preventing such illegal sale. We recognize that the sale and promotion of regulated products on social media is an area of public and regulatory interest, and we continue to follow developments closely, reassess mitigations and implement further safeguards in response. Consequently, Snap considers the dissemination of violating content that relates to other regulated or prohibited goods to give rise to a **significant** risk of harm.

DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat’s in-scope services. As set out below, while the

³¹ European Monitoring Centre for Drugs and Drug Additions, ‘*European Drug Report 2023*’, [url](#), 2023.

³² United Nations Office on Drugs and Crime, ‘*Global Study on Firearms Trafficking 2020*’, [url](#), 2020.

³³ A. Neville, ‘*European Parliamentary Research Service. Understanding EU policy on firearms trafficking*’, [url](#), February 2022.

³⁴ L. Jordan e.a., “Characteristics of Gun Advertisements on Social Media: Systematic Search and Content Analysis of Twitter and YouTube Posts”, *J Med Internet Res* (2020, 22(3)), March 2020 ([url](#)).

³⁵ European Parliamentary Research Service, ‘*Understanding EU policy on firearms trafficking*’ [url](#), 2024

³⁶ J.P. Kennedy and J.M. Wilson, ‘*Clicking Into Harm’s Way: The Decision to Purchase Regulated Goods Online*’, September 2017, [url](#).



following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap’s existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Our recommender systems are designed not to knowingly recommend content concerning the sale of prohibited products or services content, i.e. there is no interest category for this content. Detailed information regarding the mitigations relating to our recommender systems can be found in Section 5 (see the Algorithmic Systems sub-section).
(b) content moderation systems;	Snap has implemented specific proactive and reactive moderation procedures to prevent the sale of prohibited products or services. As explained in the Content Moderation Section, Snap deploys a range of automated content moderation [REDACTED] [REDACTED] to detect and moderate content relating to the sale of prohibited and regulated goods and services, and we have aggressively focused on enforcement of severe and serious harms. We provide more information in Section 5 (see Content Moderation sub-section).
(c) the applicable terms and conditions and enforcement;	Terms prohibit the sale of prohibited products or services and they are strictly enforced with the most serious consequences. Snap complies with relevant legal requirements to remove content about the sale of prohibited and regulated goods and services, and takes appropriate action against egregious or repeat violators. Snap works with law enforcement, safety organizations, and subject matter experts to continue to educate ourselves and our community, and to take appropriate action where these threats may arise on our platform. [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]
(d) systems for selecting and presenting advertisements; and	Our advertising systems require agreement to advertising policies and guidance that prohibit adverts from displaying



	information that violates the law or causes certain harms. We check advertisers are complying with their obligations via our advertising review process. Our advertising systems use a mix of automation and human review to review adverts before they are published. We provide easy mechanisms for users to hide and report advertisements that violate our policies or the law. We monitor ad rejection, reporting and enforcement data to monitor the effectiveness of our approach. We provide more information in Section 5 (see the Advertising Systems sub-section).
(e) our data related practices	We have strong data principles, practices and privacy, safety and security by design processes. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).

We have also analysed whether and how the risk of this harm is influenced by the following general factors:

Service Risk Factor	How does it apply to Snapchat and this harm?
Intentional manipulation, including inauthentic use or inauthentic use or automated exploitation of the service.	There are two key ways in which we consider our systems could be manipulated: (1) Users could seek to share novel illegal and violating material that is not detected by our automated systems. We are constantly working to adjust our systems and policies to address this. We provide more information on our approach in Section 5 (in particular the Content Moderation subsection) and Section 6 (Ongoing Risk Detection and Management). (2) Users could abuse our content moderation processes and report non-violating content / accounts in bad faith. We have processes to combat misuse and more information can be found in Section 5 (Content Moderation and Enforcement subsections).
Amplification and potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.	Snapchat's in-scope services have a number of features and design configurations that act to limit the amplification and potentially rapid and wide dissemination of content concerning the sale of prohibited products or services: 1. Snap makes it difficult for unvetted content to reach



	a large audience on Snapchat, and Snap proactively moderates Snapchat's in-scope services that provide an opportunity to reach a larger audience. As a result, this reduces the likelihood of our users encountering content concerning the sale of prohibited products or services on the platform. 2. Snapchat has also been designed to limit the prevalence of content concerning the sale of prohibited products or services. Snap has implemented specific proactive and reactive moderation procedures to prevent and remove Adult Sexual Content. We provide more information in Section 5 (see Content Moderation). 3. Content recommended to users on Spotlight and Discover, our video sharing platforms, is moderated using a combination of auto-moderation and human moderation, and is human moderated before being widely distributed. Lens and Ads are subject to review processes before submission. We provide more information in Section 5 (see Content Moderation).
Specific regional or linguistic aspects, including when specific to a Member State.	We recognise that our users may come from different Member States and content may be shared in different languages. To address this, • Snapchat is provided in multiple EU languages and our Terms are available in all EU languages. We provide more information in Section 5 (see Terms). • We deploy content moderation systems and moderators work across multiple EU languages. We provide more information in Section 5 (see Content Moderation).

Overall potential risk prioritization

Thankfully, as reported in our 2024 Report, this is still not a common issue on Snap. Our [Prevalence Testing](#) revealed that communication around Illegal goods and activities now has only a very small PVP rate [REDACTED]. However, due to the severity of some potential products and services (such as communication around dangerous or illicit drugs), prevalence is not the determinative factor for Snap's prioritization of this issue. Snap prioritizes severe harm and legal compliance over prevalence on the platform, and for this category has decided to deviate from the standard risk framework, as demonstrated in the below graph.

[illegible]

Snap is aware of the risks associated with content that facilitates the sale, promotion, or distribution of prohibited goods or services on internet platforms and services. In response, Snap has implemented robust measures designed to significantly reduce both the likelihood and potential impact of prohibited goods or services content on Snapchat.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Snapchat does not offer a marketplace for the sale of goods.
Terms and Enforcement	Yes, terms prohibit the sale of prohibited products or services



Mitigation Category	Applies to this risk?
Adapting their terms and conditions and their enforcement.	and they are strictly enforced with the most serious consequences. As explained in the Enforcement section of Section 5 of our Report, Snap complies with relevant legal requirements to remove content about the sale of prohibited and regulated goods and services, and takes appropriate action against egregious or repeat violators. Snap works with law enforcement, safety organizations, and subject matter experts to continue to educate ourselves and our community, and to take appropriate action where these threats may arise on our platform. When we identify violators engaging in the attempted buying, selling, exchanging, or facilitating sales of dangerous and illicit products and services, we disable their accounts and, in some instances, refer the conduct to law enforcement. For less severe harms, a user will be warned and their content removed. Repeat violations will result in violators' accounts being disabled.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, specific proactive and reactive moderation procedures to prevent the sale of prohibited products or services. As explained in the Content Moderation Section of Section 5 of the Report, we have proactive and reactive moderation processes in place to detect and moderate content relating to the sale of prohibited and regulated goods and services, and we have aggressively focused on enforcement of severe and serious harms.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, our algorithmic systems do not knowingly recommend content concerning the sale of prohibited products or services content, i.e. there is no interest category for this content.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of	Yes, other mitigations listed here also apply to our Advertising Systems.



Mitigation Category	Applies to this risk?
advertisements in association with the service they provide.	
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, for example, we have specific prevalence testing and monitoring moderation and enforcement data which we use to help detect and manage information related to the sales of prohibited products and services.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Yes, we cooperate with trusted flaggers in relation to the sale of prohibited products or services, in particular the National Crime Agency (NCA - particularly in relation to the sale of knives and drugs).
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various industry groups and share signals, especially in relation to drug dealers. Snap spearheaded signal sharing with Meta relating to illicit drugs content, and also collaborates with law enforcement to receive and share signals.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we have protective measures to limit Teen contact with strangers; we offer Family Center ; we make available robust reporting; and we provide guidance to parents on the web. Across Snapchat, we offer a number of resources to users to raise awareness on safety topics and protect them. As explained in the Transparency section of the Report, one of the examples of this is our in-app tool, Heads Up. This surfaces educational content from experts to Snapchatters if they try to search for drug-related content.
Protection of Minors Taking targeted measures to protect the rights of the child, including age	Yes, we have protective measures to limit teen contact with strangers, Family Center , reporting, and guidance. Our parents site provides additional guidance for parents and caregivers



Mitigation Category	Applies to this risk?
verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	on risks and support. ³⁷
<u>Content Authenticity</u> Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	General content authenticity measures taken in respect of the sale of prohibited products or services.

Conclusion

Despite the continued very low prevalence, we still consider the overall risk of the dissemination of the sale of dangerous or illicit drugs, or any other prohibited products or services that pose a threat to human life, safety, or well-being, to be in the Level 1 category due to the level of severity. Snap continues to consider that the sale of weapons poses a Level 2 overall potential risk, and a Level 3 potential overall risk in relation to other prohibited products and services. Snap continues to take steps to mitigate these harms, which has further diminished the likelihood that Snapchatters will find information related to prohibited products or services on Snapchat’s in-scope services. Snap continues to invest significant resources to further combat these harms, and are still looking to achieve further reductions in the likelihood of this risk where possible.

We have concluded therefore that Snapchat’s in-scope services have reasonable, proportionate and effective mitigation measures for dissemination of information related to the sale of prohibited products or services. There is no change in this conclusion since our 2024 Report.

³⁷ Snapchat Family Safety Hub, [url](#).



4.1.4 Dissemination of Terrorist Content

This Section 4.1.4 (Terrorism) considers the risk of harm arising from terrorism content and activity (as described in the Harm Description below) on Snapchat's in-scope services defined in Section 2 (Scope).

We have assessed this risk in accordance with the Section 3 (Methodology) as follows.

Harm Description

Terrorism refers to content that promotes or supports terrorism or other violent, criminal acts committed by individuals and/or groups to further ideological goals, such as those of a political, religious, social, racial or environmental nature. It includes any content that promotes or supports any terrorist organisation, as well as content that advances recruitment for terrorist organisations. Terrorism content includes material that promotes, incites, glorifies, or facilitates terrorist activity or ideology. This may take the form of propaganda, symbols, video or text glorifying mass violence, or recruitment communications.

We provide a summary and explanation of 'terrorism' in our explainer on [Hateful Content, Terrorism and Violent Extremism](#) and our [Transparency Report Glossary](#).

Likelihood

We have considered the likelihood of this harm in line with Section 3 (Methodology) and observed the following for 2025:

- As noted in our 2024 Report, online influences have been depicted as major drivers for the propagation and adoption of extremist ideologies, which often contain an element of collective grievance, and subsequent acts of violence.³⁸ Terrorist organisations continue to target young people, spreading propaganda especially on those social media platforms that are particularly popular among younger users, and adapting content and communication strategies to these platforms and their audiences.³⁹ It is conceivable that bad actors could disseminate Terrorist Content on Snapchat, as with any other online platform. This could include, in particular, Terrorist Content appearing in videos featured on Spotlight / Discover and extremist content and individuals promoted via Public Profiles.
- Snap internal media report analysis has not identified significant terrorism-related media referring to Snapchat. A March 2024 article in The Guardian discussed social media and radicalisation, but Snapchat was not cited. Only a couple of articles did identify activity on Snapchat:

³⁸ J.F. Bender and J. Kenyon, *Terrorism and the internet: How dangerous is online radicalization?*, Front. Psychol., 13 October 2022, [url](#).

³⁹ Europol, *European Union Terrorism Situation and Trend Report 2025*, page 10, [url](#); Council of the European Union, *Council conclusions on future priorities for strengthening the joint counterterrorism efforts of the European Union and its Member States - Council conclusions (12 December 2024)*, p. 14, [url](#).



- Essex teenage ISIS plot exposed via Snapchat - In May 2023, a 19-year-old from Essex named Matthew King used Snapchat to post an image of police officers with the caption “Target Acquired”, as part of an Islamist-inspired plot to kill soldiers or police. He carried out surveillance at several sites in London and later received a life sentence for preparing terrorist acts⁴⁰
- Spanish plane “joke” incident involving a British student - Although not prosecuted in the UK, an 18-year-old UK student made a disturbing “joke” on Snapchat in 2022 about being a Taliban member and bombing a plane. The message triggered a Spanish security response but did not result in terrorism charges in the UK (UNILAD)
- Similarly, CEP⁴¹ and Tech Against Terrorism⁴² focused on encrypted messaging platforms (e.g. Telegram), livestream services, and global video platforms. Snapchat was not referenced in any current sector-wide studies.
- In the European Union, in response to the various events influencing violent extremism and terrorism online through 2023, about 349 removal orders were sent by 6 EU member States' competent authorities to 13 online platforms under the Regulation on dissemination of terrorism content online;⁴³ none were directed at Snap.⁴⁴
- Tech against Terrorism’s 2022 report⁴⁵ also noted that file-sharing platforms (rather than social media platforms or messaging services) are increasingly being exploited by terrorist and violent extremist communities for content that is most likely to be moderated and removed from social media or messaging apps.
- [REDACTED]
- In the second half of 2024, for example, we also locked 4 accounts and removed 90 content in the EU for violations of our policy prohibiting terrorist and violent extremist content, as recorded in our [Transparency Reports](#). Where Terrorism related content was identified (either proactively via automated tools or reactively following a report), our

⁴⁰ The Standard, ‘Essex teenager planned murder of soldiers or police in Snapchat terror plot’, [url](#), 12 May 2023.

⁴¹ A. Hoffmann, A. Küsters & P. Eckhardt, *Security and Trust: An Unsolvable Digital Dilemma?*, [url](#), 11 May 2025.

⁴² [Report: Terrorist Use of End-to-End Encryption - Insights from a Year of Multi-Stakeholder Discussion](#)

⁴³ Regulation (EU) 2021/784.

⁴⁴ Violent Extremism and Terrorism Online in 2023: The Year in Review, April 2024, [url](#).

⁴⁵ Tech Against Terrorism: The Threat of Terrorist and Violent Extremist Operated Websites, January 2022, [url](#).



median turnaround time was rapid. Our [latest European Union Transparency Report](#), which covers the second half of 2024, observed a median turnaround time for our Safety teams to take enforcement action in response to proactive or reactive detections of terroristic content of **6 minutes**.

- [REDACTED]

[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

- As reported in our 2024 Report, we have previously sought independent analysis via third party intelligence vendors (SITE and MEMRI) that track extremist activity online who have verified that Snapchat does not fall into the top 100 communications platforms used by extremist groups to communicate.

These data indicate that, while there has been a slight increase in the likelihood of encountering Terrorist Content on Snapchat, it remains extremely low and it still falls within the **lowest likelihood category**.

Severity

The European Commission has stated that the presence of Terrorist Content online is a grave risk to citizens and to society at large. Terrorist and violent groups use the Internet and associated technologies for radicalization, recruitment, dissemination of propaganda, communication and mobilization. They spread their messages to intimidate, radicalize, recruit, and facilitate carrying out terrorist attacks.⁴⁶ In general, terrorist and violent content posted online can be disseminated quickly and cheaply, amplifying dangerous views and reaching (and possibly desensitizing) broad audiences.⁴⁷ An emerging risk in this context is the use of generative AI, which could enable terrorist organizations to rapidly expand their reach and influence by producing large volumes of persuasive, and often false, content that amplifies extremist narratives.⁴⁸ Security and safety, as

⁴⁶ European Commission, *Terrorist Content Online*, [url](#).
⁴⁷ OECD, *Transparency Reporting on Terrorists and Violent Extremist Content Online*, July 2021, No. 313, page 4, [url](#).
⁴⁸ C. Anthony Pfaff (Centre of Excellence Defence Against Terrorism of NATO), *The Weaponization of AI: the Next Stage of Terrorism and Warfare*, 2025, [url](#).



well as the dignity of victims, are seriously threatened by this.⁴⁹ As a result, Snap continues to consider the severity of this issue to be **serious** due to the high threat to human life, safety, or well-being.

DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat's in-scope services. As set out below, while the following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap's existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Our algorithmic systems do not knowingly recommend terrorism content – i.e., there is no 'terrorism' interest category. As explained in Section 5 (see the Algorithmic Systems sub-section), on our high-reach surfaces, like Spotlight and Discover, we take a proactive approach to moderating any content that may violate these rules prior to the content being recommended to a wide audience
(b) content moderation systems;	Snap has implemented specific proactive and reactive moderation procedures to prevent and remove Terrorist Content. Snap deploys a range of automated content moderation (which include abusive language detection, other keyword-based detection, and machine-learning-based proactive detection models) to scan Stories and Spotlight submissions. Snapchatters can report Terrorist Content to us via in-app reporting options and anyone can submit a report through the Snapchat Support Site. Our in-app reporting tool also allows users to directly report hateful content or activities that support terrorism or violent extremism. We provide more information in Section 5 (see Content Moderation sub-section).
(c) the applicable terms and conditions and enforcement;	Our Community Guidelines (which form part of our terms of service) expressly prohibit terrorist organizations, violent extremists, and hate groups from using our platform. Terrorism is considered a "severe harm" in our Community Guidelines and we respond with swift and strict consequences against violators as explained in our Severe Harms explainer . We provide more information in Section 5 (see the Terms and Enforcement sub-sections).

⁴⁹ European Commission, *Tackling Illegal Content Online Towards an enhanced responsibility of online platforms* (COM(2017) 555 final), [url](#), 28 September 2017; Europol, *European Union Terrorism Situation and Trend Report 2025*, page 11, [url](#)



(d) systems for selecting and presenting advertisements; and	Our advertising systems require agreement to advertising policies and guidance that prohibit adverts from displaying information that violates the law or causes certain harms. We check advertisers are complying with their obligations via our advertising review process. Our advertising systems use a mix of automation and human review to review adverts before they are published. We provide easy mechanisms for users to hide and report advertisements that violate our policies or the law. We monitor ad rejection, reporting and enforcement data to monitor the effectiveness of our approach. We provide more information in Section 5 (see the Advertising Systems sub-section)
(e) our data related practices	We have strong data principles, practices and privacy, safety and security by design processes. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).

We have also analysed whether and how the risk of Terrorism is influenced by the following general factors:

General Risk Factor	How does it apply to Snapchat and this harm?
Intentional manipulation, including inauthentic use or inauthentic use or automated exploitation of the service	There are two key ways in which we consider our systems could be manipulated: (1) Users could seek to share novel illegal and violating material that is not detected by our automated systems. We are constantly working to adjust our systems and policies to address this. We provide more information on our approach in Section 5 (in particular the Content Moderation subsection) and Section 6 (Ongoing Risk Detection and Management). (2) Users could abuse our content moderation processes and report non-violating content / accounts in bad faith. We have processes to combat misuse and more information can be found in Section 5 (Content Moderation and Enforcement subsections).
Amplification and potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.	Snapchat's inscope services have a number of features and design configurations that act to limit the amplification and potentially rapid and wide dissemination of Terrorist Content, in particular: 1. Snapchat is not an attractive platform for spreading Terrorist Content, in particular because it is difficult to reach a broad audience and content is deleted by default. We provide



	more information in Section 5 (see Snapchat Design / Function subsection) 2. Snap has implemented specific proactive and reactive moderation procedures to prevent and remove Terrorist Content. As explained in the Content Moderation section, Snap deploys a range of automated content moderation (which include abusive language detection, other keyword-based detection, and machine-learning-based proactive detection models) to scan Stories and Spotlight submissions. 3. Content recommended to users on Spotlight and Discover, our video sharing platforms, is moderated using a combination of auto-moderation and human moderation, and is human moderated before being widely distributed. Lens and Ads are subject to review processes before submission. We provide more information in Section 5 (see Content Moderation)
Specific regional or linguistic aspects, including when specific to a Member State.	We recognise that our users may come from different Member States and content may be shared in different languages. To address this, • Snapchat is provided in multiple EU languages and our Terms are available in all EU languages. We provide more information in Section 5 (see Terms). • We deploy content moderation systems and moderators work across multiple EU languages. We provide more information in Section 5 (see Content Moderation).

Overall potential risk prioritization

Due to the very low prevalence of extremist content on Snapchat, the overall risk would normally be assessed to be Level 3. However, the consequences of terrorism present a potential for severe harm to human life, safety or wellbeing. For that reason, Snap has decided to deviate from the standard risk framework, and has placed Terrorist Content within our **Level 2 overall potential risk prioritization category**, as outlined below. There is no change in this assessment from our 2023 Report.



Snap's Mitigations

Snap is aware of the threat posed by Terrorist Content on internet platforms and services, and the increased inherent risk arising from the types of Risk Factors highlighted by DSA. Snap has put in place significant measures designed to substantially diminish the likelihood and impact of Fraud and Spam on Snapchat. In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services. These are organised into Snap's risk assessment mitigation categories.

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. Note that the primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive. As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **link in the left hand column to a full summary** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, several aspects of Snapchat's design and function reduce the risk of terrorism activity. In particular, Snap makes it difficult for unvetted content to reach a large audience on Snapchat, and Snap proactively moderates Snapchat's in-scope services that provide an opportunity to reach a larger audience. Private messaging services require mutual friendship. Group chats are limited to a maximum of 201 users and are not publicly searchable.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes, Snap's Terms and Community Guidelines expressly prohibit terrorist organizations, violent extremists, and hate groups from using our platform. Terrorism is considered a "severe harm" in our Community Guidelines and we respond with swift and strict consequences against violators as explained in our Severe Harms explainer. Our prohibitions against Terrorism and Violent Extremism extend to all forms of content that promotes terrorism or other violent,



Mitigation Category	Applies to this risk?
	criminal acts committed by individuals or groups to further ideological goals. These rules also prohibit any content that promotes or supports foreign terrorist organizations or extremist hate groups— as designated by credible, third-party experts—as well as recruitment for such organizations or violent extremist activities. We promptly enforce against accounts found to be sending Terrorism content: • Snap removes such content for all users. • Accounts we discover engaging in prohibited terrorist activity will also be promptly disabled. • Where appropriate, accounts engaging in violation of these policies may be reported to law enforcement.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, we have specific proactive and reactive moderation procedures to prevent and remove Terrorism related content. As explained in the Moderation section in Section 5, Snap deploys a range of automated content moderation (which include abusive language detection, other keyword-based detection, and machine-learning-based proactive detection models) to scan Stories and Spotlight submissions. Snapchatters can report Terrorist Content to us via in-app reporting options and anyone can submit a report through the Snapchat Support Site. Our in-app reporting tool also allows users to directly report hateful content or activities that support Terrorism or Violent Extremism.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, unlike many of our peers, Snapchat does not offer an open news feed where unvetted publishers or individuals have an opportunity to broadcast Terrorist Content, does not offer a broad ‘reshare’ functionality that would encourage virality, and does not allow user-generated content to be recommended to a wide audience without going through review. Our algorithmic systems do not knowingly recommend Terrorism content – i.e., there is no ‘terrorism’ interest category. As explained in the Moderation section of our Existing Mitigations in Section 5, on our high-reach surfaces, like Spotlight and Discover, we take a proactive approach to moderating any content that may violate these rules prior to the content being recommended to a



Mitigation Category	Applies to this risk?
	wide audience.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, for example we monitor data to help detect and manage Terrorist Content, including data from our specific prevalence testing and enforcements (which are summarised in our Transparency Reports).
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	General trusted flagger group we currently work with to address Terrorism content in the European Union. [REDACTED] [REDACTED]
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various industry groups, including the EU Internet Forum (EUIF), that consider Terrorist Content. Note, due to the low prevalence of Terrorist Content on Snap, we do not participate in the primary multi stakeholder organization: The Global Internet Forum to Counter Terrorism (GIFCT). We also consult the expertise and work of civil rights organizations, human rights experts, law enforcement agencies, NGOs, and safety advocates to help enforce our Guidelines. Such expert knowledge comes from sources such as the Anti-Defamation League, the Southern Poverty Law Center, the Election Integrity Partnership, the Atlantic Council, the Stanford Cyber Policy Center, the members of Snap's Safety Advisory Board, and individual domain experts (including a former Ambassador to the UN Human Rights Council, leading digital rights scholars and advocates, former regulators and policymakers, and geopolitical experts).



Mitigation Category	Applies to this risk?
	We have an internal content crisis protocol that applies in the event of major incidents that may give rise to significant elevated risk of Terrorism content on Snapchat.
<u>Transparency</u> Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on harms and how to get help in our Safety Center. We make available robust reporting tools; and we provide guidance to parents on the web (see below).
<u>Protection of Minors</u> Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, we have protective measures to limit Teen contact with strangers; we offer <u>Family Center</u> , reporting, and guidance. Our new parents site provides additional guidance for parents and caregivers on risks and support. ⁵⁰
<u>Content Authenticity</u> Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	Yes, Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's inscope services.

Conclusion

Despite terrorism falling within our lowest relative likelihood category, we consider the potential risk of the dissemination of Terrorist Content to be in the serious severity category due to potential for serious harm from exposure to Terrorist Content. As a result, terrorism falls within our Level 2 overall risk prioritisation category.

⁵⁰ <https://parents.snapchat.com>.



[REDACTED]

We have traditionally considered that Snapchat’s design, deletion by default and proactive detection measures make it a challenging environment for effective dissemination of Terrorist Content on Snapchat’s in-scope services. We remain vigilant for any changes in use or increased terrorism activity on our platform. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

We have concluded therefore that Snapchat’s in-scope services continue to have reasonable, proportionate and effective mitigation measures for dissemination of Terrorist Content, but we will continue to closely monitor prevalence levels for this harm, amongst other indicators, to detect and manage any new or increased material terrorism threats on Snapchat.

4.1.5 Dissemination of content that infringes on intellectual property rights

Section 4.1.5 (IP Infringement) addresses the risk of harm associated with users uploading and distributing content that may infringe upon the intellectual property rights of others. Such infringing material could potentially surface on Snap’s public platforms, particularly within videos featured on Spotlight and Discover.

Likelihood

A joint EUIPO-Europol report in late 2024 highlighted that the widespread use of social media and influencers is acting as a catalyst for IP crime, by directly connecting consumers seeking cheap products with sellers of counterfeit or pirated goods.⁵¹

However, Snapchat’s platform architecture is fundamentally designed to limit the mass distribution of unauthorized copyrighted content. Unlike traditional content-sharing platforms that emphasize public, persistent, and easily searchable content libraries, Snapchat prioritizes ephemeral, user-to-user communication and short-lived public content. Stories and direct Snaps are time-limited and not easily redistributable, while public-facing features like Spotlight and Discover are subject to content moderation, eligibility review, and recommendation algorithms that filter for originality and community standards. This architectural focus on transient, curated sharing - combined with robust content detection and rights management mechanisms - significantly reduces the platform’s utility for large-scale copyright infringement.

⁵¹ EUIPO & Europol, *Uncovering the Ecosystem of Intellectual Property Crime - A focus on enablers and impact*, [url](#).



Snap maintains a public [Transparency Report](#) which includes data on enforcement actions related to intellectual property infringement. According to our last Report covering the second half of 2024:

- We received 1296 copyright notices; 96,8% of those requests led to the removal of some content. This compares with 1297 copyright notices - 57% of those requests leading to the removal of some content - in the second half of 2024.
- We received 169 trademark notices; 60,9% of those requests led to the removal of some content. This compares with 203 trademark notices - 23% of those requests led to the removal of some content - in the second half of 2024.

This data shows no increases in reports of intellectual property issues and a consistently low prevalence in absolute terms. Moreover, it is important to note that the percentage of requests resulting in content removal is higher, which demonstrates Snap's prompt and effective response to such reports. Considering the above, Snap continues to consider the likelihood of encountering content that infringes intellectual property on Snapchat is within the **lowest likelihood category**.

Severity

Intellectual property infringement does not generally involve (1) harms that risks significant damage to the physical or emotional well-being of Snapchatters, or (2) imminent, credible risk of severe harm, including threats to human life, safety, and well-being. However, IP rights are protected by international, EU and national laws. IP law provides protection to anyone who creates works. This means that the creator of such a work is generally the only person allowed to decide on the exploitation of the works he has created. In addition, these works are protected from misuse by others.⁵² Infringement upon trademarks or copyrights are considered to be illegal under these laws and must therefore be prohibited. As per the findings of a 2021 report published by the European Union Intellectual Property Office (EUIPO), instances of digital piracy have demonstrated a decline across various content categories. However, it is noteworthy that despite this downward trend, the persistence of digital piracy continues to be a prevailing issue within the online domain.⁵³ As a result, Snap tracks the dissemination of content that infringes on intellectual property rights to carry a risk of **significant harm**.

DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat's in-scope services. As set out below, while the following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap's existing mitigations, which are described later in this chapter and in Chapter 5.

⁵² See for example the Berner Convention, the Rome convention, TRIPS-agreement.

⁵³ EUIPO, 'Online Copyright Infringement in the European Union', December 2021, [url](#).



Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Our recommender systems make it difficult for unvetted content to reach a broad audience without moderation. Snap's algorithmic systems do not knowingly recommend content that infringes intellectual property rights, i.e. there are no interest categories relating to specific intellectual property. Detailed information relating to our recommender systems can be found in section 5 (see the Algorithmic Systems sub-section).
(b) content moderation systems;	Snap has implemented specific proactive and reactive content moderation procedures to prevent and remove content that infringes IP rights. Users can report copyright and trademark infringement. We honor copyright laws and take reasonable steps to expeditiously remove infringing material from our Services, once we become aware of it. See also Section 5 on Moderation.
(c) the applicable terms and conditions and enforcement;	Our terms prohibit intellectual property infringements and Snap strictly enforces these terms. We provide more information in section 5 (see the Terms and Enforcement sub-sections).
(d) systems for selecting and presenting advertisements; and	Our advertising systems require agreement to advertising policies and guidance that prohibit adverts from displaying information that violates the law (i.e. including copyright law) or causes certain harms. We check advertisers are complying with their obligations via our advertising review process. Our advertising systems use a mix of automation and human review to review adverts before they are published. We provide easy mechanisms for users to hide and report advertisements that violate our policies or the law. We monitor ad rejection, reporting and enforcement data to monitor the effectiveness of our approach. We provide more information in Section 5 (see the Advertising Systems sub-section).
(e) our data related practices	We have strong data principles, practices and privacy, safety and security by design processes. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).



We have also analysed whether and how the risk of this harm is influenced by the following general factors:

General Risk Factor	How does it apply to Snapchat and this harm?
Redacted text	Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text
Redacted text	Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text
Redacted text	Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text - Redacted text



We consider the dissemination of content that infringes on intellectual property rights is a **Level 3 overall potential risk** on Snapchat. Nevertheless, we take reports seriously and the reported infringement of intellectual property often leads to content removal or, in some cases, the deletion of the user's account. There is no change in this assessment since our 2024 Report.

Snapchat enables photo and video sharing and other user interaction across Snapchat's in-scope services. While this functionality has many positive benefits for users, including in particular facilitating engagement with friends and family, Snap recognises Snapchat's functionality could also be abused by bad actors for intellectual property infringement. Snap has put in place significant measures to substantially diminish the likelihood and impact of intellectual property infringement on Snapchat.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, content on Snapchat is typically short in nature, the average Snap is 10 seconds, and reporting tools help with the detection of IP infringing material.



Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes, terms prohibit IP infringements and they are strictly enforced. If someone believes that any content on Snapchat infringes their intellectual property (IP), they can report it via Snap's reporting menu or online forms for Copyright Infringement or Trademark Infringement. Snap takes reasonable steps to expeditiously remove from our Services any infringing material that we become aware of. If Snap becomes aware that a user has repeatedly infringed copyrights, we will take reasonable steps to enforce against the violator's account, in accordance with our terms.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, specific reactive moderation procedures to expeditiously remove content that infringes intellectual property rights. Snap respects the doctrine of "fair use" (when applicable) i.e., that there are certain circumstances (such as news reporting, social commentary on issues of public interest, criticism, parody, or education) where copyrighted material could be distributed without permission from, or payment to, the copyright holder.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Our algorithmic systems do not knowingly recommend content that infringes intellectual property rights, i.e. there are no interest categories relating to specific intellectual property.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in	Yes, we have a notice procedure to flag and enable us to respond to intellectual property infringements.



particular as regards detection of systemic risk.	
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	In the EU, we have collaborated with the Austrian Institute for Applied Telecommunications, the Austrian Protection Association against unfair competition and LSG Wahrnehmung von Leistungsschutzrechten GmbH. We also collaborate with The Copyright Information and Anti-Piracy Centre (CIAPC) in Finland, Rettigheds Alliancen in Denmark, ALPA (Association de Lutte Contra la Piraterie Audiovisuelle) in France.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Not applicable. We respond to reports of infringement on an individual basis.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we warn users not to publish content that infringes on intellectual property rights and we have an easily accessible reporting tool.
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	General measures relating to the protection of minors for this risk.
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality	Yes, Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's inscope services.



which enables recipients of the service to indicate such information.	
---	--

Conclusion

We consider the overall risk of the dissemination of IP infringing content to be significant. Snap has taken steps to mitigate these harms, which has substantially diminished the likelihood that Snapchatters will encounter IP infringing material. These mitigations include product and design measures like short content retention periods, some proactive moderation, and notice-and-takedown procedures.

We have concluded therefore that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures for dissemination of content that infringes intellectual property rights. There is no change in this conclusion since our 2024 Report.

4.1.6 Dissemination of Adult Sexual Content

This Section 4.1.6 (Adult Sexual Content) considers the risk of harm arising from Adult Sexual Content (as described in the Harm Description below) on Snapchat's in-scope services defined in [Section 2](#) (Scope).

We have assessed this risk in accordance with the [Section 3](#) (Methodology) as follows.

Harm Description

This section addresses pornographic content as well as commercial activities that relate to pornography or sexual interactions (whether online or offline), including the action of sending unsolicited explicit images or videos. Breastfeeding and other depictions of adult nudity in non-sexual contexts are generally permitted.

We provide a summary and explanation of 'Adult Sexual Content' in our explainer on [Sexual Content](#) and in our [Transparency Report Glossary](#).

Likelihood

We have considered the likelihood of this harm in line with Section 3 (Methodology) and observed the following for 2025:

- Estimates as to the volume of Adult Sexual Content on the Internet vary, but some historical studies have considered that around 4% of websites, 13% of web searches and 20% of mobile searches were related to Adult Sexual Content.⁵⁴ As such it is conceivable that this content could also appear on any of Snapchat's in-scope services including in

⁵⁴ Ogas, O. and S. Gaddam (2012), Boston University, *A Billion Wicked Thoughts: What the Internet Tells Us About Sex and Relationships*; and Google Inc, Columbia University and Carnegie Mellon University, *A Large Scale Study of Wireless Search Behaviour*, 2005.



particular videos on Spotlight, Discover, promoted on Public Profiles, features as part of our Lenses or as places on the Snap Map and be the subject of advertisements via Snap Ads.

- Internet Matters - Children's Wellbeing Index Report flagged how children could be exposed to sexual content online and how there is a rise of generative AI tools that significantly increases the ease of producing realistic sexual deepfakes or nude deepfakes⁵⁵.
- External researchers⁵⁶ also recognise the role played by social media platforms to facilitate social interaction and help youth meet their relational needs, provide spaces for knowledge acquisition and skill-building, allow for identity exploration and development (particularly for marginalized youth), identity exploration and development (particularly for marginalized youth), and offer avenues for civic engagement and activism in ways that resonate with this population. However, these benefits coexist with a broad array of certain risks and harms, including inappropriate content such as sexual content.
- [REDACTED]

[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]

- Where Adult Sexual Content was identified (either proactively via automated tools or reactively following a report), our median turnaround time was rapid. Our [latest European Union Transparency Report](#), which covers the second half 2024, observed a median turnaround time for our Safety teams to take enforcement action in response to proactive or reactive detections of Adult Sexual Content content of **11 minutes**.
- [REDACTED]

⁵⁵ <https://www.internetmatters.org/hub/research/childrens-wellbeing-in-a-digital-world-index-report-2025/>
⁵⁶ <https://cyberbullying.org/empowering-protecting-youth-online-legislation-hinduja-lalani-final.pdf>



- Although Adult Sexual Content is still amongst our most prevalent illegal or violating content compared with the other risks on Snapchat’s in-scope services, we are very pleased to have been able to achieve significant further reductions in prevalence since our 2024 Report. The prevalence is significantly lower than the prevalence of adult content on the Internet in general.

We have seen further, significant reductions in the prevalence of Adult Sexual Content on the in-scope services of Snapchat since our 2024 Report. This is shown in both the PVP and VVR data, and is assisted by our rapid median turnaround times in enforcing violating Adult Sexual Content once identified. As a result, while it remains our highest prevalence harm category, Adult Sexual Content now falls within our **lowest relative likelihood** category. This is a change from our 2024 Report, when this harm category fell within our medium likelihood category and demonstrates Snap’s ongoing commitment to taking reasonable, proportionate and effective measures to mitigate risks to Snapchatters and the wider community. We continue not to identify any significant volumes of Adult Sexual Content and so we still place this in our **lowest relative likelihood** category.



Severity

The European Union Agency for Human Rights⁵⁷ flagged that sexual offences – i.e., unlawful acts that result in the sexual exploitation and sexual abuse of a person – pose a significant threat to human life and well-being. The United Nations Office⁵⁸ stated that sexual offences represent an abuse of people's fundamental rights and dignity and involve the criminal exploitation of vulnerable people. Research on the effects of pornography has shown that pornography contributes to shaping people's mindsets on sexuality and on their perceptions of gender roles. Research has also suggested a link between online pornography consumption and increased physical and/or verbal violence against women⁵⁹. The European Commission⁶⁰ also flagged that, especially in relation to Teens, seeing pornography at early ages, online pornographic content can affect their views of what constitutes a healthy relationship. According to the Council of Europe's Parliamentary Assembly⁶¹, pornography often engenders and perpetuates stereotypes and that this undermines gender equality and women's self-determination.

The risk of harm to individuals that results from Adult Sexual Content varies considerably. Sexual offences i.e. unlawful acts that result in the sexual exploitation and sexual abuse of a person⁶², pose a significant threat to human life and well-being, abusing people's fundamental rights and dignity and involving the criminal exploitation of vulnerable people.⁶³ Reports show that traffickers misuse technology throughout all stages of human trafficking and for all types of exploitation including sexual exploitation and forced labour.⁶⁴ Eurostat reports an increase in the number of registered victims of human trafficking in the EU with a 10% increase in 2021 (7155). People who have not given consent to the sharing of their sexual or nude content are exposed to unique forms of abuse including the nonconsensual distribution of their content, doxing and being coerced into performing sexual acts.⁶⁵ Snap therefore considers the severity risk of Adult Sexual Content to be **serious**.

We consider the spread of sexually explicit content or depictions of nudity in which the primary intention is sexual arousal to be significantly harmful on our platform. Research on the effects of pornography has shown that pornography contributes to shaping people's mindsets on sexuality and on their perceptions of gender roles.⁶⁶ Especially in relation to Teens, seeing pornography at early ages, online pornographic content can affect their views of what constitutes a healthy

⁵⁷ European Union Agency for Human Rights, Sanctions for sexual offences, [url](#).

⁵⁸ Europol, 'Trafficking in human beings', [url](#), retrieved on 8 August 2023.

⁵⁹ G. M. Hald, N. M. Malamuth, and C. Yuen, Pornography and attitudes supporting violence against women: Revisiting the relationship in nonexperimental studies. *Aggressive Behavior* 36 2010, [url](#). Lamb & Koven 2019, p. 4. <https://pace.coe.int/en/files/29579/html>.

⁶⁰ European Commission, A Digital Decade for children and youth: the new European strategy for a better internet for kids (BIK+) COM(2022) 212 final, p. 5. [url](#), 11 May 2022.

⁶¹ Council of Europe, Parliamentary Assembly, Gender aspects and human rights implications of pornography (Resolution Resolution 2412 (2021)), [url](#), 26 November 2021.

⁶² European Union Agency for Human Rights, *Sanctions for sexual offences*, [url](#).

⁶³ Europol, 'Trafficking in human beings', [url](#), retrieved on 8 August 2023.

⁶⁴ United Nations Office on Drugs and Crime, 'Global report on trafficking in persons 2022', [url](#), 2022, p. 70.

⁶⁵ MCSA, *Sex Workers in the Digital Era*, [url](#).

⁶⁶ M. Guggisberg, *Harms associated with online pornography consumption*, [url](#).



relationship.⁶⁷ In their resolution on violent and extreme pornography, the Council of Europe's Parliamentary Assembly has stated that pornography often engenders and perpetuates stereotypes and that this undermines gender equality and women's self-determination.⁶⁸ Research has also suggested a link between online pornography consumption and increased physical and/or verbal violence against women.⁶⁹ Content can depict the objectification, humiliation, and degradation of men and/or women, as well as explicit sex scenes involving rape.⁷⁰ As a result, we consider the severity risk of pornography to be **significant**.

DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat's in-scope services. As set out below, while the following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap's existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Our recommender systems are not designed to recommend Adult Sexual Content – i.e., there is no 'Adult Sexual Content' interest category. Detailed information regarding the mitigations relating to our recommender systems can be found in Section 5 (see the Algorithmic Systems sub-section).
(b) content moderation systems;	Snap has implemented specific proactive and reactive moderation procedures to prevent and remove Adult Sexual Content. As explained in the Content Moderation Section, Snap deploys a range of automated content moderation (which includes abusive language detection, other keyword-based detection, and machine-learning-based proactive detection) to scan media uploads. Text and symbol classifiers are trained on Adult Sexual Content signals. We provide more information in Section 5 (see Content Moderation sub-section).
(c) the applicable terms and conditions and enforcement;	Our Community Guidelines (which form part of our terms of service) prohibit Adult Sexual Content for all users (both adult and Teen accounts) and they are strictly enforced. We provide more information in Section 5 (see the Terms and Enforcement

⁶⁷ European Commission, *A Digital Decade for children and youth: the new European strategy for a better internet for kids (BIK+)* COM(2022) 212 final, p. 5. [url](#), 11 May 2022.

⁶⁸ Council of Europe, Parliamentary Assembly, *Gender aspects and human rights implications of pornography* (Resolution Resolution 2412 (2021)), [url](#), 26 November 2021.

⁶⁹ G. M. Hald, N. M. Malamuth, and C. Yuen, *Pornography and attitudes supporting violence against women: Revisiting the relationship in nonexperimental studies*. *Aggressive Behavior* 36 2010, [url](#). Lamb & Koven 2019, p. 4.

⁷⁰ A.J. Bridges, 'Pornography and sexual assault', in W.T. O'Donohue & P. A. Schewe (eds.), *Handbook of Sexual Assault and Sexual Assault Prevention*, 2019, p. 129-150.



	sub-sections).
(d) systems for selecting and presenting advertisements; and	Our advertising systems require agreement to advertising policies and guidance that prohibit adverts from displaying information that violates the law or causes certain harms. We check advertisers are complying with their obligations via our advertising review process. Our advertising systems use a mix of automation and human review to review adverts before they are published. We provide easy mechanisms for users to hide and report advertisements that violate our policies or the law. We monitor ad rejection, reporting and enforcement data to monitor the effectiveness of our approach. We provide more information in Section 5 (see the Advertising Systems sub-section).
(e) our data related practices	We have strong data principles, practices and privacy, safety and security by design processes. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).

We have also analysed whether and how the risk of Adult Sexual Content is influenced by the following general factors:

Service Risk Factor	How does it apply to Snapchat and this harm?
Intentional manipulation, including inauthentic use or inauthentic use or automated exploitation of the service	There are two key ways in which we consider our systems could be manipulated: (1) Users could seek to share novel illegal and violating material that is not detected by our automated systems. We are constantly working to adjust our systems and policies to address this. We provide more information on our approach in Section 5 (in particular the Content Moderation subsection) and Section 6 (Ongoing Risk Detection and Management). (2) Users could abuse our content moderation processes and report non-violating content / accounts in bad faith. We have processes to combat misuse and more information can be found in Section 5 (Content Moderation and Enforcement subsections).
Amplification and potentially rapid and wide dissemination of illegal content and of information that is	Snapchat's inscope services have a number of features and design configurations that act to limit the amplification and potentially rapid and wide dissemination of Adult Sexual Content, 1. Snap makes it difficult for unvetted content to reach a



incompatible with their terms and conditions.	large audience on Snapchat, and Snap proactively moderates Snapchat's in-scope services that provide an opportunity to reach a larger audience. As a result, this reduces the likelihood of our users encountering Adult Sexual Content on the platform. 2. Snapchat has also been designed to limit the prevalence of sexually suggestive content. Snap has implemented specific proactive and reactive moderation procedures to prevent and remove Adult Sexual Content. We provide more information in Section 5 (see Content Moderation). 3. Content recommended to users on Spotlight and Discover, our video sharing platforms, is moderated using a combination of auto-moderation and human moderation, and is human moderated before being widely distributed. Lens and Ads are subject to review processes before submission. We provide more information in Section 5 (see Content Moderation).
Specific regional or linguistic aspects, including when specific to a Member State.	We recognise that our users may come from different Member States and content may be shared in different languages. To address this, • Snapchat is provided in multiple EU languages and our Terms are available in all EU languages. We provide more information in Section 5 (see Terms). • We deploy content moderation systems and moderators work across multiple EU languages. We provide more information in Section 5 (see Content Moderation).

Overall potential risk prioritization

Given all forms of Adult Sexual Content now fall within our lowest relative likelihood category, the overall potential risk of this Adult Sexual Content depends primarily on severity of the issue. Overall, we consider the dissemination of sexual crimes and offenses on Snapchat's in-scope services to be a **Level 2** risk on Snapchat given the potential for serious harm. We consider the dissemination of sexually explicit content or depictions of nudity to be a **Level 3** potential overall risk prioritization. There is no change in this assessment since our 2024 Report.



Snap's Mitigations

Snapchat enables photo and video sharing and other user interaction across Snapchat's in-scope services. While this functionality has many positive benefits for users, including in particular facilitating engagement with friends and family, Snap recognises Snapchat's functionality could also be abused by bad actors for the dissemination of Adult Sexual Content. Snap has put in place significant measures to substantially diminish the likelihood and impact of this risk on Snapchat.

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services. These are organised into Snap's risk assessment mitigation categories. Note that the primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive. As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **link in the left hand column to a full summary** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Several aspects of Snapchat's design and function reduce the risk of Adult Sexual Content being shared on the platform: • Snap makes it difficult for unvetted content to reach a large audience on Snapchat, and Snap proactively moderates Snapchat's in-scope services that provide an opportunity to reach a larger audience. As a result, this reduces the likelihood of our users encountering this content on the platform. • Snapchat has also been designed to limit the prevalence of sexually suggestive content.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Snap's Terms and Community Guidelines expressly prohibit Adult Sexual Content and they are strictly enforced. Activity that involves sexual exploitation or abuse, including adult sex trafficking or sexual extortion (sextortion), is considered a "severe harm" in our Community Guidelines and we respond with swift and strict consequences against violators as explained in our Severe Harms explainer .



Mitigation Category	Applies to this risk?
	Snapchatters can report Adult Sexual Content to us via in-app reporting options and anyone can submit a report through the Snapchat Support Site. Our in-app reporting tool allows users to directly report the promotion, distribution or sharing of pornographic content as well as commercial activities that relate to pornography or sexual interactions (whether online or offline). We promptly enforce against accounts found to be sending Adult Sexual Content: • Snap removes such content from Snapchat. • Snap promptly disables accounts that we determine are dedicated to sharing such content, engage in multiple violations involving Adult Sexual Content within a defined period, or engage in a serious violation involving Adult Sexual Content. • Where appropriate, for example when there is an imminent threat to human life, safety or wellbeing, accounts engaging in violation of these policies may be reported to law enforcement.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, we have specific proactive and reactive moderation procedures to prevent and remove Adult Sexual Content. As explained in the Content Moderation section in Section 5, Snap deploys a range of automated content moderation (which include abusive language detection, other keyword-based detection, and machine-learning-based proactive detection models) to scan Stories and Spotlight submissions. Text and symbol classifiers are trained on Adult Sexual Content signals.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Unlike many of our peers, Snapchat does not offer an open news feed where unvetted publishers or individuals have an opportunity to broadcast Adult Sexual Content, does not offer a broad 'reshare' functionality that would encourage virality, and does not allow user-generated content to be recommended to a wide audience without going through further review.



Mitigation Category	Applies to this risk?
	Our algorithmic systems are not designed to recommend Adult Sexual Content – i.e., there is no ‘Adult Sexual Content’ interest category. In addition, our platform has been designed to limit the prevalence of sexually suggestive content. As explained in the Moderation Section of our Existing Mitigations in Section 5, on Spotlight and Discover, we take a proactive approach to moderating any content that may violate our rules against Adult Sexual Content prior to the content being recommended to a wide audience.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Other mitigations listed here also apply to our Advertising Systems.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	For example we have specific prevalence testing and transparency reporting which we use to help evaluate and manage Adult Sexual Content.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	We cooperate with trusted flaggers in relation to Non-Consensual Intimate Image Abuse (NCII), notably Stop Fisha in France.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms	We cooperate with other providers through various industry groups, including in particular the EU Internet Forum (EUIF) which has expanded its remit to also tackle the trafficking of human beings (which is often driven by sexual crimes or pornography).



Mitigation Category	Applies to this risk?
or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	We provide guidance on harms and how to get help in our Safety Center. We make available robust reporting tools; and we provide guidance to parents on the web (see below).
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Snap also takes steps to limit the prevalence of sexually suggestive content. For example, Spotlight has been designed not to distribute sexually suggestive content to 13 - 17 Snapchat accounts and only recommends sexually suggestive content to an 18+ Snapchat account if that content has been created by a creator that the account has subscribed to or favorites suggestive content. No viewer (who is over 18 years old) who has opted in to seeing sexually suggestive content is intended to see more than one sexually suggestive content video out of seven in their Spotlight feed. If a user hides a video labeled as sexually suggestive, we stop showing them that type of content. If a user hides a creator, we stop showing them that creator.
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables	Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services.



Mitigation Category	Applies to this risk?
recipients of the service to indicate such information.	

Conclusion

All forms of Adult Sexual Content now fall within our lowest relative likelihood category. We treat sexual crimes as a Level 2 overall risk prioritization given the risk of serious harm, and other forms of Adult Sexual Content as Level 3 overall risk prioritization given the risk of significant harm. We have continued to dedicate substantial resources and taken significant steps to mitigate these risks, including further improvements to our proactive detection mechanisms. This has resulted in further substantial decreases in the prevalence rates for Adult Sexual Content, which demonstrates the effectiveness of our ongoing risk detection and management framework and procedures. We continue to work towards further reductions in the prevalence and enforced content percentages for Adult Sexual Content.

We continue to conclude that Snapchat’s in-scope services have reasonable, proportionate and effective mitigation measures for dissemination of Adult Sexual Content. As this remains our highest prevalence violating category on Snapchat’s in-scope services, we will continue to carefully monitor to achieve further falls in prevalence.

4.1.7 Dissemination of Content regarding Harassment and Bullying

This Section 4.1.7 (Harassment and Bullying) considers the risk of harm arising from harassment, bullying and stalking (as described in the Harm Description below) on Snapchat’s in-scope services.

Snap defines harassment as any unwanted behavior that could cause an ordinary person to experience emotional distress, including bullying, verbal abuse, including certain threatening behavior. The following behaviors are examples of harassment:

- Aggressive or profane name-calling.
- Shaming or embarrassing imagery directed at a person’s physical appearance, disabilities, or cognitive abilities.
- Using a new account to contact someone after previously being blocked by them.
- Attempts at intimidation.
- “Gossip girl” types of accounts whose primary purpose is to anonymously bully or shame individuals who are typically known to the account’s followers.



Snap considers shaming behavior as a form of bullying, and considers this to include any behavior intended to embarrass or humiliate the target, or make them feel as though they have done something improper or that they are unworthy of acceptance and belonging.

Harassment and Bullying content also includes content that could cause harm to the subject if not deleted as well as content that negatively focuses on a named or visible person’s physical appearance, personal traits, cognitive or physical abilities, or economic circumstance.

We provide additional guidance and definitions in our explainer on [Harassment and Bullying](#) and via our [Transparency Report Glossary](#).

Likelihood

In our 2024 report, we highlighted improvements to our reporting flow that made it easier for Snapchatters to report Harassment and Bullying. We also broadened the scope of what can be reported. As a result, while harassment remained the second most common risk category in 2024, we observed a significant reduction in its prevalence.

We are now pleased to report that Harassment has moved down to the third most common risk category. This continued decline confirms the downward trend in the prevalence of this type of content and remains significantly lower than our highest prevalence issue.

When content related to this specific Harm was detected (whether proactively through automated tools or reactively following a user report) our Safety teams acted swiftly. According to our [latest European Union Transparency Report](#) (covering the second half of 2024), the median turnaround time for enforcement action in response to proactive or reactive detections of this content was **17 minutes**.



[REDACTED]

- [REDACTED]
[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

When we consider Snapchat’s in-scope services specifically, i.e. the public parts of Snapchat which fall within the scope of the risk assessment obligations under Article 34, the likelihood of these public spaces being used for the dissemination of bullying & harassment content is even lower.

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

As a result, in respect of Snapchat’s online platforms within the scope of our DSA risk assessment only, we consider Harassment and Bullying to now fall within our **lowest likelihood category**.

Severity

We are sensitive to the reality that Harassment and Bullying are serious issues that may undermine mental health and, in the worst instances, contribute to self-harm and suicide. Recent EU studies reveal a troubling rise in online Harassment and Bullying, with young people, women, and minorities disproportionately affected. According to the European Union Agency for Fundamental Rights (FRA), 14% of people in the EU have experienced cyber harassment or bullying in the past five years.⁷¹

⁷¹ European Union Agency for Fundamental Rights, Online Content Moderation: Current Challenges in Detecting Hate Speech, [url](#), 2023.



Moreover, and as highlighted in our prior reports, the UN Special Representative of the Secretary-General on Violence Against Children recognizes that cyberbullying can inflict substantial harm, on Teens in particular, by extending its impact through the ability of reaching a broad audience at any time. The UN considers that technology amplifies the severity of harassment and bullying. Bullying, either on- or offline, is of major concern for children globally with varying rates across different countries.⁷² Furthermore, the 2023-2024 Trends and Usage Report by CyberSafeKids reported that 21% of 12-14 year olds involved in their investigation had been “bothered or upset by something seen or experienced online”. Of this group, 41% were bothered or upset by something seen or experienced on Snapchat.⁷³

Where Harassment and Bullying involve both (1) harms that risk significant damage to the physical or emotional well-being of Snapchatters, and (2) the imminent, credible risk of severe harm, including threats to human life, safety, and well-being, we treat this as a **severe** harm. In general, Snap qualifies Harassment and Bullying as **‘serious’** in terms of severity.

DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat’s in-scope services. As set out below, while the following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap’s existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Where we algorithmically recommend content on our online platforms, we take proactive measures to stop the dissemination of content that includes Harassment and Bullying. Our algorithmic systems do not knowingly recommend Harassment and Bullying – i.e., there is no “Harassment and Bullying” interest category. Harassing content violates our Community Guidelines, is removed from Snapchat, and is therefore not eligible for algorithmic recommendation. Detailed information regarding the mitigations relating to our recommender systems can be found in Section 5 (see the Algorithmic Systems sub-section).
(b) content moderation systems;	We have specific moderation procedures to prevent and remove Harassment and Bullying content. We use a mix of automation (such as abusive language detection, image recognition models, and account history) and human review to enforce our guidelines as explained in Section 5 of this Report (see the Moderation sub-section). Content on Snapchat can be reported in-app or on

⁷² UN Special Representative of the Secretary-General on Violence Against Children, ‘*Bullying and Cyberbullying*’, [url](#).

⁷³ CyberSafeKids, Trends and Usage Report Academic Year 2023-2024, p. 4 ([URL](#)).



	our web site, and “harassment” is one of the reporting reasons offered.
(c) the applicable terms and conditions and enforcement;	Our terms prohibit harassment and bullying. This is explained to users clearly in our Harassment and Bullying explainer with guidance on how we apply this policy. In Section 5 (see sub-section on Enforcement), we also explain the significant resources devoted to preventing the dissemination of content that includes Harassment and Bullying and how we strictly enforce our terms. [REDACTED]
(d) systems for selecting and presenting advertisements; and	Other mitigations listed here also apply to our Advertising Systems in relation to content regarding harassment and bullying. We provide more information in Section 5 (see the Advertising Systems sub-section).
(e) our data related practices	We have strong data principles, practices and privacy, safety and security by design processes. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).

We have also analysed whether and how the risk of this harm is influenced by the following general factors:

General Risk Factor	How does it apply to Snapchat and this harm?
[REDACTED]	[REDACTED]





Snap is aware of the serious harm posed by Harassment and Bullying. In line with our commitment to user safety and compliance with the DSA, we have implemented a broad range of measures to significantly reduce the likelihood and impact of harassment-related offences occurring on Snapchat's in-scope services.

Mitigation Category	Applies to this risk?
Snapchat Design and Function	Yes, fundamental design decisions mean that content



Mitigation Category	Applies to this risk?
Adapting the design, features or functioning of their services, including their online interfaces.	constituting Harassment and Bullying can be easily reported.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes, our terms prohibit Harassment and Bullying. This is explained to users clearly in our Harassment and Bullying explainer with guidance on how we apply this policy. In our Enforcement section, we also explain the significant resources devoted to preventing the dissemination of content that includes Harassment and Bullying and how we strictly enforce our terms. [REDACTED]
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, we have specific moderation procedures to prevent and remove Harassment and Bullying content. We use a mix of automation (such as abusive language detection, image recognition models, and account history) and human review to enforce our guidelines as explained in the Content Moderation Section of this Report. Content on Snapchat can be reported in-app or on our web site, and “harassment” is one of the reporting reasons offered.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, where we algorithmically recommend content on our online platforms, we take proactive measures to stop the dissemination of content that includes Harassment and Bullying. Our algorithmic systems do not knowingly recommend Harassment and Bullying – i.e., there is no “Harassment and Bullying” interest category.



Mitigation Category	Applies to this risk?
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, for example we have specific prevalence testing and monitoring moderation and enforcement data which we use to help detect and manage Harassment and Bullying risk. Our Safety Advisor Board also has several anti-bullying experts which we call on for independent review and expertise.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Yes, we cooperate with many trusted flaggers in the EU. For example, Snap partners with the Bee Secure helpline in Luxembourg , with E-Enfance 3018 in France, and the Danish Centre for Digital Youth Care (CfDP).
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Snap is a member of a number of EU trade associations to contribute to the policy debate to support the development of a proportionate regulatory framework to promote online safety. Also, Snap has set up a number of crisis management protocols to help the organization swiftly tackle unexpected incidents and help minimize their impact on our service, users and operations.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on our terms, harms, moderation and enforcement practices, as well as how to get help in our Safety Center and via in-app resources (Here For You and Safety Snapshot). We have partnered with local organisations, such as the Diana Award, to raise awareness of Harassment and Bullying.
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, we have protective measures to limit teen contact with strangers; we offer Family Center ; we make available robust reporting; and we provide guidance to parents and teens, including the safety measures and resources highlighted in the Transparency mitigation section above, such as the Harassment and Bullying explainers.



Mitigation Category	Applies to this risk?
	Our parents site provides additional guidance for parents and caregivers on risks and support. ⁷⁴
<u>Content Authenticity</u> Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	General content authenticity measures taken in respect of Harassment and Bullying. Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services. This includes applying an AI sparkle icon in specific situations, such as our Bitmoji Backgrounds. We continue to assess whether to include such an icon on a case-by-case basis, considering whether generated images are photorealistic.

Conclusion

Harassment and Bullying remains one of our higher prevalence issues. There also remains significant societal concern regarding the issue of Harassment and Bullying (although this often concerns out of scope services such as private messaging). Given these factors and the severity of the harm that can result in some Harassment and Bullying cases (which we consider to be serious), we consider the dissemination of content on Snapchat's in-scope services to have a **Level 2 overall risk prioritisation level**. Note: As with other harms, where there is an imminent risk of severe harm from Harassment and Bullying, this category is treated as an overall Level 1 potential risk.

We have taken significant measures to prevent harassment and bullying, including clear guidance on our rules and how we enforce them, easy to access reporting tools and rapid response times to address violating content. Our investigations have shown the prevalence of Harassment and Bullying on Snapchat's in-scope services has continued to fall.

While we see higher reports of Harassment and Bullying relating to out-of-scope services, such as private messaging, we have confirmed that relatively few enforcement cases relate to Snapchat's in-scope services. As a result, we have moved the relative likelihood of Harassment and Bullying on Snapchat's in-scope services from medium to lowest likelihood.

⁷⁴ Snapchat Family Safety Hub, [url](#).



[REDACTED]

As a result, we have concluded that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures for dissemination of Harassment and Bullying content.

4.1.8 Dissemination of content that glorifies Self-Harm, including Suicide

This Section 4.1.8 (Self-Harm & Suicide) considers the risk of harm arising from Self-Harm and Suicide content (as described in the Description below) on Snapchat's in-scope services.

Harm Description

Self-harm and Suicide refers to content relating to suicide [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

[REDACTED]

[REDACTED]

We provide a summary and explanation of 'Self-Harm and Suicide' in our explainer on [Threats, Violence and Harm](#) and our [Transparency Report Glossary](#).

Likelihood

The risk of young people encountering content that promotes glorifies self-harm, including the promotion of self-injury, suicide, eating disorders, body image dissatisfaction and distorted values and attitudes online and on social media in general has been identified in several studies. For example, a recent study by the European Parliament's Culture and Education (CULT) Committee highlighted the risk of young people encountering such material.⁷⁵ Research published at the beginning of the year by [Reset](#) (currently led by former Hillary Clinton advisor Ben Scott) indicates that some online platforms have a high prevalence of harmful content such as pro-restrictive eating disorder and pro-suicide/self-harm material and a lack of appropriate mitigation measures.⁷⁶

⁷⁵ European Parliament, requested by the CULT Committee, 'The influence of social media on the development of children and young people', [url](#), 2023.

⁷⁶ Reset research on risks to minors and DSA compliance, 16 Jan 2024, <https://www.reset.tech/>



[REDACTED]

We have taken note of the Samaritans' Report on experiences with self-harm on social media, which found that platforms such as Instagram and TikTok often recommend self-harm content on "Explore" or "For You" around 83% of the time—without user intent.⁷⁷ While Snapchat was not mentioned, we take these signals seriously. Furthermore, A Sky article cites a statement from the Molly Rose Foundation criticizing Snap for failing to identify remove sufficient content related to suicide and self-harm. The foundation's criticism is based on Snap's Statements of Reasons sent to the Commission's Transparency API, as well as Snap's Transparency Report. It observed that only 2% of reported sSelf-Harm and Suicide content on Snapchat was taken down between January and July of last year.⁷⁸ We disagree with the foundation's findings which are based on an assumption that Snapchat poses the same risk as other platforms and therefore should be expected to have the same level of enforcement. It does not take into account the effectiveness of Snapchat's design and mitigations in preventing the wide-spread dissemination of this such violating content on Snapchat's in-scope services.

Since our 2024 Report, our [Transparency Report](#) continues to show that "Self-Harm & Suicide" is an issue that still leads to a moderate volume of content and account enforcements. In the second half of 2024, we received 307,660 reports related to Self-Harm and Suicide (up from 188,124,785 reports in the second half of 2023), enforcing against approximately 32,841 pieces of content and 13,885 accounts. However, as we noted in the 2024 Report, those figures relate to Snapchat in general. When we consider Snapchat's in-scope services specifically, i.e. the public parts of Snapchat which fall within the scope of the risk assessment obligations under Article 34, the likelihood of these public spaces being used for the dissemination of content that glorifies Self-Harm is very low. As in 2024, we rejected far fewer Snaps on Spotlight and Discover and therefore we consider the moderate enforcement rates for Harassment and Bullying primarily concern services that are out of scope of this Report.

[REDACTED]

Moreover, when content related to this specific harm was detected (whether proactively through automated tools or reactively following a user report), our Safety teams were generally able to act swiftly to verify and enforce violating content. According to our [latest European Union Transparency Report](#) (covering the second half of 2024), the median turnaround time for enforcement action in response to proactive or reactive detections of this content was **17 minutes**.

⁷⁷ [Samaritans Report \(2022\) How social media users experience self-harm and suicide content](#)

⁷⁸ News Sky, Snapchat 'asleep at the wheel when it comes to suicide and self-harm content' says childrens charity", 13 January 2025 [URL](#).



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

As a result, we consider the risk of dissemination of content glorifying Self-Harm and Suicide to fall within our **lowest likelihood category** for Snapchat’s in-scope services.

Severity

The extent of the harm risked by content glorifying self-harm is significant. In relation to content on social media in general, historical reports have indicated a particular impact this content can have on the mental health of Teens.⁷⁹ While social media and other online services may not cause self-harm, there is a potential for content glorifying self-harm to potentially trigger an existing vulnerability, such as an eating disorder.⁸⁰ We note also that in May 2023 the (US) Surgeon General’s Advisory issued a warning regarding social media and youth mental health, in particular with regards to exposure to hate-based content and suicide/self-harm-related material.⁸¹

As a result, we consider the severity of harm risked from Self-Harm and Suicide (including content relating to self-injury, suicide or eating disorders) to be “**serious**”, Where the dissemination of content that indicates an imminent, credible risk of severe harm, including threats to human life, safety, and well-being, we consider the severity of harm risked to be **severe** (as explained in our [severe harm](#) explainer) In practice, we devote enforcement resources to preventing the dissemination of content that glorifies Self-Harm, including the promotion of self-injury, suicide or eating disorders.

⁷⁹ The Guardian, ‘Facebook aware of Instagram’s harmful effect on teenage girls, leak reveals’, [url](#), 14 September 2021; E. Bozola e.a., ‘The Use of Social Media in Children and Adolescents: Scoping Review on the Potential Risks’, [url](#), August 2022.; A.M. Memom e.a., ‘The role of online social networking on deliberate self-harm and suicidality in adolescents: A systematized review of literature’, [url](#), October 2018.

⁸⁰ The New York Times, ‘Eating Disorders and Social Media Prove Difficult to Untangle’, [url](#), October 2021. Note Group Chats are private messaging groups, and do not qualify as an online platform.

⁸¹ <https://www.hhs.gov/sites/default/files/sq-youth-mental-health-social-media-advisory.pdf>



DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat's in-scope services. As set out below, while the following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap's existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	On Snapchat's services where we algorithmically recommend content, we take proactive measures to stop the dissemination of content that glorifies Self-Harm and Suicide, including the promotion of self-injury, suicide or eating disorders. Self-Harm and Suicide content violates our Community Guideline, is removed from Snapchat, and is therefore not eligible for algorithmic recommendation. Our algorithmic systems do not knowingly recommend content glorifying Self-Harm i.e. there is no 'glorifying self-harm' interest category. [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] Detailed information regarding the mitigations relating to our recommender systems can be found in Section 5 (see the Algorithmic Systems sub-section)
(b) content moderation systems;	We have specific proactive and reactive moderation procedures to prevent and remove content that promotes Self-Harm and Suicide. We use a mix of automation (such as abusive language detection, image recognition models, and account history) and human review to enforce our guidelines. We continue to make improvements to our proactive detection tools, including with respect to Self-Harm and Suicide content, to help ensure a low or negligible level of Self-Harm and Suicide content on the surfaces on Snapchat that allow users to share Public Content. We also proactively scan Limited Broadcast Content in Friends Stories for certain severe harms including Self-Harm content. Content on Snapchat can be reported in-app or on our web site, and "Self-Harm and Suicide" is one of the reporting reasons offered. Snapchatters can also report Self-Harm and Suicide content to us via in-app reporting options and anyone can submit a report through the Snapchat Support Site.



	We provide more information in Section 5 (see Content Moderation sub-section).
(c) the applicable terms and conditions and enforcement;	Snap's Terms prohibit the dissemination of content that promotes Self-Harm and Suicide. [REDACTED] [REDACTED] [REDACTED] [REDACTED] ■ [REDACTED] ■ [REDACTED] ■ [REDACTED] ■ [REDACTED] [REDACTED] [REDACTED] We promptly enforce against accounts found to be sending prohibited Self-Harm and Suicide content and Snap removes such content for all users. [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]
(d) systems for selecting and presenting advertisements; and	Our advertising systems require agreement to advertising policies and guidance that prohibit adverts from displaying information that violates the law or causes certain harms. We check advertisers are complying with their obligations via our advertising review process. Other mitigations listed here apply to our Advertising Systems too. We provide more information in Section 5 (see the Advertising Systems sub-section).
(e) our data related practices	We have strong data principles, practices and privacy, safety and security by design processes. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).

We have also analysed whether and how the risk of this harm is influenced by the following general factors:

Service Risk Factor	How does it apply to Snapchat and this harm?
---------------------	--



Intentional manipulation, including inauthentic use or inauthentic use or automated exploitation of the service	There are two key ways in which we consider our systems could be manipulated: (1) Users could seek to share novel Self-Harm and Suicide material that is not detected by our automated systems. We are constantly working to adjust our systems and policies to address this. We provide more information on our approach in Section 5 (in particular the Content Moderation subsection) and Section 6 (Ongoing Risk Detection and Management). (2) Users could abuse our content moderation processes and report non-violating content / accounts in bad faith. We have processes to combat misuse and more information can be found in Section 5 (Content Moderation and Enforcement subsections).
Amplification and potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.	Snapchat's inscope services have a number of features and design configurations that act to limit the amplification and potentially rapid and wide dissemination of content glorifying Self-Harm and Suicide, in particular: 1. Snapchat is not an attractive platform for spreading content glorifying Self-Harm and Suicide, in particular because it is difficult to reach a broad audience and content is deleted by default. We provide more information in Section 5 (see Snapchat Design / Function subsection) 2. Snap has implemented specific proactive and reactive moderation procedures to prevent and remove such content. As explained in the Content Moderation Section, Snap deploys a range of automated content moderation (which includes abusive language detection, other keyword-based detection, and machine-learning-based proactive detection). We provide more information in Section 5 (see Content Moderation) 3. Content recommended to users on Spotlight and Discover, our video sharing platforms, is moderated using a combination of auto-moderation and human moderation, and is human moderated before being widely distributed. Lens and Ads are subject to review processes before submission. We provide more information in Section 5 (see Content Moderation)
Specific regional or linguistic aspects, including when specific to a Member State.	We recognise that our users may come from different Member States and content may be shared in different languages. To address this, • Snapchat is provided in multiple EU languages and our Terms are available in all EU languages. We provide more information in Section 5 (see Terms). • We deploy content moderation systems and moderators work across multiple EU languages. We provide more information in Section 5 (see Content Moderation).



Although the prevalence of content that glorifies Self-Harm and Suicide on Snapchat’s in-scope services is considered to be at the lowest level of all our risks, due to the potential for severe and serious harms to be caused, we have chosen to elevate the risk prioritization for these risks. Snap will always consider the dissemination of content that indicates an imminent, credible risk of severe harm, including threats to human life, safety, and well-being, as Level 1 (as explained in our [severe harm](#) explainer), and we devote significant resources to combatting this type of harm.

The risk to users and, in particular, young people encountering content that promotes or glorifies Self-Harm and Suicide, including the promotion of self-injury and suicide online and on social media is well-known. We take this issue very seriously and have put in place significant measures to substantially diminish the likelihood and impact of Self-Harm and Suicide content on Snapchat.

129



specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, several aspects of Snapchat's design and function reduce the risk of Self-Harm and Suicide activity: • Unlike many of our peers, Snap does not offer the option to live-stream or have an open news feed where unvetted publishers or individuals have an opportunity to broadcast prohibited Self-Harm and Suicide content nor does Snapchat offer a 'reshare' functionality that would encourage virality. Snap also does not allow user-generated content to be recommended for distribution to a large audience without going through human review. • Private messaging services require mutual friendship. Group chats are limited to a maximum of 201 users and are not publicly searchable.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Snap's Terms prohibit the dissemination of content that promotes self-harm and suicide. This includes: Suicide (the act of intentionally causing one's death); and depictions of threats of suicide; Suicide attempts; Suicidal ideation; Suicide encouragement/tips and Praising/glorifying self-harm. Glorifying suicide or self-harm is prohibited and includes: • Content depicting self-harm or suicide posted for shock value rather than expressing a cry for help. • Content promoting suicidality or eating disorders by glamorizing, romanticizing, or normalizing the act. • Encouraging, instructing, and/or asking users to harm themselves. We allow some discussion (such as news or public issue commentary) of self-harm and suicide. We promptly enforce against accounts found to be sending prohibited Self-Harm and Suicide content and Snap removes such content for all users. When we learn of content suggesting that there is an emergency situation involving imminent danger of death or serious



Mitigation Category	Applies to this risk?
	bodily injury involving any person, we proactively escalate the report to law enforcement as appropriate. We have processes for referring such content to the relevant law enforcement authorities in the EU and government agencies in the rest of the world.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, we have specific proactive and reactive moderation procedures to prevent and remove content that promotes self-harm. As described in the Content Moderation Section of this Report, we use a mix of automation (such as abusive language detection, image recognition models, and account history) and human review to enforce our guidelines. We continue to make improvements to our proactive detection tools, including with respect to Self-Harm and Suicide content, to help ensure a low or negligible level of Self-Harm and Suicide content on the surfaces on Snapchat that allow users to share Public Content. We also proactively scan Limited Broadcast Content in Friends Stories for certain severe harms including Self-Harm content. Content on Snapchat can be reported in-app or on our web site, and “Self-Harm and Suicide” is one of the reporting reasons offered. Snapchatters can also report Self-Harm content to us via in-app reporting options and anyone can submit a report through the Snapchat Support Site. Snap cannot use proactive moderation procedures with respect to Private Services due to our legal obligations. Chat / Group chat services are ephemeral by default. However, Snapchat does not use full end-to-end encryption and private messages will be retained and available to moderators for moderation if reported (and to law enforcement if requested/escalated).



Mitigation Category	Applies to this risk?
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, on Snapchat's services where we algorithmically recommend content, we take proactive measures to stop the dissemination of content that glorifies self-harm, including the promotion of self-injury, suicide or eating disorders. Our algorithmic systems do not knowingly recommend content glorifying self-harm i.e. there is no 'glorifying self-harm' interest category. We allow some discussion (such as news or public issue commentary) of self-harm, suicide, or eating disorders, when the discussion is not glorifying such behavior. Even so, to the extent we observe it, we mark this content as "sensitive" internally and adjust our algorithmic systems so that content recommendations are not too dense with this kind of content, as it may be troubling in excess.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, for example we have specific prevalence testing and monitoring moderation and enforcement data which we use to help detect and manage Self-Harm and Suicide.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	We do not have specific Trusted Flaggers on Suicide and Self-Harm but we do work with Trusted Flaggers on child safety which may raise this.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various industry groups. For example, we are a member of Thrive - an industry signal-sharing initiative focused on Suicide and Self-Harm content.



Mitigation Category	Applies to this risk?
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on our terms, harms, moderation and enforcement practices, as well as how to report and get help in our Safety Center . In Snapchat we provide a number of tools to users. For example: • If a user searches for certain Self-Harm and Suicide related terms we will surface our Here For You tool and may be routed to suicide helplines in their region. We work with third-party mental health groups to develop these supportive materials and resources. • Snap includes help resources within rejection reasons. For example, Snap has established a self-harm flow for Lenses, which includes escalation to the Trust & Safety team, sending help resources, and escalation to authorities. Lenses that are rejected, although few in number, include help resources within the rejection reason.
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, we have protective measures to limit Teen contact with strangers; we offer Family Center ; we make available robust reporting; and we provide guidance to parents on the web. Our parents site provides additional guidance for parents and caregivers on risks and support. ⁸²
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	General content authenticity measures are taken [REDACTED]. Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services.

⁸² Snapchat Family Safety Hub, [url](#).



Conclusion

Content that glorifies self-harm is categorized within the Extremely Low likelihood category for Snapchat's in-scope services. However, this content falls within our 'serious harm' category and as a result we have decided to categorize it as a Level 2 overall potential risk, even though our risk matrix would suggest a lower category. We always treat content relating to suicide and other situations involving imminent, credible risk of harm as a Level 1 overall potential risk. In response, we have significant dedicated mitigation measures, including clear prohibitions, guidance, proactive and reactive moderation, reporting tools, sensitive content recommendation limits and cooperation with trusted flaggers. Our prevalence rates for Self-Harm content on Snap's in-scope services continue to decline and are at very low levels. This is further supported by low violative view rates for Self-Harm content.

As a result, we have concluded that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures for dissemination of content glorifying Self-Harm (including the promotion of self-injury, suicide or eating disorders). Snap monitors this category to confirm whether further mitigating measures might be required. There is no change in this conclusion since our 2024 Report.

4.1.9 Dissemination of content relating to violent or dangerous behavior

This Section 4.1.9 (Violent or Dangerous Behaviours) considers the risk of harm arising from violent or dangerous behaviours (as described in the Harm Description below) on Snapchat's in-scope services defined in [Section 2](#) (Scope).

We have assessed this risk in accordance with the [Section 3](#) (Methodology) as follows.

Harm Description

This section addresses violence (including physical and psychological violence beyond the category of Harassment already referred to in this Section 4) and risky/dangerous behaviours & activities. This includes:

- **Violent, Threatening, or Abusive Behaviour:** This includes any content that expresses an intent to inflict serious physical or emotional harm on an individual or group, or damage to their property beyond the category of Harassment already referred to in this Section 4.
- **Threats and Intimidation:** Any form of explicit or implied threat—whether of violence, retaliation, or psychological harm. This includes content that attempts to coerce or manipulate others through fear, blackmail, or control.
- **Coercive and Controlling Behaviour:** Content that reflects or promotes coercive tactics, including manipulation, isolation, surveillance, or restriction of a person's freedom or autonomy.
- **Vigilante Activity:** This includes coordinated efforts to intimidate or take physical action against individuals or communities outside proper legal process.



- Encouraging or Engaging in Dangerous Behaviour: Content that promotes or participates in risky activities likely to be imitated, and that could result in serious harm. This includes reckless driving, unsafe challenges, or any behaviour that endangers personal or public safety.
- Glorification or Incitement of Harmful Behaviour: This includes content that glorifies or encourages violence or abuse, including towards people or animals. This includes Snaps depicting gratuitous violence, graphic injury, animal cruelty, or other forms of distressing content.

We provide a summary and explanation of ‘Violent or Dangerous Behaviours’ in our explainer of [Threats, Violence & Harm](#) & [Harassment and Bullying](#) (limited to verbal abuse and threats) and our [Transparency Report Glossary](#).

Likelihood

Research has found that risky and criminal behavior is in danger of becoming normalized among a generation of young people. The online environment, and the dissemination of content encouraging or engaging in violent or dangerous behavior, is reported to play a significant role in this trend.⁸³ Algorithms and the interactions facilitated by online platforms in general have been found to have been used by radical groups to recruit vulnerable individuals to their cause. In 2017, the UN Secretary-General addressed the problem of online violence in his report to the UN General Assembly on ‘The Safety of Journalists and the Issue of Impunity’ (A/72/290).⁸⁴ This appears to have resulted in the sharp growth of violent events and deteriorating online discourse.⁸⁵

For this Report, it is worth noting that in Q3 2023, our ongoing monitoring identified an uptick in the prevalence of violating content views for the violent and disturbing category. In response, we launched new proactive detection mechanisms to target violent and disturbing content. Immediately after launch, the team were able to enforce significantly more proactively detected Snaps daily. Consequently, as of July 2024, we were pleased to report that prevalence for violent or disturbing content (as well as for Dangerous Activities) has seen a further substantial fall [REDACTED].

Since our 2024 Report,

- We have seen further falls in the PVP levels for violent and disturbing content [REDACTED]. This is a result of our specific efforts to reduce exposure to illegal and other violating content falling within our ‘Violent and disturbing content’ category. Our

⁸³ The Guardian (D. Milmo), *Risky online behaviour ‘almost normalised’ among young people*, 5 Dec 2022, [url](#).

⁸⁴ UN General Assembly, *The safety of journalists and the issue of impunity: Report of the Secretary-General*, A/72/290, [url](#).

⁸⁵ Habib, e.a., ‘*Making a Radical Misogynist: How online social engagement with the Manosphere influences traits of radicalization*’, February 2022, [url](#).



proactive content moderation has successfully evolved to reduce the prevalence of violent and disturbing content.

- Dangerous Activities saw a recent spike in prevalence in January 2025. Dangerous Activities is a broad category. The spike was determined to be caused by our platform policy and trust and safety teams making some re-adjustments to our policy concerning enforcement of Snaps taken while driving. This re-adjustment has since been completed and prevalence data relating to dangerous activities has returned to expected levels as at 30 April 2025 (i.e. a similar low level as we observed in July 2024)

On Snapchat, our [Transparency Report](#) shows that, in the EU, Snap received 172,880 reports on Threats and Violence in the second half of 2024 with action against 17,136 of content and 903 accounts in the second half of 2024. Where Violent and Dangerous Content was identified (either proactively via automated tools or reactively following a report), our median turnaround time was rapid. Our [latest European Union Transparency Report](#), which covers the second half 2024, observed a median turnaround time for our Safety teams to take enforcement action in response to proactive or reactive detections of Violent and Dangerous content of **24 minutes**.

-

-

--	--	--	--



As a result [REDACTED]
[REDACTED] we continue to place the dissemination of content encouraging or engaging in Violent and Dangerous Behaviour within the **lowest likelihood category** relative to other risks on Snapchat’s in-scope services. We continue to note that all of the risks we track on Snapchat have a relatively low prevalence compared to the prevalence of these issues elsewhere online and offline.

Severity

The extent to which a person’s access to violence content causes violent conduct and harm has been the subject of debate for a long time. Researchers have linked certain types of internet use to increased aggressive behavior. For example, youths who perpetrated serious crimes were significantly more likely to have viewed violent online content.⁸⁶

Snap considers that the spectrum of “encouraging or engaging in violent or dangerous behavior” can vary considerably and covers a broad range of content types:

- Content relating to imminent, credible threats such as school or other mass shooting and bombing threats, although this is mainly a US-related risk and less relevant for EU users. Snap considers credible imminent threats to human life to constitute a severe harm.
- Viral “challenges” may cause injury (for example, the “Milk Crate Challenge” of 2021). Since well before the existence of social media, some people have sought out videos of other people getting hurt. This content ranges from horrifying shock content, to relatively tame comedic pratfalls and minor injuries.
- Local reports of teen users being blackmailed or extorted over Snapchat, which may spill over into real-world intimidation and violence.⁸⁷

Snap considers that these issues can vary considerably in severity, from severe to significant and our teams are trained to distinguish between harmful and harmless content. Overall, Snap considers content encouraging or engaging in Violent and Dangerous Behaviour to fall within our **significant harm** category.

DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA

⁸⁶ M.L. Ybarra, M. Diener-West, D. Markow, *Linkages between internet and other media violence with seriously violent behavior by youth*, Pediatrics, 2008/122, p. 929-937.
⁸⁷ Dutch News, *Dozens of teenagers blackmailed via Snapchat, police warn*, 20 January 2025, [url](#).



influence the risk of this harm on Snapchat's in-scope services. As set out below, while the following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap's existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Unlike many of our peers, Snapchat does not offer an open news feed where unvetted publishers or individuals have an opportunity to broadcast violent or dangerous behaviours, does not offer a broad 'reshare' functionality that would encourage virality, and does not allow user-generated content to be recommended to a wide audience without going through further review. Our algorithmic systems do not knowingly recommend violent or dangerous behaviours. Detailed information regarding the mitigations relating to our recommender systems can be found in Section 5 (see the Algorithmic Systems sub-section).
(b) content moderation systems;	We have specific proactive and reactive moderation procedures to prevent and remove content showing violent or dangerous behaviours.
(c) the applicable terms and conditions and enforcement;	Snap's Terms and Community Guidelines expressly prohibit violent or dangerous behaviour and it is strictly enforced. Certain activities that involve violent or dangerous behaviours, including sharing credible, imminent threats to human life, safety, or well-being, specific threats of violence or other serious criminal activities are a "severe harm" in our Community Guidelines and we respond with swift and strict consequences against violators as explained in our Severe Harms explainer. Snapchatters can also report violent or dangerous behaviours to us via in-app reporting options and anyone can submit a report through the Snapchat Support Site. We provide more information in Section 5 (see the Terms and Enforcement sub-sections).
(d) systems for selecting and presenting advertisements; and	Our advertising systems require agreement to advertising policies and guidance that prohibit adverts from displaying information that violates the law or causes certain harms. We check advertisers are complying with their obligations via our advertising review process. Other mitigations listed here apply to our Advertising Systems too. We provide more information in Section 5 (see the Advertising Systems



	sub-section).
(e) our data related practices	We have strong data principles, practices and privacy, safety and security by design processes. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).

We have also analysed whether and how the risk of Violent or Dangerous Behaviours is influenced by the following general factors:

General Service Risk Factor	How does it apply to Snapchat and this harm?
Intentional manipulation, including inauthentic use or inauthentic use or automated exploitation of the service	There are two key ways in which we consider our systems could be manipulated: (1) Users could seek to share novel material relating to violent or dangerous behavior that is not detected by our automated systems. We are constantly working to adjust our systems and policies to address this. We provide more information on our approach in Section 5 (in particular the Content Moderation subsection) and Section 6 (Ongoing Risk Detection and Management). (2) Users could abuse our content moderation processes and report non-violating content / accounts in bad faith. We have processes to combat misuse and more information can be found in Section 5 (Content Moderation and Enforcement sub-sections).
Amplification and potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.	Snapchat's inscope services have a number of features and design configurations that act to limit the amplification and potentially rapid and wide dissemination of content glorifying Self-Harm, in particular: 1. Snapchat is not an attractive platform for spreading content relating to violent or dangerous behavior, in particular because it is difficult to reach a broad audience and content is deleted by default. We provide more information in Section 5 (see Snapchat Design / Function subsection) 2. Snap has implemented specific proactive and reactive moderation procedures to prevent and remove such content. As explained in the Content Moderation Section, Snap deploys a range of automated content moderation (which includes abusive language detection, other keyword-based detection, and machine-learning-based proactive detection). We provide more information in



Overall potential risk prioritization



Snap's Mitigations

Snapchat enables photo and video sharing and other user interaction across Snapchat's in-scope services. While this functionality has many positive benefits for users, including in particular facilitating engagement with friends and family, Snap recognises Snapchat's functionality could allow content relating to violent or dangerous behaviour to be disseminated. Snap has put in place significant measures to substantially diminish the likelihood and impact of this content on Snapchat.

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services. These are organised into Snap's risk assessment mitigation categories. Note that the primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive. As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **[link in the left hand column to a full summary](#)** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, several aspects of Snapchat's design and function reduce the risk of Violent and Dangerous Behaviour being shared on the platform: • Snap makes it difficult for unvetted content to reach a large audience on Snapchat, and Snap proactively moderates Snapchat's in-scope services that provide an opportunity to reach a larger audience. As a result, we experience very few instances of Violent and Dangerous Behaviour on Snapchat.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes, Snap's Terms and Community Guidelines expressly prohibit Violent and Dangerous Behaviour and it is strictly enforced. Certain activities that involve Violent and Dangerous Behaviour, including sharing credible, imminent threats to human life, safety, or well-being, specific threats of violence or other serious criminal activities are a "severe harm" in our Community Guidelines and we respond with swift and strict consequences against violators as explained in our Severe Harms explainer . Snapchatters can also report Violent and Dangerous Behaviour to us via in-app reporting options and anyone can submit a report through the Snapchat Support Site. Our in-app reporting tool allows users to directly report threats, violence or dangerous behaviours. We promptly enforce against accounts found to be sharing this



Mitigation Category

Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.

Testing and adapting their algorithmic systems, including their recommender systems.

Applies to this risk?

content:

- Snap removes such content for all users.
- Accounts we discover engaging in prohibited activities will also be promptly disabled.
- Where appropriate, accounts engaging in violation of these policies are reported to law enforcement as appropriate.

Yes, we have specific proactive and reactive moderation procedures to prevent and remove content showing Violent and Dangerous Behaviour. As explained in the [Content Moderation](#) section in Section 5 of this Report, Snap deploys a range of automated content moderation (which include abusive language detection, other keyword-based detection, and machine-learning-based proactive detection models) to scan Stories and Spotlight submissions. Image classifiers are trained on violent behaviour signals.

Unlike many of our peers, Snapchat does not offer an open news feed where unvetted publishers or individuals have an opportunity to broadcast Violent and Dangerous Behaviour, does not offer a broad 'reshare' functionality that would encourage virality, and does not allow user-generated content to be recommended to a wide audience without going through further review.

[illegible]



Mitigation Category	Applies to this risk?
	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] As explained in the Content Moderation Section of our Existing Mitigations in Section 5, on Spotlight and Discover, we take a proactive approach to moderating any content that may violate these rules prior to the content being recommended to a wide audience.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, for example we have specific prevalence testing and transparency reporting which we use to help detect and manage Violent and Dangerous Behaviour.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	We do not have a specific trusted flagger on Violent and Dangerous Behaviour in general but we do engage with several trusted flaggers on specific behaviors. For example, for certain threats, abusive and coercive behaviours we cooperate with Refuge.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of	We are not working with other providers on Violent and Dangerous Behaviour specifically. However, Snap is a member of a number of EU trade associations to contribute to the policy debate to support the development of a proportionate regulatory framework to promote online safety.



Mitigation Category	Applies to this risk?
conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Also, Snap has set up a number of crisis management protocols to help the organization swiftly tackle unexpected incidents and help minimize their impact on our service, users and operations.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on harms and how to get help in our Safety Center . We make available robust reporting tools; and we provide guidance to parents on the web (see below).
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, we have protective measures to limit Teen contact with strangers; we offer Family Center, reporting, and guidance. Our new parents site provides additional guidance for parents and caregivers on risks and support.⁸⁸ In addition, In particular, we would note that sensitive content distribution is limited on both Spotlight and Discover: • In Spotlight, we limit the distribution of sensitive content based on the following rules: ◦ We do not recommend sensitive content to users under 18 by default. ◦ We do not recommend sensitive content to new users (i.e. users with less than 200 views in the past 28 days). ◦ For all other users, by default, we ensure the initial video watched in a session is not sensitive and after that we ensure that sensitive content is only shown sparingly (i.e. 1 in 7 videos). • In Discover, as in Spotlight, we limit the display of sensitive content for all users. We also do not show sensitive content to users under 18 by default and the display of sensitive content can be disabled entirely in the Family Center.
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely	Yes, Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services.

⁸⁸ <https://parents.snapchat.com>.



Mitigation Category	Applies to this risk?
appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	

Conclusion

Dissemination of content encouraging or engaging in Violent and Dangerous Behaviour on Snapchat’s in-scope services is one of our lowest likelihood category risks. We recognize that the potential harm arising from such content can be significant and we have therefore tracked this risk with an overall Level 3 potential risk rating. We devote significant resources to enforcing against truly harmful or shocking content encouraging or engaging in Violent and Dangerous Behaviour as summarized above. Our prevalence testing shows the prevalence of this type of content to be failing on Snapchat’s in-scope services.

As a result, we have concluded that Snapchat’s in-scope services have reasonable, proportionate and effective mitigation measures for dissemination of content encouraging or engaging in Violent and Dangerous Behaviour. Snap monitors this category to confirm whether further mitigating measures might be required. There is no change in this conclusion since our 2024 Report.

4.1.10 Dissemination of Harmful False Information

This Section 4.1.10 (False Information) considers the risk of harm arising from the dissemination of harmful false content (as described in the Harm Description below) on Snapchat’s in-scope services.

Harm Description

Snap describes False Information as content that is false or misleading and causes harm or is malicious. Harmful False Information may be observed in content denying tragic events (e.g. Holocaust denial), promoting unsubstantiated medical claims, impersonating others in harmful ways, misleading users about democratic processes (e.g. election times or procedures), or misrepresenting public interest issues.

[REDACTED]

[REDACTED]



[REDACTED]

- [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

We provide a summary and explanation of this harm as part of our explainer on [Harmful False or Deceptive Practices](#) and our [Transparency Report Glossary](#).

Likelihood

Our external research shows that “Fake news,” (online) “disinformation” and “deep fakes” have gained a lot of attention in the media and academic and political debate over the last years.⁸⁹ Misinformation can include publishing false news articles to win elections and the impersonation of a celebrity in order to promote products or services, promote propaganda, or simply for attention. It may be a malicious effort to harass someone. It may be an attempt to extort someone into paying to regain access to an account or online identity. Such information could conceivably be present in videos published on Spotlight and For You, promoted in Stories on Public Profiles, in places and Snaps featured on Snap Map and in Lenses published via Lens Studio. Harmful false advertising might include ads for content that mimics the appearance or function of Snapchat features or formats or political advertising with false statements and slogans regarding important societal issues.

In practice, the dissemination of harmful misinformation is still not common on Snapchat. As explained in our [Transparency Reports](#), False information continues to account for only 0.1% of the total of all content enforced on Snapchat. This figure remained steady throughout 2024. We track Impersonation separately, and it similarly accounts for a very low percentage of our enforcement actions (0.1% in the second half of 2024)) Lenses with this type of information are rarely submitted.

In our 2024 Report we observed that our [Prevalence Testing](#) showed a very low prevalence of ‘Harmful False Information’ (0.0001% PVP) in July 2024 which remains the same by April 2025. Moreover, when content related to this specific Harm was detected (whether proactively through automated tools or reactively following a user report) our Safety teams acted swiftly. According to our [latest European Union Transparency Report](#) (covering the second half of 2024), the median

⁸⁹ T. McGonagle, “‘Fake news’: False fears or real concerns?”, 35 *Netherlands Quarterly of Human Rights* (No. 4, December 2017), p. 203-209. Katie Pentney, ‘Tinker, Tailor, Twitter, Lie: Government Disinformation and Freedom of Expression in a Post-Truth Era’ (2022) 22 *Human Rights Law Review*, 1-29; Paulo Cavaliere, ‘The Truth in Fake News: How Disinformation Laws Are Reframing the Concepts of Truth and Accuracy on Digital Platforms’, *European Convention on Human Rights Law Review*, (2022) 3(4), 481-523.



[REDACTED]

turnaround time for enforcement action in response to proactive or reactive detections of this content was **2 minutes**.

[REDACTED]

[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

As a result, we continue to place dissemination of harmful misinformation into our **lowest likelihood** category relative to other risks.⁹⁰

Severity

False information does not present an *imminent* threat to human life. However, Snap recognises there is a risk that malicious false information could be harmful. This includes:

- Disinformation: content that is deliberately false and designed to cause harm;
- Misinformation: content that is false or misleading in ways that can cause harm, shared without knowledge that the information is false and harmful; and
- Mal-information: genuine information that is shared with an intent to cause harm

The common denominator of these categories of false information content is that their false nature can be difficult to identify, because content can be manipulated, disguised as credible information or even allowed under free speech protection. Because of the diverse forms that false information can take, it is more reasonable to classify Harmful False Information as "general" harm. It can be observed that in recent years more and more political initiatives are emerging to combat Harmful False Information, demonstrating the need to combat the spread of this type of content.⁹¹ To this end, the EU has launched a (Strengthened) Code of Practice on

⁹⁰ This classification is also supported by the fact that Snapchat was not included in the report issued by the European Commission: European Commission, Directorate-General for Communications Networks, Content and Technology, *Digital Services Act – Application of the risk management framework to Russian disinformation campaigns*, Publications Office of the European Union, 2023 ([url](#)).

⁹¹ European Commission, ‘*Tackling online disinformation*’, overview of initiatives ([url](#)).



Disinformation,⁹² several reputable international institutions have recognized the need for intermediaries to take action to restrict third party content in their Joint Declaration on “Fake News”, Disinformation and Propaganda.⁹³

As a result, we continue to classify “Harmful False Information” in general as having a risk of **significant harm** relative to other risks.

DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat’s in-scope services. As set out below, while the following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap’s existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Our recommender system presents political content only from trusted, verified creators and partners. Our algorithmic systems do not knowingly recommend content encouraging or engaging in misinformation i.e. there is no ‘misinformation’ interest category. We take steps to prevent content with misleading or sensationalist headlines. Detailed information regarding the mitigations relating to our recommender systems can be found in Section 5 (see the Algorithmic Systems sub-section).
(b) content moderation systems;	We use specific proactive and reactive moderation procedures to prevent and remove misinformation. In particular, Discover features content only from approved media publishers and significant content creators. Our entertainment platform for user-generated content, Spotlight, is proactively and a priori moderated before content can reach a wide audience. We provide more information in Section 5 (see Content Moderation sub-section)
(c) the applicable terms and conditions and enforcement;	Our terms prohibit misinformation. We have a specific Harmful False or Deceptive Information explainer which explains our approach to enforcement. We provide more information in Section 5 (see the Terms and Enforcement sub-sections).

⁹² European Commission, ‘2022 Strengthened Code of Practice on Disinformation’, June 2022 ([url](#)).

⁹³ Declaration by the United Nations Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe Representative on Freedom of the Media, the Organization of American States (OAS) Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples’ Rights Special Rapporteur on Freedom of Expression and Access to Information, March 2017 ([url](#)).



(d) systems for selecting and presenting advertisements; and	Ads are reviewed for false, misleading and misinformation as part of Snap's Advertising Policies. Any reviewed ad that has false, misleading or misinformation will be rejected. Our moderation team also reviews reported ads, so if an ad were to be approved for misinformation and is reported, our team will take another look. Every political, health or sensitive issue ad is reviewed by humans on the ad review team. We reject ads that contain unsubstantiated or false claims. All political ads are logged in our political ads library. We provide more information in Section 5 (see the Advertising Systems sub-section).
(e) our data related practices	We have strong data principles, practices and privacy, safety and security by design processes. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).

We have also analysed whether and how the risk of this harm is influenced by the following general factors:

General Service Risk Factor	How does it apply to Snapchat and this harm?
Intentional manipulation, including inauthentic use or inauthentic use or automated exploitation of the service	There are two key ways in which we consider our systems could be manipulated: (1) Users could seek to share novel illegal and violating material that is not detected by our automated systems. We are constantly working to adjust our systems and policies to address this. We provide more information on our approach in Section 5 (in particular the Content Moderation subsection) and Section 6 (Ongoing Risk Detection and Management). (2) Users could abuse our content moderation processes and report non-violating content / accounts in bad faith. We have processes to combat misuse and more information can be found in Section 5 (Content Moderation and Enforcement subsections).
Amplification and potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.	Snapchat's in-scope services have a number of features and design configurations that act to limit the amplification and potentially rapid and wide dissemination of Harmful False Information: (1) Snapchat is not an attractive platform for spreading Harmful False Information, in particular because it is



	difficult to reach a broad audience and content is deleted by default. We provide more information in Section 5 (see Snapchat Design / Function subsection) (2) Snap has implemented specific proactive and reactive moderation procedures to prevent and remove harmful, false and deceptive information. We provide more information in Section 5 (see Content Moderation). (3) Content recommended to users on Spotlight and Discover, our video sharing platforms, is moderated using a combination of auto-moderation and human moderation, and is human moderated before being widely distributed. Lens and Ads are subject to review processes before submission. We provide more information in Section 5 (see Content Moderation).
Specific regional or linguistic aspects, including when specific to a Member State.	We recognise that our users may come from different Member States and content may be shared in different languages. To address this, • Snapchat is provided in multiple EU languages and our Terms are available in all EU languages. We provide more information in Section 5 (see Terms). • We deploy content moderation systems and moderators work across multiple EU languages. We provide more information in Section 5 (see Content Moderation).

Overall potential risk prioritization

Harmful False Information classifies as a **Level 3 overall potential risk** on Snapchat, but is a risk that we apply significant resources to mitigate against, from the design of our platform to the ways we carefully review content before it has an opportunity to reach a wide audience. There is no change in this assessment from our 2024 Report.



Snap's Mitigations

Snap is aware of the serious harm posed by False Information content online. In line with our commitment to user safety and legal compliance, we have implemented a broad range of measures to significantly reduce the likelihood and impact of this harm occurring on Snapchat's in-scope services.

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive. As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **[link in the left hand column to a full summary](#)** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, Snapchat is not an attractive platform for spreading misinformation, in particular because it is difficult to reach a broad audience and content is deleted by default. Snap has made conscious design decisions to restrict the ability for content to go viral and limiting the remix functionality to specific content types and applying short retention to content.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes, our terms prohibit misinformation. We have a specific Harmful False or Deceptive Information explainer which explains our approach to enforcement.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, we use specific proactive and reactive moderation procedures to prevent and remove misinformation. In particular, Discover features content only from approved media publishers and significant content creators. Our entertainment platform for user-generated content, Spotlight, is proactively and a priori moderated before content can reach a wide audience.



Mitigation Category	Applies to this risk?
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, our algorithmic systems do not knowingly recommend content encouraging or engaging in misinformation i.e. there is no ‘misinformation’ interest category. We take steps to prevent content with misleading or sensationalist headlines.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems. Every political, health or sensitive issue ad is reviewed by humans on the ad review team. We reject ads that contain unsubstantiated or false claims. All political ads are logged in our political ads library.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, for example we have specific prevalence testing and monitoring moderation and enforcement data which we use to help detect and manage harmful false misinformation.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Yes, we cooperate with trusted flaggers, our trusted flaggers may also report misinformation, but this rarely happens because of the limited amount of misinformation on the platform.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Snap has not yet signed up to be a member of the EU disinformation code. We have limited exposure to the risk and use our limited resources to focus on other codes relating to risks more relevant to Snapchat’s in-scope services. However, Snap works closely with French regulator Arcom, which monitors industry action against misinformation. We have also worked closely with the Commission and other stakeholders during the recent EU elections.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on our terms, harms, moderation and enforcement practices (see corresponding annexes, as well as how to and how to get help in our Safety Center. This includes a specific Harmful False or Deceptive Information explainer.



Mitigation Category	Applies to this risk?
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, we have protective measures to limit Teen contact with strangers; we offer Family Center ; we make available robust reporting; and we provide guidance to parents on the web. Our parents site provides additional guidance for parents and carers on risks and support. ⁹⁴
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	We recognise the risk that generative AI could be used to generate harmful false misinformation, including deep fakes. Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content; and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services.

Conclusion

We recognise a risk of significant harm that could arise from Harmful False Information. In practice, Snapchat's in-scope services have very little exposure to Harmful False Information. It is one of our lowest likelihood categories of risks. As a result we track this risk as an overall significant potential risk. Snapchat has significant measures in place to prevent harmful misinformation, in particular the design and function of Snapchat's in-scope services which limits the spread of content, limits the places where user generated can reach a broader audience and targets proactive moderation at those areas to prevent harmful misinformation from becoming viral. We have a rapid response time when Harmful False Information does slip through.

As a result, we have concluded that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures for dissemination of Harmful False Information. There is no change in this conclusion from our 2024 Report.

4.1.11 Dissemination of Fraud and Spam

This Section 4.1.11 (Fraud and Spam) considers the risk of harm arising from Fraud and Spam (as described in the Harm Description below) on Snapchat's in-scope services defined in [Section 2](#) (Scope).

⁹⁴ <https://parents.snapchat.com>.



[REDACTED]

We have assessed this risk in accordance with the [Section 3](#) (Methodology) as follows.

Harm Description

Snap describes Fraud and Spam as content from users, media partners and advertisers that falls within one of the following two categories:

- [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
 - [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
 - [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
 - [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

We provide a summary and explanation of ‘Fraud and Spam’ in our explainer of [Harmful, False or Deceptive Practices](#) and our [Transparency Report Glossary](#).

Likelihood

We have considered the likelihood of this harm in line with Section 3 (Methodology) and observed the following for this 2025 Report:



- Fraud and Spam are forms of cybercrime, which is a growing problem in the European Union.⁹⁵ Online fraud takes many forms and Europol has stated that “the growing e-commerce industry will result in a parallel growth of card—not-present fraud, especially as industry measures at preventing card-present fraud become more effective.”⁹⁶ The European Securities and Markets Authority (ESMA) has also invited Snap and other online platforms, through a 2025 letter, to consider the measures identified by International Organization of Securities Commissions (IOSCO), which highlighted the global concern regarding online harm linked to financial misconduct⁹⁷.
- In addition, evidence⁹⁸ submitted to a UK Home Affairs Committee inquiry⁹⁹ into fraud by Alison Thewliss MP (SNP, Glasgow Central) quoted figures from TSB on the prevalence of fraud on different platforms. She stated that 70% of the frauds that TSB was picking up were being perpetrated on Meta —24% on Facebook and 46% on Instagram— **4% on Snapchat** and 23% across other platforms. This is further evidence that indicates that Fraud has low prevalence on Snapchat in general.
- Snap has also been closely monitoring emerging trends identified at the UK level, with the potential for these to spread across the EU. A recent research report by the Alan Turing Institute (31 March 2025)¹⁰⁰ highlights that AI-enabled crime is on the rise. While still at an early stage, there is growing evidence of a significant acceleration in such activities, particularly in areas including financial crime, child sexual abuse material, phishing, and romance scams. With respect to Snapchat specifically, this report refers to interviewee concerns about the dissemination of material on private messaging services like Snapchat. While we take these concerns very seriously, private messaging services on Snapchat are out of scope of this Report.
- With regards to Fraud and Spam in the content present on Snapchat’s in-scope services:
 - [REDACTED]
 - Our [latest European Union Transparency Report](#) shows that “Fraud and Spam” reports continue to lead to a moderate volume of content and account enforcements on our content online platforms, such as Spotlight and Discover.

⁹⁵ Europol, Cybercrime, ‘EU Policy Cycle – EMPACT’, January 2022, [url](#).

⁹⁶ Europol, ‘Payment fraud’, [url](#).

⁹⁷ https://www.esma.europa.eu/sites/default/files/2025-05/ESMA35-1872330276-2397_-_ESMA_Letter_to_Snap.pdf

⁹⁸ UK House of Commons Home Affairs Committee, Home Affairs Committee. Oral evidence: Fraud, HC 125, Wednesday 22 November 2023, [url](#).

⁹⁹ UK House of Commons Home Affairs Committee, [Committee Press Release](#).

¹⁰⁰ AI and Serious Crime Online, March 2025, <https://cetas.turing.ac.uk/publications/ai-and-serious-online-crime>



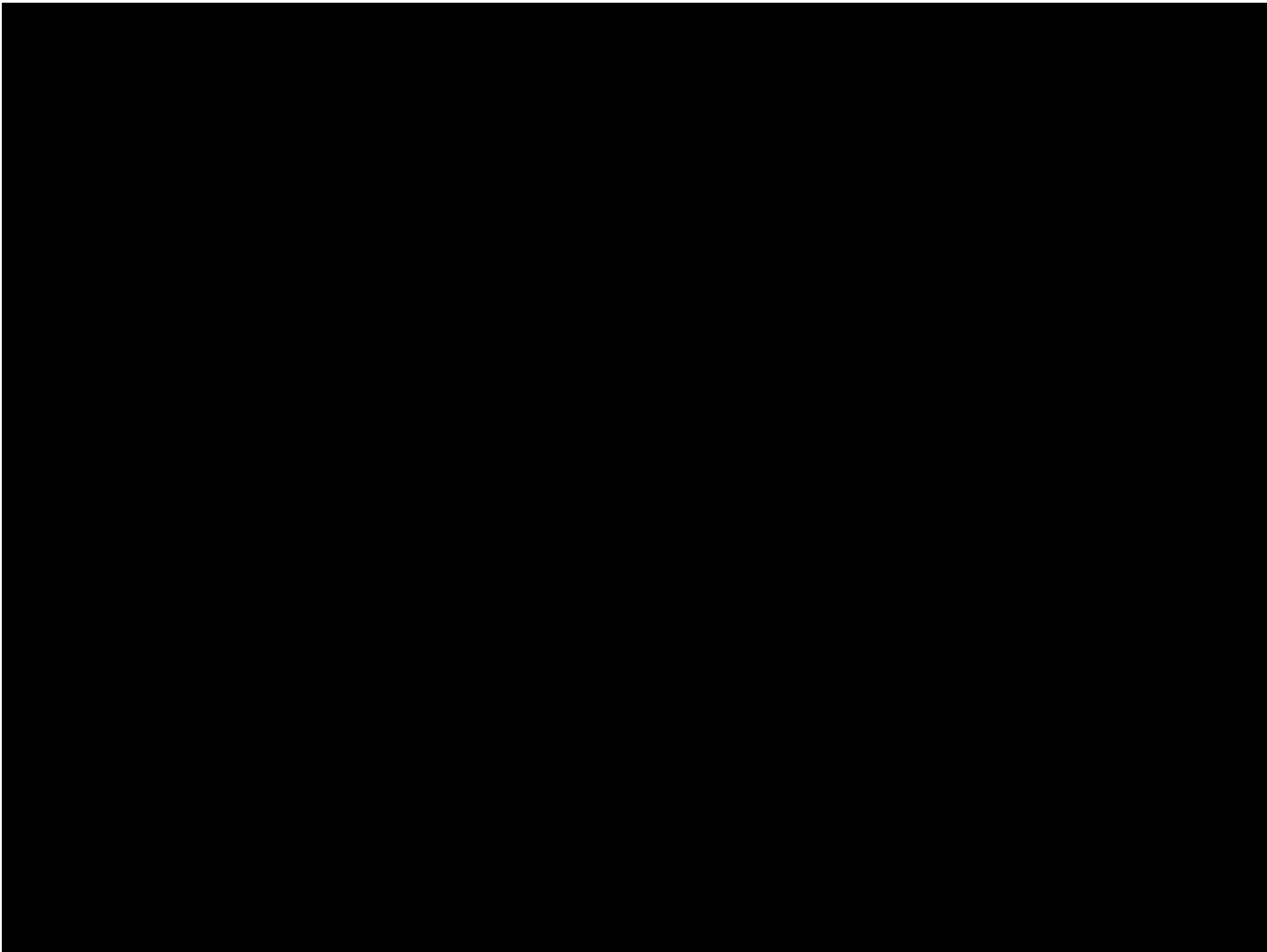
- [REDACTED]

§ 87(2)(b) [REDACTED]

§ 87(2)(b) [REDACTED]

[REDACTED]
 [REDACTED]
 [REDACTED]

-
- | Group | Bar 1 | Bar 2 | Bar 3 |
|-------|-----------|-----------|-----------|
| 1 | Very Long | Very Long | Very Long |
| 2 | Very Long | Very Long | Very Long |
| 3 | Very Long | Very Long | Very Long |
| 4 | Very Long | Very Long | Very Long |
| 5 | Very Long | Very Long | Very Long |
| 6 | Very Long | Very Long | Very Long |
| 7 | Very Long | Very Long | Very Long |
| 8 | Very Long | Very Long | Very Long |
| 9 | Very Long | Very Long | Very Long |
| 10 | Very Long | Very Long | Very Long |
| 11 | Very Long | Very Long | Very Long |
| 12 | Very Long | Very Long | Very Long |



- [Redacted text block containing multiple lines of text, some indented with bullet points]

Overall, given the further reductions in prevalence of Fraud and Spam content on Spotlight and Discover and new evidence regarding the prevalence of Fraud and Spam in advertising post-publication, we consider the dissemination of Fraud and Spam in terms of relative likelihood across all of Snapchat’s in-scope services to fall within the **lowest likelihood category**. This is a



change from our conclusion in our 2024 report where we considered that Fraud and Spam for advertising fell within the medium likelihood category based solely on ad rejection rates.

Severity

We recognise that as societies are becoming highly dependent on information and communication technologies they are also increasingly vulnerable to cybercrime.¹⁰¹ Cybercrime is a growing problem for countries, in most of which internet infrastructure is well developed and payment systems are online.¹⁰² Many of these cybercrimes involve the misrepresentation of identity. When a malicious actor impersonates an individual, this can cause financial loss, reputational loss, or cyber damage. Online fraud and forgery has also been one of the focus areas for the European Commission in means of new laws and non-legislative actions.¹⁰³

Additionally, the European Commission has stated that spam is a major phenomenon that “undermines consumer confidence in electronic communications. This represents a massive invasion of privacy; consumer fraud; an unregulated wave of harmful content received by Teens; higher business costs; lower productivity and an overall brake on the growth of the information society as a whole”.¹⁰⁴

We consider severe harm to include both (1) harms that risk significant damage to the physical or emotional well-being of Snapchatters, and (2) the imminent, credible risk of severe harm, including threats to human life, safety, and well-being. Spam is not classified as a “severe” harm according to these criteria, but that does not mean we do not take it seriously. **Fraud and Spam are classified as a “significant harm”** relative to the other risks we track.

Overall potential risk prioritization

We consider Fraud and Spam to represent one of the **Level 3 overall potential risks** compared to other more severe harms. Nevertheless, we take these risks very seriously and devote significant resources to protecting our users from Fraud and Spam wherever possible, from user-generated content to advertising. There is no change in this assessment since our 2024 Report, save to reflect the fact that Fraud and Spam in relation to both ads and content is now assessed to fall within the the Lowest Relative Likelihood category.

¹⁰¹ A. Nunzi, ‘Cybercrime: A new challenge for the European Union’, 2012, [url](#).
¹⁰² Europol, Cybercrime, ‘EU Policy Cycle – EMPACT’, January 2022, [url](#).
¹⁰³ European Commission, ‘Cybercrime’, [url](#).
¹⁰⁴ European Commission, ‘Protecting Privacy and Fighting Spam’ January 2006, [url](#).



Snap's Mitigations

Snap is aware of the threat posed by content including Fraud and Spam on internet platforms and services, and the increased inherent risk arising from the types of Risk Factors highlighted in the DSA. Snap has put in place significant measures designed to substantially diminish the likelihood and impact of Fraud and Spam on Snapchat. In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat’s in-scope services. These are organised into Snap’s risk assessment mitigation categories.

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat’s in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. Note that the primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive. As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **link in the left hand column to a full summary** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Several aspects of Snapchat’s design and function reduce the risk of Fraud and Spam being shared on the platform: • Snap makes it difficult for unvetted content to reach a large audience on Snapchat, and Snap proactively moderates Snapchat’s in-scope services that provide an opportunity to reach a larger audience. As a result, we see a relatively low volume of Fraud and Spam on Snapchat. • Snapchat is not an attractive platform for spreading Fraud and Spam, in particular because there is no e-commerce function and Snap makes it difficult for unvetted content to reach a broad audience without moderation.



Mitigation Category	Applies to this risk?
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Snap's Terms and Community Guidelines expressly prohibit Fraud and Spam and it is strictly enforced. Snapchatters can report Fraud and Spam to us via in-app reporting options and anyone can submit a report through the Snapchat Support Site. We promptly enforce against accounts found to be sharing this content: • Snap removes such content from Snapchat. • Snap promptly disables accounts that we determine are dedicated to sharing such content, engage in multiple violations involving Fraud and Spam within a defined period, or engage in a serious violation involving Fraud and Spam
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	We have specific proactive and reactive moderation procedures to prevent and remove content involving Fraud and Spam. As explained in the Content Moderation section in Section 5 of this Report, Snap deploys a range of automated content moderation features (which include abusive language detection, other keyword-based detection, and machine-learning-based proactive detection models) to scan Public Stories and Spotlight submissions. Text and symbol classifiers are trained on Fraud and Spam signals.



Mitigation Category	Applies to this risk?
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Unlike many of our peers, Snapchat does not offer an open news feed where unvetted publishers or individuals have an opportunity to broadcast content including Fraud and Spam, does not offer a broad ‘reshare’ functionality that would encourage virality, and does not allow user-generated content to be recommended to a wide audience without going through further review. Our algorithmic systems are designed not to recommend content including Fraud and Spam. As explained in the Content Moderation section in Section 5 of this Report, on Spotlight and Discover, we take a proactive approach to moderating any content that may violate our rules against Fraud and Spam prior to the content being recommended to a wide audience. In addition, when we consider whether to allow content for algorithmic recommendations, we apply additional rules. Content is “Not Eligible for Recommendation” when it contains engagement bait. This means content where the intent is not to entertain or inform the viewer, but to manipulate them to boost the Snap’s views or interactions. Engagement bait often sets up an expectation that never pays off. Some examples: • A “wait for it” caption, but “it” never happens. • Challenges based on nonexistent Snapchat features, such as, “Snapchat won’t let you like this 10 times”. • Attempts to leverage likes or shares, such as, “If this gets 20,000 likes, I’ll shave my head”. • Attempt to trick people into re-watching or pausing a Snap via long blocks of text, brief glimpses of something, or “spot the difference” games. • Misleading or sensationalized headlines.
Advertising Systems	Yes, other mitigations listed here also apply to our



Mitigation Category	Applies to this risk?
Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Advertising Systems. Snap's Advertising Policies detail the criteria that our automation and human review teams apply while considering whether to allow or reject an ad on our platform. Our advertising policies prohibit Deceptive Content, including: • Ads that are false or misleading, including deceptive claims, offers, functionality, or business practices. • Unauthorized or undisclosed sponsored content. • Promotion of fraudulent goods or services, including counterfeit documents or certificates, or counterfeit products. • Creating or sharing content that mimics the appearance or function of Snapchat features or formats. • Ads that contain deceiving calls to action, or lead to landing pages unrelated to the brand or content being advertised. • Cloaking, otherwise restricting landing page access, or modifications to URL content following submission in an attempt to circumvent review. • Ads that encourage dishonest behavior. (e.g., ads for fake IDs, plagiarism, essay writing services). • Non-delivery of goods, or misrepresented shipping delays or inventory constraints. The advertising policies for financial products and services add further detail about the kind of deceptive content that is prohibited. • Ads for financial products and services must clearly and prominently disclose all applicable material terms and conditions to consumers prior to the submission of an application. • Ads for loans must disclose, among other things, APR, repayment period, fees and costs, penalties, and the contact information of the lending institution.

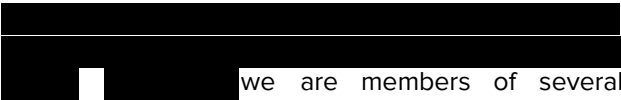


Mitigation Category	Applies to this risk?
	• Ads for products intended for a limited audience should only be targeted to that audience. For example, if a credit card offer is limited to individuals over the age of 18, the offer's ad campaign must be age targeted to 18+. • Ads for certain complex financial products, which may include cryptocurrency wallets and trading platforms, require prior approval from Snap. • We prohibit: ◦ Get-rich-quick offers, pyramid schemes, or other deceptive or too-good-to-be true financial offers (see General Requirements: Fraud for more details). ◦ Promising guaranteed financial returns on speculative investments. ◦ Ads that promote particular securities or that provide or allege to provide insider tips. ◦ Payday loans or predatory lending. For commercial promotion within content from media partners or users, we apply our Commercial Content Policy. The Commercial Content Policy outlines rules to protect Snapchatters from potentially misleading References to Snap. Commercial content must not suggest an affiliation with or endorsement by Snap or its products. This means that commercial content must not use any Snap-owned trademark, Bitmoji artwork or representations of the Snapchat user interface, except as permitted in the Snapchat Brand Guidelines or the Bitmoji Bitmoji Brand Guidelines. Commercial content must also not contain altered or confusingly similar variations of any Snap-owned trademark. The Commercial Content Policy also prohibits Deceptive Content, which includes: • False or misleading content, including deceptive claims, offers, functionality, or business practices.



Mitigation Category	Applies to this risk?
	• Promotion of fraudulent goods or services, including counterfeit documents or certificates, or counterfeit products. • Creating or sharing content that mimics the appearance or function of Snapchat features or formats. • Deceiving calls to action, or bait-and-switch links to landing pages unrelated to the brand or content being promoted. • Cloaking, otherwise restricting landing page access, or modifications to URL content following submission in an attempt to circumvent review. • Encouraging dishonest behavior. (e.g., commercial content related to fake IDs, plagiarism, essay writing services). • Non-delivery of goods, or misrepresented shipping delays or inventory constraints. • Products or services principally dedicated to selling counterfeit products, such as imitations of designer or officially-licensed products. • Products or services with false celebrity testimonials or usage. • Deceptive financial products such as, payday loans, predatory lending, insider tips relating to financial products or services, get-rich-quick offers, pyramid schemes or other deceptive or too-good-to-be true financial offers.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	For example, we have specific prevalence testing and transparency reporting which we use to help evaluate and manage Fraud and Spam. Prevalence testing is generally not used for ads since they are prescreened and there is a higher bar for bad actors for ads since it requires payment configurations.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the	We cooperate with trusted flaggers, our trusted flaggers may also report fraud spam, but this is not generally the focus of their efforts.



Mitigation Category	Applies to this risk?
implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	 we are members of several organizations and trade associations that tackle online issues facing the industry.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	We provide guidance on harms and how to get help in our Safety Center . We make available robust reporting tools; and we provide guidance to parents on the web (see below). We have also supported the StopThinkFraud campaign on Snapchat.
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	We have protective measures to limit Teen contact with strangers; we offer Family Center , reporting, and guidance. Our new parents site provides additional guidance for parents and caregivers on risks and support. ¹⁰⁵
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services.

DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat's in-scope services. As set out below, while the

¹⁰⁵ <https://parents.snapchat.com>.



following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap's existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Our recommender systems make it difficult for unvetted content, including Fraud and Spam, to reach a broad audience without moderation. Our algorithmic systems are designed not to recommend content including Fraud and Spam. Detailed information relating to our recommender systems can be found in section 5 (see the Algorithmic Systems sub-section).
(b) content moderation systems;	Snap has implemented specific proactive and reactive content moderation procedures to prevent and remove Fraud and Spam. As explained in Section 5 (see Moderation), Snap for example deploys a range of automated content moderation features (that include text and symbol classifiers that are trained on spam and fraud signals, as well as labels, filters and back end rules).
(c) the applicable terms and conditions and enforcement;	Our Community Guidelines (which form part of our service) prohibit Fraud and Spam and are strictly enforced. We provide more information in section 5 (see the Terms and Enforcement sub-sections).
(d) systems for selecting and presenting advertisements; and	Our advertising systems require agreement to advertising policies and guidance that prohibit adverts from displaying information that violates the law or causes certain harms. We check advertisers are complying with their obligations via our advertising review process. Our advertising systems use a mix of automation and human review to review adverts before they are published. We provide easy mechanisms for users to hide and report advertisements that violate our policies or the law. We monitor ad rejection, reporting and enforcement data to monitor the effectiveness of our approach. We provide more information in Section 5 (see the Advertising Systems sub-section).
(e) our data related practices	We have strong data principles, practices and privacy, safety and security by design processes. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).

We have also analysed whether and how the risk of Fraud and Spam is influenced by the following general factors:



Service Risk Factor	How does it apply to Snapchat and this harm?
Intentional manipulation, including inauthentic use or automated exploitation of the service	There are two key ways in which we consider our systems could be manipulated: • Users could seek to share novel illegal and violating material that is not detected by our automated systems. We are constantly working to adjust our systems and policies to address this. We provide more information on our approach in Section 5 (in particular the Content Moderation subsection) and Section 6 (Ongoing Risk Detection and Management). • Users could abuse our content moderation processes and report non-violating content / accounts in bad faith. We have processes to combat misuse and more information can be found in Section 5 (Content Moderation and Enforcement subsections).
Amplification and potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.	Snapchat's in-scope services have a number of features and design configurations that act to limit the amplification and potentially rapid and wide dissemination of Fraud and Spam, 1. Snapchat is not an attractive platform for spreading Fraud and Spam, in particular because it is difficult to reach a broad audience and content is deleted by default. We provide more information in Section 5 (see Snapchat Design / Function subsection) 2. Snap has implemented specific proactive and reactive moderation procedures to prevent and remove Fraud and Spam. We provide more information in Section 5 (see Content Moderation). 3. Content recommended to users on Spotlight and Discover, our video sharing platforms, is moderated using a combination of auto-moderation and human moderation, and is human moderated before being widely distributed. Lens and Ads are subject to review processes before submission. We provide more information in Section 5 (see Content Moderation).
Specific regional or linguistic aspects, including when specific to a Member State.	We recognise that our users may come from different Member States and content may be shared in different languages. To address this, • Snapchat is provided in multiple EU languages and our Terms are available in all EU languages. We provide more information in Section 5 (see Terms). • We deploy content moderation systems and moderators work across multiple EU languages. We provide more information in Section 5 (see Content Moderation).



Conclusion

We have assessed Fraud and Spam across both content and advertising to fall within our lowest likelihood category. Given the risk of significant harm arising from Fraud and Spam, we categorize this issue as a Level 3 overall risk prioritization. We handle significant volumes of enforcement and rejections every month. Our prevalent testing shows this is working, with significant further reductions in the prevalence of Fraud and Spam on public content surfaces of Snapchat. Ad rejection rates have remained consistent with the overall increase in ads reviewed. However, we see very low reports of violating ads, post-publication, including with respect to Fraud and Spam.

As a result, we have concluded that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures for Fraud and Spam. We are pleased with the further reductions in the prevalence of Fraud and Spam on Snapchat's public content surfaces and the new evidence we have gathered which demonstrate that our ad review and rejection processes are reasonable, proportionate and effective.

4.1.12 Dissemination of information related to Other Illegal Activities

This Section 4.1.12 (Other Illegal Activities) considers the risk of harm arising from other illegal content and activity (as described in the Harm Description below), beyond the illegal harm categories already referred to in this Section 4, on Snapchat's in-scope services defined in [Section 2](#) (Scope).

We have assessed this risk in accordance with the Section 3 (Methodology) as follows.

Harm Description

As we allow users to publish content, we recognise that it is possible that information related to Other Illegal Activities not already captured by our other Snap Harm Categories above may be disseminated on Snapchat.

While laws and regulations differ in jurisdictions throughout the world – and Snapchat is a global community – our policies generally prohibit any activity that undermines public safety or violates human rights, the laws of the United States or the laws of the country in which the user is located. In all cases, prohibited illegal activities will include promotion of criminal activity (including facilitation or participation in cybercrime).

Likelihood

Our [prevalence](#) measurement and [transparency reporting](#) continue to track the prevalence of known significant issues that could potentially impact online platforms, including Snapchat, as informed by our work with [Trusted Flaggers](#), [industry groups](#) and our [safety advisory board](#) and internal [cross functional working groups](#). These categories are already addressed above, as in our 2024 Report, and we are not currently aware of other significant issues.



With the introduction of the Digital Services Act, we introduced a new reporting option to report ‘illegal content’ in general. We have seen very few reports being made through this reporting option and, when it is used, the quality of the reports are very low (for example, missing key information to be able to identify the content in question and/or the illegal nature of the activity or the report concerned harmless activity) and are usually not actionable. With regards to the very small number of reports that were actionable, almost all of these related to our existing illegal and other violating content categories referred to above. We have not observed any significant new illegal activity categories.

As a result we still believe the dissemination of information related to Other Illegal Activities to fall within the **lowest likelihood category** relative to other risks identified by Snap.

Snap has assessed the possibility of illegal content and activities occurring on Snapchat that is not already covered by the Snap Harm Categories:

- We did not identify any external evidence that suggested there was a significant new category of illegal content or activity occurring on Snapchat. This includes consultation with our Safety Advisory Board as referred to in Section 6 (Ongoing Risk Detection and Mitigation).
- With regards to Internal Evidence:
 - With the introduction of the DSA, we introduced a new reporting option to report ‘illegal content’ in general. It is worth noting that we have seen relatively few reports being made through the existing DSA reporting option and we have not observed any significant new activity that is not covered by our existing Snap Harm Categories / Community Guidelines categories.
 - We have assessed our community support requests and we have not observed any significant new activity that is not covered by our existing Snap Harm Categories / Community Guidelines categories.
 - As [Section 5.10](#) shows, we do not engage trusted flaggers for illegal content in general. We receive low volumes of reports relating to illegal activity from our trusted flaggers.

Severity

The extent of harm that might be risked by information relating to Other Illegal Activities would depend on the issue. Snap has specific categories for risks concerning the dissemination of information which are most relevant to online platforms. As a result, we categorize the risk of harm in general from information relating to Other Illegal Activities as **significant**. Snap would consider the issue of illegal activity to be **severe** where the content includes a credible threat to human life, safety, or well-being.



DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat's in-scope services. As set out below, while the following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap's existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Our recommender systems are designed not to knowingly recommend any illegal or harmful content or activities. Detailed information regarding the mitigations relating to our recommender systems can be found in Section 5 (see the Algorithmic Systems sub-section)
(b) content moderation systems;	Snap has implemented specific proactive and reactive moderation procedures to prevent and remove illegal and harmful content. Detailed information regarding Snap's content moderation practices can be found in Section 5 (see Content Moderation sub-section).
(c) the applicable terms and conditions and enforcement;	Snap's Community Guidelines (which form part of our terms of service) cover a wide range of topics, including illegal or regulated activities in general. Our terms are strictly enforced. Our policies generally prohibit any activity that undermines public safety or violates human rights, the laws of the United States or the laws of the country in which the user is located. In all cases, prohibited illegal activities will include promotion of criminal activity (including facilitation or participation in cybercrime). With the introduction of the DSA, we introduced a new reporting option to report 'illegal content' in general. We provide more information in Section 5 (see the Terms and Enforcement sub-sections).
(d) systems for selecting and presenting advertisements; and	Our advertising systems require agreement to advertising policies and guidance that prohibit adverts from displaying information that violates the law or causes certain harms. We check advertisers are complying with their obligations via our advertising review process. Our advertising systems use a mix of automation and human review to review adverts before they are published. We provide easy mechanisms for users to hide and report advertisements that violate our policies or the law. We monitor ad rejection, reporting and enforcement data to monitor the effectiveness of our approach. We provide more information in Section 5 (see the Advertising Systems



We have also analysed whether and how the risk of Other Illegal Activities is influenced by the following general factors:

171



	[REDACTED] [REDACTED] [REDACTED]
[REDACTED] [REDACTED] [REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]

Overall potential risk prioritization

In general Snap assesses the overall potential risk of the dissemination of this type of content to be **Level 3** i.e. Snap’s lowest risk category compared to other risks. As in other cases, where any issue arises that poses an imminent and credible threat to human life, safety, or well-being, Snap treats this issue with a **Level 1** overall potential risk. There is no change in this assessment since our 2023 Report.

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
	[REDACTED]			

It is possible, despite Snap’s terms and policies prohibiting such practices, as well as Snap’s [Moderation](#) and [Enforcement](#) mechanisms, that malicious actors will find ways to circumvent Snap’s enforcement mechanisms and practices in order to engage in illegal activity, which could then appear on Snap’s public surfaces. Snap removes illegal content and activity as we become aware of it, cooperates with law enforcement, and disables the accounts of egregious or repeat violators.



Snap's Mitigations

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive. As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **[link in the left hand column to a full summary](#)** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, Snapchat's in-scope services have been adapted to include proactive moderation for some Other Illegal Activities.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes, our terms prohibit Other Illegal Activities and they are strictly enforced. Our legal team, supported by external counsel as needed, reviews reports of new issues to confirm illegality and appropriate enforcement action.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, general proactive and reactive moderation procedures to prevent and remove Other Illegal Activities.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, our algorithmic systems do not knowingly recommend information relating to illegal activity i.e. there is no 'illegal activity' interest category.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at	Yes, other mitigations listed here also apply to our Advertising Systems.



limiting or adjusting the presentation of advertisements in association with the service they provide.	
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	We rely on our Trusted Flaggers , industry groups and our safety advisory board and internal cross functional working groups to ensure we are prioritizing the right issues. With the introduction of the Digital Services Act, we have introduced a new reporting option to report 'illegal content' in general, and we expect to use data gathered from this option to provide us with greater visibility on the prevalence of information relating to other illegal activity on Snapchat.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Yes, we cooperate with trusted flaggers who are able to flag Other Illegal Activities.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various industry groups on prominent issues facing online platforms.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on our terms, harms, moderation and enforcement practices (see the Annex), as well as how to and how to get help in our Safety Center.
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, we have protective measures to limit Teen contact with strangers; we offer Family Center ; we make available robust reporting; and we provide guidance to parents on the web. Our parents site provides additional guidance for parents and carers on risks and support. ¹⁰⁶
Content Authenticity	No specific content authenticity measures taken in respect of

¹⁰⁶ <https://parents.snapchat.com>.



Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	information relating to illegal activities. Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's inscope services.
--	---

Conclusion

We prohibit the dissemination of information relating to illegal activities and criminal activity in our [Terms](#). We specifically track the issues relating to the dissemination of information which we consider to have the most relevance to online platforms, such as Snapchat. We treat other dissemination issues as a Level 3 overall potential risk compared to other harms. We regularly review our risk categories using our Risk Detection and Management processes. We have seen few reports using our new option to report 'other illegal activity' and have seen any new categories emerge as the vast majority are either not actionable or relate to one of our existing categories.

As a result, we have concluded that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures for information relating to Other Illegal Activities. Since our 2024 Report, we have monitored DSA enquiries, including to our new reporting option, but have so far not identified any new trends or the need to establish new illegal or harmful content categories.

4.2 Category 2: Negative Effects on Fundamental EU Rights

(Article 34.1.b / DSA Recital 81)

In this part of the Report, we explain the results of our assessment on actual or foreseeable negative effects of Snapchat's in-scope services on our Fundamental EU Rights as required by Article 34.1.b and Recital 81 of the Digital Services Act. Those Fundamental EU Rights are set out in the Charter of Fundamental Rights of the European Union (the "Charter")¹⁰⁷. We have assessed in particular the rights to human dignity, freedom of expression and of information, including media freedom and pluralism, private life, data protection, non-discrimination and consumer protection. We also consider the rights of the child, including how easy it is for Teens to

¹⁰⁷ Charter of Fundamental Rights of the European Union ([url](#)).



understand the design and functioning of the service, as well as how Teens can be exposed through their service to content that may impair Teens' health, physical, mental and moral development. Such risks may arise, for example, in relation to the design of online interfaces which intentionally or unintentionally exploit the weaknesses and inexperience of Teens or which may cause addictive behavior.

Note that for all harms, where there is (1) a risk of significant damage to the physical or emotional well-being of Snapchatters, and (2) imminent, credible risk of severe harm, including threats to human life, safety, and well-being, we treat this as a severe harm and an Level 1 overall risk prioritization.

Category 2 - Negative effects on Fundamental EU Rights				
Category	Relative likelihood of risk occurring on Snapchat	Harm classification industry wide	Risk Prioritization	Conclusion
4.2.1 Right to Human Dignity	Extremely Low Likelihood	Severe harm industry wide	Level 1	Low Risk / Reasonable, proportionate and effective
4.2.2 Right to Freedom of Expression	Extremely Low Likelihood	Significant harm industry wide	Level 3	Low Risk / Reasonable, proportionate and effective
4.2.3 Right to Private Life	Extremely Low Likelihood	Serious harm industry wide	Level 2	Low Risk / Reasonable, proportionate and effective
4.2.4 Right to Data Protection	Low Likelihood	Severe harm industry wide	Level 1	Low Risk / Reasonable, proportionate and effective
4.2.5 Right to Non-Discrimination and Freedom of Religion	Extremely Low Likelihood	Serious harm industry wide	Level 3	Low Risk / Reasonable, proportionate and effective
4.2.6 Children's Rights	Extremely Low Likelihood	Severe harm industry wide	Level 1	Low Risk / Reasonable, proportionate and effective and we are actively participating in efforts to develop



				an EU wide guidance to assess if further industry measures are needed.
4.2.7 Right to Consumer Protection	Extremely Low Likelihood	Significant harm industry wide	Level 3	Low Risk / Reasonable, proportionate and effective
4.2.8 Right to Property	N/A. Already covered under Category 1: Dissemination of content that infringes on intellectual property rights			

4.2.1 Right to Human Dignity

All public spaces displaying user generated content have the potential for the dissemination of content that may undermine Human Dignity. We recognise that without mitigation such content could conceivably appear in any of Snapchat's in-scope services displaying user generated content, from videos featured on Spotlight / Discover, to Place Stories on Snap Map. Advertising could, for example, include Hate Speech or discriminatory elements. Snapchat, as with other platforms that host user generated content, may be used to spread content that undermines respect for Human Dignity. All of the risks we track on Snapchat have a low prevalence compared to the prevalence of these issues elsewhere online and offline. To aid our prioritization, our methodology seeks to assess the relative likelihood between the risks we track (even though all are low in absolute terms).

Without mitigations, this could include content that promotes:

- Human trafficking and/or the sale of coerced sex;
- Child sexual abuse material;
- Terrorism;
- Self-harm, including the promotion of self-injury, suicide or eating disorders;
- Incitement to violence or hatred directed against a group of persons or a member of a group based on any of the grounds referred to in Article 21 of the Charter.

As Snap takes these issues very seriously and has implemented several levers to prevent this content from being distributed on the platform

Likelihood

We have assessed the relative likelihood of Snapchat's inscope services disseminating content that may undermine Human Dignity, based on Policy Violating Prevalence (PVP) via random sampling and our transparency report data in Section 4.1 of this Report, as follows:



Category	Relative likelihood of risk occurring on Snapchat
4.1.1 Dissemination of Child Sexual Abuse Material	Extremely Low
4.1.2 Dissemination of Illegal Hate Speech	Extremely Low
4.1.4 Dissemination of Terrorist Content	Extremely Low
4.1.6 Dissemination of Adult Sexual Content	Extremely Low
4.1.8 Dissemination of content that glorifies Self-Harm, including the Promotion of Self-Injury, Suicide or Eating Disorders	Extremely Low

We therefore continue to assess that the relative likelihood that the in-scope services of Snapchat would have an actual or foreseeable negative effects of the Right to Human Dignity falls within our **Extremely low likelihood category**.

Severity

The Council of Europe recognises that the misuse of social media can trigger numerous harmful consequences, including the risk threatening human dignity, and flagged examples such as Hate Speech, incitement to violence and discrimination, etc.¹⁰⁸ We have assessed the severity of harm caused by each of these categories of content that significantly undermines the Right to Human Dignity in Section 4.1 of this Report, as follows:

Category	Relative likelihood of risk occurring on Snapchat
4.1.1 Dissemination of Child Sexual Abuse Material	Severe harm industry wide
4.1.2 Dissemination of Illegal Hate Speech	Significant harm industry wide
4.1.4 Dissemination of Terrorist Content	Serious harm industry wide
4.1.6 Dissemination of Adult Sexual Content	Serious harm industry wide
4.1.8 Dissemination of content that glorifies Self-Harm, including the promotion of Self-Injury, Suicide or Eating Disorders	Serious harm industry wide

Our assessment shows a variety of harm ranging from significant to the most severe. We continue to choose to assess the category of actual or foreseeable negative effects of the Right to Human Dignity using the highest severity rating of the categories we have assessed. Several of them, including CSEAI and human trafficking, have been identified as ‘Severe Harms’ in our [Community](#)

¹⁰⁸ Council of Europe, ‘Social media: social threads or threats to human rights’ January 2017, [url](#).



Guidelines. We consider severe harm to include both (1) harms that risk significant damage to the physical or emotional well-being of Snapchatters, and (2) the imminent, credible risk of severe harm, including threats to human life, safety, and well-being. We consider these types of harms to merit a heightened level of scrutiny, as well as swift, strict, and permanent consequences for violators. Given the outsize potential for harm for some of the risks to Human Dignity, Snap continues to consider this risk to fall within the **severe** risk of harm category.

DSA Risk Factors

In accordance with Article 34(2), our risk assessment also addresses our recommender systems, content moderation systems, applicable terms and conditions, systems for the selection and presenting of advertisements and any of our data-related practices. These factors have been listed and explained in Section 3 on the Risk Assessment Methodology and applied throughout Section 4. We also considered the risk factors in the context of the Right to Human Dignity. As many of these factors pose similar risks, and are mitigated in a horizontal manner (i.e. by measures that cut across all systemic risks), we did not include them in this section again but refer to our explanation in Section 4.1.1 on Child Sexual Abuse, 4.1.2 on Hate Speech, Section 4.1.4 on Terrorist Content, Section 4.1.6 on Adult Sexual Content, and Section 4.1.8 on Self-Harm and Suicide.

Overall potential risk prioritization

Although the prevalence of content that negatively impacts users' Rights to Human Dignity has been assessed to be in the lowest likelihood category, we have assessed the severity of this risk to be severe. As a result, overall, we consider the negative effects on the Right to Human Dignity to fall within the **Level 1 overall potential risk prioritization category**. As described in our risk methodology in Section 1, we assess overall potential risk on a case by case basis and Snap reserves the option to deviate from the overall potential risk matrix we use as a guide. This is one of the cases where we have chosen to deviate. There is no change in this assessment from our 2024 Report.



Snap's Mitigations

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive.

As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **[link in the left hand column to a full summary](#)** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, Snapchat's in-scope services have been adapted to include proactive moderation for CSEAI and other illegal content that undermines Human Dignity. We also have tools within the app where individuals can report this type of activity to our Trust and Safety team. When our Trust and Safety team recognizes a Snapchatter in distress, they can forward self-harm prevention and support resources, and notify emergency personnel when appropriate. For example, if a user searches for suicide related terms we will surface our Here For You tool.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes, our Terms prohibit CSEAI and other illegal content that undermines Human Dignity and they are strictly enforced.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, specific proactive and reactive moderation procedures to prevent CSEAI and other illegal content that undermines Human Dignity. We have terms in place to prevent Media Partners from publishing illegal or harmful content on Discover. All Media Partners are vetted prior to being permitted to distribute their content broadly on Snapchat by a team of editors. Media Partners go through an editorial review of their content, a reputational search (to evaluate if a Media Partner



	has a history damaging press, legal actions, etc.), and compliance review before they're able to distribute content. Media partners are proactively moderated, and the content of their shows/editions are reactively moderated. Senior partner managers will relay feedback to Media Partners to remove or change content. If a partner refuses, we could just remove it ourselves, but partners typically comply.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, our algorithmic systems do not knowingly recommend content that would negatively affect the Right to Human Dignity i.e. there are no interest categories that we consider to negatively affect Human Dignity.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, we have specific prevalence testing and transparency reports for CSEAI, Terrorist Content, and other illegal content that undermines Human Dignity.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Yes, we cooperate with trusted flaggers in relation to CSEAI and other illegal content that undermines Human Dignity.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various industry groups [REDACTED]
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on our terms, harms, moderation and enforcement practices (see the Annex), as well as how to get help in our Safety Center. For example, if a user searches for suicide related terms we will surface our Here For You tool.



Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, we have protective measures to limit Teen contact with strangers; we offer Family Center ; we make available robust reporting; and we provide guidance to parents on the web. Our parents site provides additional guidance for parents and carers on risks and support. ¹⁰⁹
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	Some content authenticity measures have been taken in respect of content that undermines Human Dignity. Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services.

Conclusion

Snap considers risks to Human Dignity to have a Level 1 overall potential risk. In response it has put in place a range of mitigation measures. This includes in particular our proactive content [Moderation](#) which is designed to detect and prevent CSEAI from appearing on each of Snapchat's in-scope services. For example, our automated and human review on Spotlight. Our prevalence testing has allowed us to improve this proactive content moderation. As a result, we've reduced the prevalence of CSEAI and other content that may undermine human dignity on Snapchat's in-scope services to the lowest likelihood level. See [Section 4.1.1](#) (dissemination of child sex abuse material).

Similarly, dissemination of Terrorist Content is not prevalent on Snap. [REDACTED]

[REDACTED]

[REDACTED]

4.2.2 Right to Freedom of Expression and Assembly

Snapchat is a platform whose mission is to empower people to express themselves, live in the moment, learn about the world, and have fun together. By design, the platform itself presents an opportunity to enhance the freedom of expression, information and assembly of Snapchatters.

¹⁰⁹ <https://parents.snapchat.com>.



However, Snap, alongside other digital platforms hosting user-generated content, presents some risk to these rights and freedoms. [REDACTED]

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

Likelihood

All of the risks we track on Snapchat have a low prevalence compared to the prevalence of these issues elsewhere online and offline. To aid our prioritization, our methodology seeks to assess the relative likelihood between the risks we track (even though all are low in absolute terms).

It is difficult to quantify the likely risk of negative impact on Freedom of Expression and Assembly. Algorithmic biases and self-censorship are difficult to detect. We rely on user feedback and testing to flag significant incidents. At present, we are not aware of any significant bias of self-censorship issues in the algorithms used by Snapchat's in-scope services. We continue to monitor the number and nature of the general community support requests we receive and this data does not identify any trend that suggests Snapchat may be negatively impacting Freedom of Expression and Assembly. Our [Transparency Reports](#) continue to show that we receive low incidents of illegal content reports from recipients of Snapchat or authorities that we chose not to take action [REDACTED] Based on the



[REDACTED]

lack of reporting Snap has received and the [overall design of Snapchat](#) (which does not generally provide a platform for political public content in general), we deem this the **extremely low likelihood category in terms of likelihood**.

Severity

Snap has assessed information published by governments and other third party sources and considers that if content that may undermine human dignity were to materialise on an online platform it would fall within our **significant harm category**. However, Snapchat generally is not a platform for political or activist content and so the impact on freedom of expression and assembly is unlikely to be severe on Snapchat compared with other spaces on the internet dedicated to such content.

[REDACTED]

DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat’s in-scope services. As set out below, while the following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap’s existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
---------------------	--

[REDACTED]

111



(a) the design of recommender systems and any other relevant algorithmic system;	Our recommender systems are designed not to knowingly recommend content that impedes the right to Freedom of Expression and Assembly. Detailed information regarding the mitigations relating to our recommender systems can be found in Section 5 (see the Algorithmic Systems sub-section).
(b) content moderation systems;	Snap has implemented specific proactive and reactive moderation procedures to make it difficult for unvetted content to reach a large audience on Snapchat, and Snap proactively moderates Snapchat's in-scope services that provide an opportunity to reach a larger audience. We provide more information in Section 5 (see Content Moderation sub-section).
(c) the applicable terms and conditions and enforcement;	Our Community Guidelines (which form part of our terms of service) clearly define certain topics which we prohibit, including false information that threatens Public Health (e.g. COVID-19 vaccinations), civic processes, or that denies tragic events (like the Holocaust). These terms are strictly enforced given the risk of severe harm. We provide more information in Section 5 (see the Terms and Enforcement sub-sections).
(d) systems for selecting and presenting advertisements; and	Our advertising systems require agreement to advertising policies and guidance that prohibit adverts from displaying information that violates the law or causes certain harms. We check advertisers are complying with their obligations via our advertising review process. Our advertising systems use a mix of automation and human review to review adverts before they are published. We provide easy mechanisms for users to hide and report advertisements that violate our policies or the law. We monitor ad rejection, reporting and enforcement data to monitor the effectiveness of our approach. We provide more information in Section 5 (see the Advertising Systems sub-section).
(e) our data related practices	We have strong data principles, practices and privacy, safety and security by design processes. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).

We have also analysed whether and how the risk of serious Negative Effects on the Right to Freedom of Expression and Assembly is influenced by the following general factors:

General Risk Factor	How does it apply to Snapchat and this harm?
---------------------	--



Intentional manipulation, including inauthentic use or inauthentic use or automated exploitation of the service	There are two key ways in which we consider our systems could be manipulated: (1) Users could seek to share novel illegal and violating material that is not detected by our automated systems. We are constantly working to adjust our systems and policies to address this. We provide more information on our approach in Section 5 (in particular the Content Moderation subsection) and Section 6 (Ongoing Risk Detection and Management). (2) Users could abuse our content moderation processes and report non-violating content / accounts in bad faith. We have processes to combat misuse and more information can be found in Section 5 (Content Moderation and Enforcement subsections).
Amplification and potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.	Snapchat's in-scope services have a number of features and design configurations that act to limit the harms to the exercise of the Right to Freedom of Expression and Assembly, in particular: 1. Snapchat is generally not a place for political or activist public content. Such content is not eligible for promotion on Spotlight and user content on Discover is only from a small number of popular, entertaining community creators and their content is moderated by humans against our Content Guidelines. 2. Snap makes it difficult for unvetted content to reach a large audience on Snapchat, and Snap proactively moderates Snapchat's in-scope services that provide an opportunity to reach a larger audience. We provide more information in Section 5 (see Snapchat Design / Function subsection). 3. Content recommended to users on Spotlight and Discover, our video sharing platforms, is moderated using a combination of auto-moderation and human moderation, and is human moderated before being widely distributed. Lens and Ads are subject to review processes before submission. We provide more information in Section 5 (see Content Moderation)
Specific regional or linguistic aspects, including when specific to a Member State.	We recognise that our users may come from different Member States and content may be shared in different languages. To address this, • Snapchat is provided in multiple EU languages and our Terms are available in all EU languages. We provide more information in Section 5 (see Terms).





Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, several aspects of Snapchat's design and function reduce the risk for the Freedom of Expression and Assembly: • Snap makes it difficult for unvetted content to reach a large audience on Snapchat, and Snap proactively moderates Snapchat's in-scope services that provide an opportunity to reach a larger audience. • Our platform is generally not a place for political or activist public content. Such content is not eligible for promotion on Spotlight and user content on Discover is only from a small number of popular, entertaining community creators and their content is moderated by humans against our Content Guidelines. • All Media Partners are vetted prior to being permitted to distribute their content broadly on Snapchat by a team of editors. These Partners include news organizations, which are subject to their own professional rules. Media Partners go through an editorial review of their content, a reputational search (to evaluate if a publisher has a history damaging press, legal actions, etc.), and compliance review before they're able to distribute content. As a result, we provide a balanced approach to political and activist public content on Snapchat that is designed to limit the sources of such information to professional media partners. • As explained when discussing the dissemination of content that infringes on intellectual property rights, Snap respects the doctrine of "fair use," i.e., that there are certain circumstances (such as news reporting, social commentary on issues of public interest, criticism, parody, or education) where excerpts of copyrighted material could be distributed without permission from or payment to the copyright holder. This helps reinforce the rights of Freedom of Expression and the Freedom of Assembly.
Terms and Enforcement	Snap's Terms and Community Guidelines clearly define



Mitigation Category	Applies to this risk?
Adapting their terms and conditions and their enforcement.	certain topics which we prohibit, including false information that threatens Public Health (e.g. COVID-19 vaccinations), civic processes, or that denies tragic events (like the Holocaust). We expressly prohibit harmful false or deceptive information and they are strictly enforced. We respond with swift and strict consequences against violators as explained in our explainer. We promptly enforce against accounts found to be sending terrorism content: • Snap removes such content for all users. • Accounts we discover engaging in prohibited terrorist activity will also be promptly disabled. • Where appropriate, accounts engaging in violation of these policies may be reported to law enforcement.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, we have specific proactive and reactive moderation procedures to prevent and remove violative content. As explained in the Content Moderation section in Section 5 of this Report, Snap deploys a range of automated content moderation (which include abusive language detection, other keyword-based detection, and machine-learning-based proactive detection models) to scan Stories and Spotlight submissions. Snapchatters can report violative content to us via in-app reporting options and anyone can submit a report through the Snapchat Support Site. Our in-app reporting tool also allows users to directly report violative content.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, unlike many of our peers, Snapchat does not offer an open news feed where unvetted publishers or individuals have an opportunity to broadcast violative content, does not offer a broad 'reshare' functionality that would encourage virality, and does not allow user-generated content to be recommended to a wide audience without going through review. Our algorithmic systems do not knowingly recommend violative content. As explained in the Content Moderation section in Section 5 of this Report, on our high-reach



Mitigation Category	Applies to this risk?
	surfaces, like Spotlight and Discover, we take a proactive approach to moderating any content that may violate these rules prior to the content being recommended to a wide audience. The pool of content recommended by our algorithmic systems does not generally include political or other important societal matters regardless of where they fall on the political spectrum.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, for example we monitor data to help detect and manage content that may present a risk for the Right to Freedom of Expression, including data from our specific prevalence testing and enforcements (which are summarised in our Transparency Reports).
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	No, we do not work with trusted flaggers for users' rights to Freedom of Expression and access to accurate information.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, our Crisis Protocols handle issues related to users' rights to Freedom of Expression and access to accurate information. Note, we will continue to reassess and explore the opportunity to join the EU disinformation code.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on harms and how to get help in our Safety Center. We make available robust reporting tools; and we provide guidance to parents on the web (see below).



Mitigation Category	Applies to this risk?
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, we have protective measures to limit Teen contact with strangers; we offer Family Center , reporting, and guidance. Our new parents site provides additional guidance for parents and caregivers on risks and support. ¹¹²
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	Yes, Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services.

Conclusion

Snap continues to consider the overall risk to be within the level 3 risk prioritization category given the stakes and the severity of threats to Freedom of Expression, despite low prevalence and robust protections in place. Snap's mission is to be an expressive platform where users can be their authentic self, and we view our obligation to facilitate Freedom of Expression as foundational. While harms to Freedom of Expression are hard to detect, and we are not aware of any significant bias of self-censorship issues in the algorithms used by Snapchat's in-scope services, we provide avenues for our users to report these issues to us, and we value and respect user feedback. We continually evaluate and evolve our algorithms, including to reduce perceived biases, and monitor for and respond to events that could impact Freedom of Expression. We couple this with enforcement of our [Terms](#) and our robust [Moderation](#) practices to provide a platform where users feel free to express themselves in the world.

We have concluded therefore that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures to address risks relating to Freedom of Expression. Snap monitors its impact on this fundamental right category to confirm prevalence continues to decline, or whether further mitigating measures might be required. There is no change in this conclusion from our 2024 Report.

¹¹² <https://parents.snapchat.com>.



4.2.3 Right to Private Life

We understand well that online platforms can be used to spread content that undermines respect for private and family life (Right to Private Life), and that such content can have traumatic consequences if not properly mitigated. On Snapchat, without mitigations, content that undermines private and family life and personal data privacy could conceivably appear in any of Snapchat's in-scope services displaying user generated content, including information in videos featured on Spotlight / Discover and Snap Map. Snapchat's platform architecture, combined with its commitment to responsible policy enforcement across our content surfaces, establishes safeguards against negative impacts to the private life of users.

Likelihood

In our 2024 Report, we explained that:

- Our prevalence testing as at 30 July 2024 showed that “invasion of privacy” had seen a further, substantial fall [REDACTED] It is now at a very low level.
- [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED] Overall we have seen a slight increase in line with expectations and we continue to receive low numbers of privacy-related queries.¹¹³

In our 2025 Report, we observed the following:

- As of 30 April 2025, our prevalence testing indicated a slight increase in reports of “invasion of privacy,” though the overall level remains very low at 0.0109%.
- [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

We therefore continue to assess that the relative likelihood that the in-scope services of Snapchat would have an actual or foreseeable Negative Effects on the Right to Private Life falls within our **Extremely low likelihood category**.

¹¹³ Note that these numbers exclude requests we receive from automated services for data requests, and general support tickets.



Severity

Snap takes risks to privacy very seriously. Online platforms can be used to spread content that undermines respect for private and family life, home and communications, and personal data including by distributing surreptitious or non-consensual imagery of intimate activities, private movements, or identifying information; distributing deceptive content that involves the impersonation of friends, celebrities, public figures for harmful, non-satirical purposes; or content depicting Teens without parental consent. Without mitigations, content that undermines the Right to Private Life could find reach, particularly if it is of a nature of high public interest (such as celebrity voyeurism or doxxing) or if it appeals to the prurient interests of some people (e.g., certain non-consensual intimate imagery), which is why Snap enforces against this content robustly.

The severity of this risk lies not only in the immediacy of harm but in its potential to cause enduring psychological, reputational, and social damage. Victims may suffer from anxiety, fear, social isolation, or ongoing harassment - especially when content is widely circulated or permanently archived online. In the case of minors, the impact can be particularly profound, potentially affecting their development, education, and digital confidence. The nature of such violations, especially when tied to sensitive or identity-compromising content, underscores the need for swift enforcement and proactive prevention. Snap recognizes this and treats violations of privacy-related rights as a serious threat to user safety on the platform and consequently consider it to fall within our **serious harm category**.

DSA Risk Factors

In accordance with Section 3 (Methodology), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat's in-scope services. As set out below, while the following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap's existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Our recommender systems are designed not to knowingly recommend harmful content which impedes the exercise of the Right to Private Life. Detailed information regarding the mitigations relating to our recommender systems can be found in Section 5 (see the Algorithmic Systems sub-section).
(b) content moderation systems;	Snap has implemented specific proactive and reactive moderation procedures to prevent and remove harmful content which could impede the exercise of the Right to Private Life. As explained in the Content Moderation Section, Snap deploys a range of specific proactive and reactive



We have also analysed whether and how the risk of Negative Effects on the Right to Private Life is influenced by the following general factors:

194





Given the stakes and the severity of threats to Private Life, Snapchat assesses the overall risk to be within the **Level 2** category, despite low prevalence and robust protections in place. As described in our [risk methodology](#) in Section 1, we assess overall potential risk on a case by case basis and Snap reserves the option to deviate from the overall potential risk prioritization matrix we use as a guide. This is one of the cases where we have chosen to deviate. There is no change in this assessment from our 2024 Report.

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, Privacy by Design is Snap's approach to building products that consider user privacy from inception. Each product is subject to a PASS Review (Privacy Assessment System) to ensure that our products do not misuse user-data.
Terms and Enforcement	Yes, for example, our Community Guidelines prohibit impersonation, our Commercial Content Policy prohibits



Adapting their terms and conditions and their enforcement.	non-consensual sexual material and our Spotlight Terms require “ <i>you must have any necessary third-party rights including, without limitation, music copyrights and rights of publicity, for all content in your Snaps</i> ”.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, specific proactive and reactive moderation procedures to protect the privacy interests of our community. Users have the ability to report Snaps and the reporting menu includes options such as “They leaked / are threatening to leak my nudes”, , “It involves a child”
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, our algorithmic systems do not categorize or recommend content that violates users’ Right to Private Life. For example, we have terms, moderation and enforcement to prevent distribution of illegal / violating content. We also do not process sensitive category information.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems. For example, Snap ensures that ads shown are in line with its Snap Advertising Policies which states that advertisements do not collect sensitive information or special category of data. We also ensure advertisers are not targeting specific individuals on our platform and that users do not feel like their privacy is being compromised by our advertising.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, we have specific prevalence testing and Transparency Reports for sexual content and intrusion of privacy. We also monitor privacy-related inquiries as detailed above.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with	Yes, we cooperate with trusted flaggers in relation to sexual content and Teen safety which may impact users’ Right to Private Life.



Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various industry groups e.g. CIPL, FPF. Our content moderation policies provide de facto content moderation crisis protocol.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on our terms, harms, moderation and enforcement practices (see the Annex), as well as how to and how to get help in our Safety Center. Our Privacy Center offers a suite of information on our products, users' choices to safeguard their privacy and how to contact us.
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, we have protective measures in place for Teens. Our reporting menu also includes the option to report "It involves a child". Our Family Center includes resources and guidance for Teens and their parents or trusted adults. Our parents site provides additional guidance for parents and carers on risks and support.¹¹⁴
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	Yes, Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services.

¹¹⁴ <https://parents.snapchat.com>.



Conclusion

Snap continues to consider the overall risk prioritization to be within the **Level 2 category** given the stakes and the severity of threats to privacy life, despite low prevalence. [REDACTED]

[REDACTED] However, privacy is the first of Snap's four core platform governance values. We have robust protections in place, including clear terms and moderation. Snap enforces against these content violations robustly. [REDACTED]

[REDACTED] We also mitigate risks through intentional product design choices and collaborate with experts, think tanks and researchers on human rights, privacy and online safety to inform our approach.

We have concluded therefore that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures to address risks relating to the Right to Private Life. There is no change in this conclusion from our 2024 Report.

4.2.4 Right to Data Protection

We understand well the importance of ensuring that personal data is collected, processed or secured appropriately. Depending on how and the extent to which Snapchatters use our platforms, significant volumes of the content published on Snapchat's in-scope services, including on Spotlight / Discover and Snap Map, is user generated images and videos.

Snapchat recognises that rights of personal data may be infringed by excessive use of data shared with Snapchat for personalized advertising. Additionally, this right may be impacted by disseminating microtargeted ads, or ads that attempt to incite users, or by collection of sensitive information by advertisers.

In its research on online advertising, the OECD identifies misleading advertising online as a particular concern, and flags that consumers may not be able to identify some forms of online advertising and that this is a significant issue as misleading online advertising could reduce consumer sentiment and trust online.¹¹⁵ The research also raises concern that online advertising may prey on consumer biases and vulnerabilities to sell and in extreme cases can threaten individuals through malicious advertising. It also raises concern that individuals may suffer harm from increased data collection associated with the most aggressive forms of online advertising.

¹¹⁵ OECD, 'Online advertising - Trends, benefits and risks for consumers', January 2019, [url](#).



These issues are echoed in research by the EP in which privacy risks are found to arise from excessive behavioral targeting.¹¹⁶ The research considered that obfuscation and location based targeting of advertisements reduce consumer choice. Excessive behavioral targeting techniques including exploiting consumers behavior biases via the use of dark patterns. The research raised the concern for discrimination and harmful targeting of vulnerable consumers through Real-Time Bidding (RTB). It considered that data-driven ad distribution through RTB processes had the potential for structural discrimination or harmful targeting of vulnerable consumers. For example, if past browsing behavior relating to sensitive matters were used to target advertising e.g. bad eating habits or a potential addiction to drinking or gambling, this could increase exposure of vulnerable persons to advertisement in these fields. This could pose a serious risk to consumers' wellbeing. The research also highlighted the increasing prevalence of malvertising and spearphishing which threatens cybersecurity.

An article¹¹⁷ published by the International Association of Privacy Professionals indicates that consumers are increasingly protective of their personal data. As revealed by the IAPP [Privacy and Consumer Trust Report 2023](#), 68% of consumers globally are either somewhat or very concerned about their privacy online. The recent popularization of generative AI tools was highlighted as one of the newest factors to drive these concerns, with 57% of consumers globally agreeing that AI poses a significant threat to their privacy.

Additionally, this right might be impacted by the dissemination of harmful ads, for example ads that use Hate Speech or contain discriminatory elements, or targets a specific audience based on discriminatory parameters. Also, the design of the algorithmic systems used for advertising could harm human dignity by inadvertently targeting specific groups by basing targeting decisions on biased data. Without mitigations, these ads and related systems could severely impact human dignity.

Likelihood

All of the risks we track on Snapchat have a low prevalence compared to the prevalence of these issues elsewhere online and offline. To aid our prioritization, our methodology seeks to assess the relative likelihood between the risks we track (even though all are low in absolute terms).

Snapchat handles a significant volume of personal data relating to individuals in the European Union. Depending on how and the extent to which Snapchatters use Snapchat, this could be limited to basic account information or it could extend, for example, to published images and videos and metadata about the Snapchatter's interaction with such content. Significant volumes of the content published on Snapchat's in-scope services is user generated images and videos which might be related to individual Snapchatter creators and/or others. It is therefore more likely

¹¹⁶ N. Fourberg e.a., on 'Online advertising: the impact of targeted advertising on advertisers, market access and consumer choice', 2021, [url](#).

¹¹⁷ Consumer Perspectives of Privacy and Artificial Intelligence, February 2024, [url](#).



than not that Snapchat's in-scope services could cause an impact on an individual's data protection rights if such personal data is not collected, processed or secured appropriately.

In practice however, we continue to observe a low volume of data protection issues on Snapchat:

- Our prevalence testing showed that the Policy Violating Prevalence of "invasion of privacy" content (as per Section 6.4 on [Prevalence Testing](#) chapter) has been maintained at an extremely low level through 2024 and the first quarter of 2025:



- Snap also receives consistently low numbers of privacy-related queries from users and the broader community:



However, despite low prevalence of privacy issues and low numbers of privacy related queries from our users, a significant volume of personal data is being processed by Snap in relation to the in-scope services on Snapchat. As a result, we continue to assess this risk falls within our **Low Likelihood category**.

Severity

The impact on an individual's data protection rights caused by a data security breach or other incident involving inappropriate processing of personal data would depend on the nature of the information involved. In respect of a security breach, the European Data Protection Board in its Guidelines 01/2021 on Examples regarding Personal Data Breach (14 December 2021)¹¹⁸ provides a range of examples relevant to Snapchat such as: (1) on one hand, a Ransomware attack with proper backup and without exfiltration giving rise to only minor consequences and no significant effect on data subjects; and (2) on the other hand, highly confidential personal data sent by mistake giving rise to high risk to data subjects.

¹¹⁸ European Data Protection Board, 'Guidelines 01/2021 on Examples regarding Personal Data Breach', 14 December 2021, [url](#).



Regarding Snap's processing of personal data, Snapchat's in-scope services do not fall within the three examples in Article 35.3 of the General Data Protection Regulation 2016/679 (GDPR) of high risk processing. Snap also does not fall within the high risk profiling example noted in Guidelines WP 248 of the Article 29 Working¹¹⁹ concerning large scale profiling by social media companies with sensitive or highly confidential information. Nevertheless Snap considers there is a potential for its processing of personal data in relation to Snapchat's in-scope services to cause severe impact on data protection rights, if not designed appropriately.

Noting the above, Snap continues to qualify the severity of this risk as ‘**severe**’.

DSA Risk Factors

In accordance with Article 34(2), our risk assessment also addresses our recommender systems, content moderation systems, applicable terms and conditions, systems for the selection and presenting of advertisements and any of our data-related practices. These factors have been listed and explained in Section 3 on the Risk Assessment Methodology and applied throughout Section 4. We also considered the risk factors in the context of the Right to Data Protection. As many of these factors pose similar risks, and are mitigated in a horizontal manner (i.e. by measures that cut across all systemic risks), we did not include them in this section again but refer to our explanation in Sections 4.2.3 on the Right to Private Life.

Overall potential risk prioritization

Considering the extent of the personal data being processed by Snapchat and our assessment of the risk of severe harm from Negative Effects on the Right to Data Protection within the European Union, we have assessed this to be a **Level 1 overall potential risk prioritization**, notwithstanding the low incident of privacy related queries from recipients for Snapchat's in-scope services. There is no change in this assessment from our 2024 Report.

[illegible]

¹¹⁹ Article 29 Data Protection Working Party, *Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is “likely to result in a high risk” for the purposes of Regulation 2016/679*, 13 October 2017, [url](#).



Snap's Mitigations

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive.

As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **link in the left hand column to a full summary** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, Snapchat is a platform with strong privacy principles. These principles are reflected into the architecture of our platform. Product changes are subject to privacy by design reviews and we maintain data protection impact assessments. For example, our Lenses only require object detection rather than facial identification. Lenses can tell what is or isn't a face, they do not identify specific faces, limiting data processing for the use of Lenses. Snap does not use any data collected by Lenses to customize the content that the user sees in Spotlight or Discover, nor is any data collected for advertising purposes. Besides, voice data collection of Snapchatters in the EU is off by default; it is only used to provide the service.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes, our Privacy Center provides a suite of policies, including our Privacy Policy and they are enforced. In our Content Guidelines for Recommendation Eligibility we inform creators "We inform these standards with proactive moderation using technology and human review" and "you must have any necessary third-party rights including, without limitation, music copyrights and rights of publicity, for all content in your Snaps" This prevents any risk that users may not be aware that their content submitted to Spotlight is subject to automated and human review, and prohibits creators from depicting individuals in content without necessary rights.



	In our Snap Spotlight Submission and Revenue Terms we state “You understand that Snaps you submit to Spotlight are Public Content and may be visible to all Snapchat users, as well as non-Snapchat users on other services and websites”. This prevents the risk of creators being unaware that their Stories submitted to Spotlight become public and informs users that their content may be saved off Snapchat.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, specific proactive and reactive moderation procedures to prevent and remove content that violates users’ Right to Data Protection. For example, on Discover, Media Partners are proactively moderated and only a small pool of Snapchatters are shown in Discover (“Snap Stars” or “Popular Users”).
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, our algorithmic systems do not categorize or recommend content that violates users’ Right to Data Protection. For example, users can opt out from personalized recommendations based on inferred interest and we do not process sensitive category information.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, our Advertising Systems has a suite of protections including: • No microtargeting • We offer controls to turn off most personalized ads. Users can learn more about their choices here How to Adjust My Advertising & Interest Preferences on Snapchat. • We ensure that sensitive data is not being used for ad targeting • We continue to trial evolving privacy enhancing technologies, such as third party data clean rooms, to provide advertisers with options to further minimize the privacy impact of Snap ad services.
Ongoing Risk Detection and Management	Yes, we consult with experts and our community, and we also monitor and respond to privacy-related inquiries.



Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	No, we don't cooperate with trusted flaggers in relation to data protection violations.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various industry groups e.g. CIPL, FPF. We also have a well-established protocol to deal with privacy incidents, as well as a Security Incident Response Policy.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on privacy protection in our Privacy Center. For example, we explain to users How We Rank Content in Discover , How We Rank Content on Spotlight – Snapchat Support and Snapchat Ads Privacy & Transparency .
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, we have protective measures to limit disclosure of Teens' data and we avoid nudge techniques to encourage Teens to change their privacy settings and select less privacy-enhancing choices. We offer Family Center ; we make available robust reporting; and we provide guidance to parents on the web. Our parents site provides additional guidance for parents and carers on risks and support. ¹²⁰
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or	No specific content authenticity measures taken in respect of users' Right to Data Protection.

¹²⁰ <https://parents.snapchat.com>.



events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	
--	--

Conclusion

Snap considers the likelihood and the serious nature of the impacts on the Right to Data Protection within the European Union to fall within our Level 3 overall potential risk of Snapchat’s in-scope services, despite robust protections in place. Depending on how and the extent to which Snapchatters use these platforms, significant volumes of the content published on Snapchat’s in-scope services is user generated images and videos. It is therefore more likely than not that Snapchat’s in-scope services could negatively affect an individual’s data protection rights if such personal data is not collected, processed or secured appropriately, which is why Snap enforces its [privacy principles](#) robustly. Privacy is central to Snapchat’s values. We put significant thought and consideration into our [privacy principles](#) and those principles are reflected into the architecture of our platform. We have a cross-functional group responsible for compliance with our [privacy and safety by design principles](#), we review product changes for impact to data protection rights and we maintain Data Protection Impact Assessments of our processing of personal data where appropriate to ensure we are confident this will not result in a high risk to the rights and freedoms of individuals. We receive a low level of data protection queries as a result of the robust protections in place.

We have concluded therefore that Snapchat’s in-scope services have reasonable, proportionate and effective mitigation measures against the risk of Negative Effects on the Right to Data Protection. There is no change in this conclusion from our 2024 Report.

4.2.5 Right to Non-Discrimination and Freedom of Religion

We understand well that online platforms can be used to spread content that contains or promotes discrimination for example by using discriminatory characteristics for targeting ads, biased algorithms used for recommender systems and content moderation, the spread of discriminatory content, facilitating online harassment, disproportionately reporting accounts of individuals from marginalized (religious) communities based on user reports, etc. This risk poses a serious threat to the rights of EU citizens who are already vulnerable to abuse and have encountered discrimination and marginalization historically. Without mitigations, content that undermines the Right to Non-Discrimination and Freedom of Religion could conceivably appear



[REDACTED]

in any of Snapchat’s in-scope services displaying user generated content, including information in videos featured on Spotlight / Discover and Snap Map.

Likelihood

All of the risks we track on Snapchat have a low prevalence compared to the prevalence of these issues elsewhere online and offline. To aid our prioritization, our methodology seeks to assess the relative likelihood between the risks we track (even though all are low in absolute terms).

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED] We

have assessed the relative likelihood of Snapchat’s inscope services disseminating content relating to matters that may undermine the Right to Non-Discrimination and Freedom of Religion, based on Policy Violating Prevalence (PVP) via random sampling and our Transparency Report data in Section 4.1 of this Report, as follows:

Category	Relative likelihood of risk occurring on Snapchat
4.1.2 Dissemination of Illegal Hate Speech	Lowest Relative Likelihood
4.1.4 Dissemination of Terrorist Content	Lowest Relative Likelihood

In addition, although we placed the risks to the Right to Data Protection in the highest relative likelihood category in Section 4.2.4, we have noted that the overall number of privacy and data protection related queries we received is very low. When focused on algorithmic bias specifically, we have not received any material volume of queries, which is not surprising as we noted in Section 4.2.2 (Right to Freedom of Expression and information) that Snap’s inscope services do not generally provide a platform for political public content.

We therefore continue to assess that the relative likelihood that the in-scope services of Snapchat would have an actual or foreseeable Negative Effects of the Right to Human Dignity falls within **Extremely Low Likelihood category**.

Severity

This risk poses a serious threat to the rights of EU citizens, many of whom have historically faced discrimination and marginalization. Recent examples underscore how algorithmic systems can perpetuate bias even today. For instance, the Netherlands Institute for Human Rights found that the Breeze dating app's matching algorithm discriminated against non-white users, resulting in a



formal order to address the bias.¹²¹ In the regulatory sphere, the European Commission has demanded that X (formerly Twitter) provide internal documentation on how its recommendation systems may be amplifying far-right content—underscoring concerns about algorithmic moderation potentially skewing democratic discourse.¹²² These examples further reflect the scrutiny that algorithmic bias has attracted, within both legal frameworks and regulatory action, particularly under the evolving mandates of the EU’s AI Act and GDPR.

As in prior reports, we have assessed the severity of harm caused by categories of content that may undermines the Right to Non-Discrimination and Freedom of Religion in Section 4.1 of this Report, as follows:

Category	Relative likelihood of risk occurring on Snapchat
4.1.2 Dissemination of Illegal Hate Speech	Significant harm industry wide
4.1.4 Dissemination of Terrorist Content	Serious harm industry wide

Given these circumstances, Snap qualifies this risk of harm relating to negative impacts on the Right to Non-Discrimination and Freedom of Religion to be **‘serious’**.

DSA Risk Factors

In accordance with Article 34.2, our risk assessment also addresses our recommender systems, content moderation systems, applicable terms and conditions, systems for the selection and presenting of advertisements and any of our data-related practices. These factors have been listed and explained in Section 3 on the Risk Assessment Methodology and applied throughout Section 4. We also considered the risk factors in the context of the Right to Non-Discrimination and Freedom of Religion. As many of these factors pose similar risks, and are mitigated in a horizontal manner (i.e. by measures that cut across all systemic risks), we did not include them in this section again but refer to our explanation in Sections 4.1.2 on Hate Speech and Section 4.1.4 on Terrorist Content.

Overall potential risk prioritization

Snap would consider this risk to the right to non-discrimination and freedom of religion to fall be within the **Level 3** category overall. Although we consider the risk to fall within our serious harm category, there is arelatively low prevalence of Hate Speech, terrorism and bias concerns on the platform. There is no change in this assessment from our 2024 Report.

¹²¹ Tim de Jonge, Frederik Zuiderveen Borgesius, ‘Mitigating Digital Discrimination in Dating Apps -- The Dutch Breeze case’, [url](#), (2024)
¹²² Financial Times, ‘Brussels orders X to hand over documents on algorithm’,[url](#), (2025)



Snap's Mitigations

Practically, our in-app reporting tool allows users to directly report hateful content or activities that support terrorism or violent extremism. On our high-reach surfaces, like Spotlight and Discover, we take a proactive approach to moderating any content that may violate these rules. When hateful content is reported, our teams will remove any violating content and users who engage in repeated or egregious violations will have their account access locked. As an additional measure, we encourage Snapchatters to block any users who make them feel unsafe or uncomfortable.

- We survey a subset of our users quarterly to understand whether they find their time spent on our experience entertaining and satisfactory. We use this to track whether our product changes are improving viewers' overall perception of the app.
- We provide a diversity of perspectives. We have multiple programs to foster a more diverse content community and surface different perspectives (e.g. [Black accelerator program](#).)
- We ensure there is always a large mix of content from creators from viewers' home country and content in the language in which they have set their device.



- We add diversity to every viewer's feed in terms of the account they see, and the categories of content we surface to them. This prevents users from entering an echo chamber or filter bubble of seeing the same content repeatedly. We use machine learning to understand content categories and diversify it.

Modifying facial features or overlaying cultural elements in Snapchat's Lenses may reinforce discriminatory ideas based on appearance or ethnicity and promote harmful imagery. Also, Lenses incorporating cultural symbols or references might lack proper context and sensitivity. The [Lens Studio Submission Guidelines](#), reiterated our Community Guidelines and spelled out that the following categories of Lenses are prohibited:

- Content that demeans, defames, or promotes discrimination or violence on the basis of any of the identities listed in our Community Guidelines
- Examples: slurs, stereotypes, hate symbols, the promotion of hateful conspiracy theories, the glorification of atrocities or historical hatemongers

Snap designs every Lens with race, gender, ethnicity and cultural norms in mind. Snap leverages its ever-growing diverse training datasets, as well as feedback from community members. If a Lens does not resonate with our community, as expressed through a high ratio of user reports, we take that feedback into consideration and will re-review the Lens with a goal to leave as-is, modify, or remove.

If a Lens is appropriate, but could theoretically be misused by someone, that alone is not sufficient to reject a Lens. Snap considers current and historical global events when releasing a Lens, and delays or denies amplification to Lenses that may be deemed insensitive due to broader social occurrences throughout the world. We aim for our Lenses to celebrate individuality and diversity without altering a user's skin tone or features to mimic another ethnicity or race. We are committed to improving our technology to uphold these values. . Snap presents religious and cultural iconography in a respectful manner, with feedback solicited from internal and external subject matter experts. This means Snap is especially thoughtful around holiday or event-based content, including the geography in which a Lens will launch. Also, Snap ensures that a Lens is not deceptive. Snap uses signifiers and watermarks where there may be questions of creative authenticity. Snap tests Lenses on photos/videos of and in real life settings with diverse groups of people to accurately enforce our policies.

As reported in our 2024 [Diversity Annual Report](#), we know DEI is critical for long-term growth - whether it's Snapchatters demanding products to meet their diverse needs, or the desire to reach new and different markets. It also highlights in particular two new initiatives since our 2023 Report that showcase examples of how empathy can inspire new perspectives and tangible business impact.



- **Snap Out Loud** - Snapchat is a platform that celebrates authenticity. That's why our team created an AR experience to spotlight the different communities who share the LGBTQIA umbrella. Led by SnapPride, this Snap Show educates Snapchatters about the meaning of LGBTQIA, and celebrates the people who make up the community. Snapchatters were welcomed into seven separate spaces, denoted by the letters of the acronym, to explore each world. 25 million unique users were reached across 11 countries and the lens was shared one million times.



- **8 Mars 8 femmes** - In France, only 10% of statues in public spaces honor female figures. SnapWomen partnered with the Sales and AR Studio team in Paris to launch an AR activation on International Women's Day 2023 across 8 major cities. Called 8 Mars 8 femmes, or 8th of March, 8 women, the activation featured female AR statues next to male ones to celebrate great women in history who were never given appropriate credit for their impact. These AR statues are permanently activated, and honor Josephine Baker, Olympe de Gouges, Manon Tardon, Hubertine Auclert, Simone Veil, Françoise de Graffigny, Élisabeth Vigée Le Brun, and Simone de Beauvoir. The initiative received widespread media coverage in France. Snap's AR Studio team was honored to win the global Drum Award for Marketing for their outstanding creativity.



Specific Mitigations.

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive. As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **[link in the left hand column to a full summary](#)** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, Snapchat's in-scope services have been adapted to include proactive moderation for hateful content or activities supporting terrorism or violent extremism. We also work with civil society organizations to ensure our policies are enforced responsibly. Product Inclusion helps us create equitable experiences by intentionally involving and considering marginalized groups



	at critical moments throughout the product development process.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes, Snap's Terms and Community Guidelines prohibit Hate Speech or content that demeans, defames, or promotes discrimination or violence on the basis of race, color, caste, ethnicity, national origin, religion, sexual orientation, gender identity, disability, or veteran status, immigration status, socio-economic status, age, weight, or pregnancy status. We strictly enforce these rules.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, specific proactive and reactive moderation procedures to prevent and remove hateful content or activities supporting terrorism or violent extremism. We provide in-app reporting for hateful content or activities supporting terrorism or violent extremism.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, our algorithmic systems do not categorize or recommend hateful content or activities supporting terrorism or violent extremism.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems. In order to ensure we are not using discriminatory targeting models particularly when there is significant legal impact to the consumers, we offer special targeting models that do not include gender or age, which we require for advertisers who are advertising in the housing, credit or employment (HCE) spaces, so that discriminatory factors will not go into who sees these ads. We do not allow advertisers to build audiences for their ads based on their own data about our teenage users regardless of those user's own ad settings (i.e. activity data from the advertisers own online properties and the advertiser's own customer lists).
Ongoing Risk Detection and Management	Yes, we have specific prevalence testing and transparency reporting for Hate Speech, terrorist and violent extremist



Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	content.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Yes, we cooperate with trusted flaggers in relation to illegal hate speech, terrorist and violent extremist content.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various industry groups e.g. EU Internet Forum. We have signed onto the EU Hate Speech Code.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on our terms, harms, moderation and enforcement practices (see the Annex), as well as how to and how to get help in our Safety Center. We provide in-app reporting for hateful content or activities supporting terrorism or violent extremism.
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, we have protective measures in place for Teens. For example, Teens cannot create public profiles and if they post to Spotlight or Snap Maps their profile details are anonymized as an extra precaution. Our reporting menu also includes the option to report “It involves a child”. We hope protections like these help protect Teens from hateful content. Our Family Center includes resources and guidance for Teens and their parents or trusted adults. Our parents site provides additional guidance for parents and carers on risks and support. ¹²³
Content Authenticity	No specific content authenticity measures taken in respect of Hate Speech, terrorist and violent extremist content.

¹²³ <https://parents.snapchat.com>.



Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	However, Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services.
--	---

Conclusion

Snap considers the overall risk to be within the Level 3 category taking account of the harm that risks to the Right to Non-Discrimination and Freedom of Religion may cause and the low prevalence for Hate Speech on the platform. In practice Snap has substantial protective measures in place. Snap works with civil society organizations, like Access Now, human rights experts, law enforcement agencies, NGOs, and safety advocates to make sure we are calibrating wherever necessary to ensure that our products and policies function to keep Snapchatters safe. Our in-app reporting tool allows users to directly report hateful content or activities that support terrorism or violent extremism. On our high-reach surfaces, like Spotlight and Discover, we take a proactive approach to moderating any content that may violate these rules. Further, our diversity and inclusion efforts continue to help us create equitable experiences and build inclusive products.

We have concluded therefore that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures to protect users' Right to Non-Discrimination and Freedom of Religion. There is no change in this conclusion from our 2024 Report.

4.2.6 Children's Rights

We understand that online platforms can impact Children's Rights. This is a risk we take seriously as Snap's priority is protecting the safety and wellbeing of our users whilst ensuring they have a positive experience online. Privacy, safety and security are key values of the company and at the core of our value proposition to our users.

The 'rights of the child' under the Charter¹²⁴ comprises two elements that are relevant to Snapchat's in-scope services:

1. Children have the right to such protection and care as is necessary for their well-being; and

¹²⁴ Art 24, Charter of Fundamental Rights of the European Union (CFREU), [url](#).



2. Children have the right to express their views freely and have those views taken into consideration on matters which concern them in accordance with their age and maturity.

In respect of element 1, we address the well-being of children when considering Category 4 of the DSA risks in particular parts of the [Negative Effects on Minors](#) and [physical and mental wellbeing](#) elsewhere in this Section 4. **This section therefore focuses on element 2 i.e. risks to children's rights of expression.**

Likelihood

As explained in [Snapchat Community](#) as part of our Introduction to this Report, Snapchat is used by a wide demographic, with 18-24 year olds still making up the highest percentage of users of Snapchat. Nevertheless, there is still a percentage of our users who are Teens (13-17). Therefore we still consider that Teens using Snapchat are just as likely to be exposed to freedom of expression issues identified in this Report as other members of the Snapchat Community as follows:

Risk Category	Relative likelihood of risk occurring on Snapchat	Relative likelihood of Negative Effect on children
Right to Freedom of Expression	Extremely Low Likelihood	Extremely Low Likelihood

As a result, we continue to conclude that the relative likelihood of a risk of Negative Effects on children and Teens falls within the **Lowest Relative Likelihood** category.

Severity

We assessed the risk of harm from the Right to Freedom of Expression to fall within our significant harm classification. However, we take the safety and wellbeing of the youngest members of our community very seriously and recognise that this group is particularly vulnerable and if a particular risk materializes, there is an increased risk that the severity of the harm they suffer is higher. For Freedom of Expression, we consider this as follows:

Risk Category	Harm classification industry wide	Is the industry wide severity risk higher for children and Teens?
Right to Freedom of Expression	Significant harm industry wide	Yes, Snap considers that it is vital that children and Teens are able to access online platforms and participate in lawful online debate and dialogue to learn, have their views heard and develop their own values and identities, regardless of their



		ability to pay.
--	--	-----------------

As a result, we have chosen to place the severity of harm arising from an issue that negatively affects Children’s Rights in our ‘**severe**’ category.

DSA Risk Factors

In accordance with Article 34(2), our risk assessment also addresses our recommender systems, content moderation systems, applicable terms and conditions, systems for the selection and presenting of advertisements and any of our data-related practices. These factors have been listed and explained in Section 3 on the Risk Assessment Methodology and applied throughout Section 4. We also considered the risk factors in the context of Children’s Rights. As many of these factors pose similar risks, and are mitigated in a horizontal manner (i.e. by measures that cut across all systemic risks), we did not include them in this section again but refer to our explanation in Section 4.4.3 on Negative Effects on Minors.

Overall potential risk

Although the relative likelihood for the Negative Effects on Children’s Rights falls within our Extremely Low Likelihood category, Snap considers the risk of harm to fall within the severest category. Consequently, Snap considers this to be a **Level 1** overall potential risk for Snapchat’s in-scope services. There is no change in this assessment from our 2024 Report.

As described in our [Risk Methodology section](#), we assess overall potential risk on a case by case basis and Snap reserves the option to deviate from the overall potential prioritization risk matrix we use as a guide. This is one of the cases where we have chosen to deviate.

[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
	[Redacted]	[Redacted]	[Redacted]	[Redacted]
	[Redacted]	[Redacted]	[Redacted]	[Redacted]
		[Redacted]	[Redacted]	[Redacted]
[Redacted]				

Snap's Mitigations

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat’s in-scope services, using the defined list of mitigations set out in Article 35 of the



DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive.

As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **link in the left hand column to a full summary** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, several aspects of Snapchat's design and function reduce the risk for the Freedom of Expression and Assembly: • Snap makes it difficult for unvetted content to reach a large audience on Snapchat, and Snap proactively moderates Snapchat's in-scope services that provide an opportunity to reach a larger audience. • Our platform is generally not a place for political or activist public content. Such content is not eligible for promotion on Spotlight and user content on Discover is only from a small number of popular, entertaining community creators and their content is moderated by humans against our Content Guidelines. • All Media Partners are vetted prior to being permitted to distribute their content broadly on Snapchat by a team of editors. These Partners include news organizations, which are subject to their own professional rules. Media Partners go through an editorial review of their content, a reputational search (to evaluate if a publisher has a history damaging press, legal actions, etc.), and compliance review before they're able to distribute content. As a result, we provide a balanced approach to political and activist public content on Snapchat that is designed to limit the sources of such information to professional media partners. • As explained when discussing the dissemination of content that infringes on intellectual property rights, Snap respects the doctrine of "fair use," i.e., that there are certain circumstances (such as news reporting, social commentary on issues of public interest, criticism, parody, or education) where excerpts of copyrighted material could be distributed without permission from or payment to the copyright holder.



Mitigation Category	Applies to this risk?
	This helps reinforce the rights of Freedom of Expression and the Freedom of Assembly.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Our Terms clearly define certain topics which we prohibit, including false information that threatens Public Health (e.g. COVID-19 vaccinations), civic processes, or that denies tragic events (like the Holocaust). Yes, Snap's Terms and Community Guidelines expressly prohibit harmful false or deceptive information and they are strictly enforced. We respond with swift and strict consequences against violators as explained in our Explainer. We promptly enforce against accounts found to be sending Terrorist Content: • Snap removes such content for all users. • Accounts we discover engaging in prohibited terrorist activity will also be promptly disabled. • Where appropriate, accounts engaging in violation of these policies may be reported to law enforcement.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, we have specific proactive and reactive moderation procedures to prevent and remove violative content. As explained in the Moderation section in Section 5, Snap deploys a range of automated content moderation (which include abusive language detection, other keyword-based detection, and machine-learning-based proactive detection models) to scan Stories and Spotlight submissions. Snapchatters can report violative content to us via in-app reporting options and anyone can submit a report through the Snapchat Support Site. Our in-app reporting tool also allows users to directly report violative content.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, unlike many of our peers, Snapchat does not offer an open news feed where unvetted publishers or individuals have an opportunity to broadcast violative content, does not offer a broad 'reshare' functionality that would encourage virality, and does not allow user-generated content to be recommended to a wide audience without going through review. Our algorithmic systems do not knowingly recommend violative content. As explained in the Moderation section of our Existing



Mitigation Category	Applies to this risk?
	Mitigations in Section 5, on our high-reach surfaces, like Spotlight and Discover, we take a proactive approach to moderating any content that may violate these rules prior to the content being recommended to a wide audience. The pool of content recommended by our algorithmic systems does not generally include political or other important societal matters regardless of where they fall on the political spectrum.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, for example we monitor data to help detect and manage content that may present a risk for the Right to Freedom of Expression, including data from our specific prevalence testing and enforcements (which are summarised in our Transparency Reports).
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	No, we do not work with trusted flaggers for users' rights to Freedom of Expression and access to accurate information. However we are working with trusted flaggers on children's safety in general.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, our Crisis Protocols balance Teen's rights to Freedom of Expression with access to accurate information. We have recently exercised these protocols successfully during the French riots in June 2023. Note, we are actively working to support efforts to agree an EU Age appropriate design code to protect Children's Rights.
Transparency	Yes, we provide guidance on harms and how to get help in our



Mitigation Category	Applies to this risk?
Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Safety Center. We make available robust reporting tools and we provide guidance to parents on the web (see below).
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	We have protective measures to allow Teens to express themselves without the pressures of friends lists, comments and likes. We have community, ad and content guidelines that are specific to Teens. Snap also adopted protective measures to limit Teen contact with strangers; we offer Family Center , reporting, and guidance. Our new parents site provides additional guidance for parents and caregivers on risks and support. ¹²⁵
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	Yes, Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's inscope services.

Conclusion

Snap considers Negative Effects on Children's Rights to be a lower likelihood risk but one that has a risk for severe harm industry wide, without appropriate mitigations. As a result we treat this as one of our highest priority risks, with a Level 1 Risk Prioritization. Snap is designed to fairly apply rules on content publication and provide an appropriate environment for Teens to exercise expression and assembly on Snapchat's in-scope services (and Snapchat as a whole). As explained in the [Freedom of Expression](#) and [Protection of Minors](#) section of the Report, this includes adapting our systems to limit the access of Teen accounts to higher risk features and content, like public profiles and sexually suggestive content, as well providing Teens and Families with accessible guidance and tools for the use of Snapchat and ensuring our [Terms](#), [Moderation](#) and [Enforcement](#) also operate fairly.

¹²⁵ <https://parents.snapchat.com>.



[REDACTED]

We have concluded therefore that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures to protect against Negative Effects on Children's Rights. In addition, we continue to assess Snapchat against the Commission's new Article 28 Guidelines on ensuring a high level of privacy, safety and security for minors.

4.2.7 Right to Consumer Protection

We understand that without mitigations online platforms can be used to spread content that contains false or misleading information that can harm consumers. This risk to the Right to Consumer Protection poses a serious threat to the rights of EU citizens who may be vulnerable to deception or invasion of privacy. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Likelihood

Snap has implemented safeguards, both through product design and policy enforcement, to effectively diminish the likelihood that the Right to Consumer Protection is violated on the platform. We monitor the number and nature of Privacy and Data Protection requests we receive.

In our 2024 Report, we explained that:

- Our prevalence testing as at 30 July 2024 showed that "invasion of privacy" had seen a further, substantial fall [REDACTED] It is now at an Extremely Low PVP level.
- In July 2023-June 2024, [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED] Overall we have seen a slight increase in line with expectations and we continue to receive low numbers of privacy-related queries.

In this reporting period, we observed the following:

- As of 30 April 2025, our prevalence testing indicated a slight increase in reports of "invasion of privacy," though the overall level remains low at [REDACTED]



- In July 2024-June 2025 [REDACTED]
[REDACTED]
[REDACTED] This reflects a slight decrease in privacy operations requests compared to our previous reporting period, alongside a considerable increase in GDPR- or Member State-related queries submitted through the DPO channel. Despite this increase, the overall volume remains relatively low, [REDACTED]
[REDACTED]

We monitor the number of community support requests we receive relating to the European Union. Note, these figures (and the privacy figures above) concern the requests which we review manually and excludes automated responses. We have observed a steady decrease in EU consumer support requests:

- [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

Finally, we also monitor for significant changes in our ad review processes on a global basis:

- [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

- [REDACTED]
[REDACTED]
[REDACTED]



Given these safeguards we have in place, and the low, decreasing level of consumer queries and consistent levels of ad rejections for our main fraud category, Snap considers that the risk to the right of consumer protection falls within the **Extremely Low Likelihood level**.

Severity

As described above, Snap recognises that the processing and sharing of extensive user data can lead to disproportionate personalized ad targeting.¹²⁶ Additionally, misleading or false advertisements can harm consumers, which harm is recognized generally by the European Union, for example within its Audiovisual Media Services Directive which also covers ad transparency, and in relation to political ads.¹²⁷ National legislators have also recognized the need for transparency, see for example the Dutch ACM's recent guidelines for promoting a transparent and fair online platform economy for businesses.¹²⁸

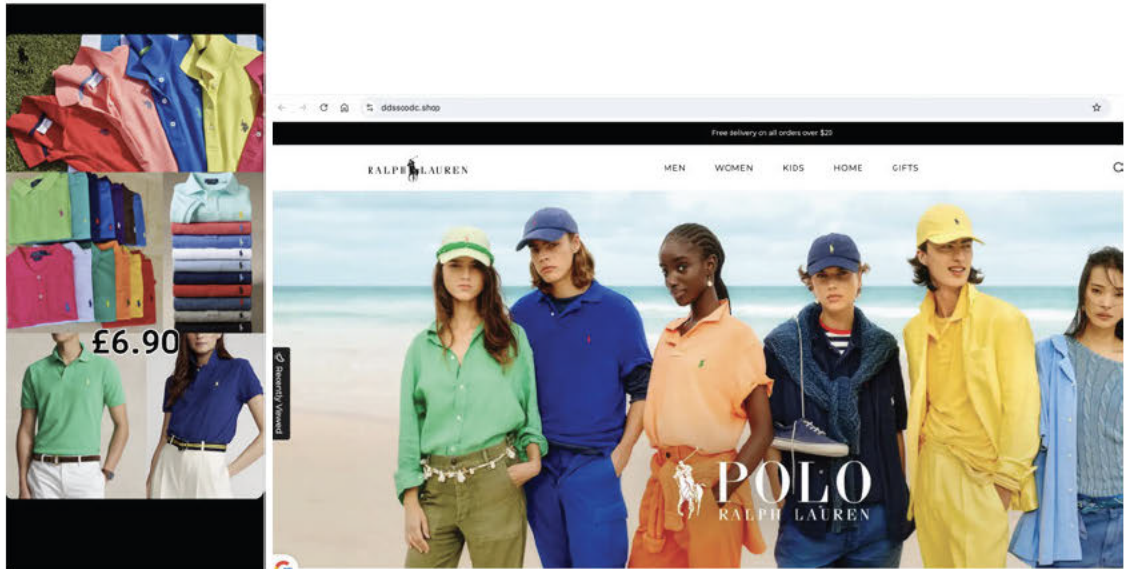
An example of a fraudulent ad that would be rejected by Snap is shown below. The advert itself uses a low quality top Snap with too good to be true prices (\$6.90 for Polo brand shirts). The advert links to a URL is suspicious, and ultimately redirects to a website with what appears to be a fake Ralph Lauren landing page.¹²⁹ Products listed on this page are again below \$10. **It is unlikely any customer placing orders will receive any goods, if any goods are shipped it is extremely unlikely they will be authentic. It is also very likely any credit card used on this page would be stolen and would result in a chargeback.**

¹²⁶ A.M. Correa, 'Regulating targeted advertising: Addressing discrimination with transparency, fairness, and auditing tests remedies', 2022, 46 Computer Law & Security Review, [url](#).

¹²⁷ See for example Regulation (EU) 2024/900 of the European Parliament and of the Council of 13 March 2024 on the transparency and targeting of political advertising, [url](#).

¹²⁸ Authority for Consumers & Markets, 'ACM Guidelines for Promoting a transparent and fair online platform economy for businesses', April 2023, [url](#). See also: ACM, AFM, AP, CvdM, 'Basic principles for effective online transparency' July 2023, [url](#).

¹²⁹ <https://mart.hdebrnz.shop/ralphlauren> which redirects to <https://www.ddssoodc.shop/>.



As a result Snap treats risk of harm from a negative impact on the Right to Consumer Protection as **significant** due to the potential harm it can cause to users.

DSA Risk Factors

In accordance with Article 34(2), our risk assessment also addresses our recommender systems, content moderation systems, applicable terms and conditions, systems for the selection and presenting of advertisements and any of our data-related practices. These factors have been listed and explained in [Section 3 on the Risk Assessment Methodology](#) and applied throughout Section 4. We also considered the risk factors in the context of the Right to Consumer Protection. As many of these factors pose similar risks, and are mitigated in a horizontal manner (i.e. by measures that cut across all systemic risks), we did not include them in this section again but refer to our explanation in Section 4.1.3 on the Sale of Prohibited Goods and Services and Section 4.2.3 on the Right to Private Life.

Overall potential risk prioritization

Given we have assessed the potential for negative impacts to the Right to Consumer Protection to fall within the lowest likelihood category and to have a risk of significant harm, we consider this risk to fall within our **Level 3 potential risk prioritization category**, given the Extremely Low Likelihood and significant harm categorization.. There is no change in this assessment from our 2023 Report.



Snap's Mitigations

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive.

As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **link in the left hand column to a full summary** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

DSA Mitigation	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, Snapchat's in-scope services have been adapted to include proactive moderation for ads that violate our policies, or false or deceptive content. Snap places a strong emphasis on its adherence to Article 25 DSA concerning dark patterns. Consequently, this constitutes a strategic mitigation measure aimed at mitigating the potential impediment to the Right to Consumer Protection. Snap is committed to ongoing monitoring of this aspect to ensure continued compliance and effectiveness. We also require a specific minimum audience of [REDACTED] Snapchatters to prevent advertisers from manipulating small audiences.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes, Snap has invested considerable resources in developing and enforcing advertising policies that safeguard the Right to Consumer Protection. We have robust ad policies to prevent inappropriate and illegal advertising on our platform. Our Community Guidelines also prohibit spreading false



	information that causes harm or is malicious, impersonation, i.e., attempting to deceive people about who you are, and disallow spam and other deceptive practices. Our Commercial Content Policy also disallows false or misleading content, including deceptive claims, offers, functionality, or business practices, promotion of fraudulent goods or services, products or services with false celebrity testimonials or usage, deceptive financial products, and other similar content.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, we use a combination of automated and human review to prevent and remove ads that violate our policies or the law from appearing on Snapchat. This also includes ensuring inappropriate ads are not targeted at Teens. Additionally, all ads can be flagged by Snapchatters in the app as being inappropriate along with the reason for the violation. We also don't allow user-generated political content from being promoted on Spotlight. We take these measures in order to circumvent the spread of harmful and false content.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	See Advertising Systems below.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	To address potential risks with targeted advertisements, and to ensure advertisers are not manipulating small audiences with micro-targeted campaigns, most of the ads on Snapchat, including all political ads, require a specific minimum of [REDACTED] Snapchatters to be targeted. We also offer special targeting models that do not include gender or age for advertisers who are advertising in the housing, credit or employment (HCE) spaces so that discriminatory factors will not go into who sees these ads. Lastly, to ensure that users have choice about use of their personal data for targeting ads, we allow users to control the data that's used to determine the ads they see. In the EU, we offer controls to turn off most personalized ads and for other regions users can restrict our use of third party data and being included in advertiser supplied audience matches for ads targeting.
Ongoing Risk Detection and Management	Yes, we have specific prevalence testing and Transparency Report false information, impersonation, spam and other



Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	regulated goods.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Yes, we cooperate with trusted flaggers in relation to regulated goods.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various industry groups e.g. CIPL, FPF.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, to ensure users know when content is commercial in nature, we automatically place an “Ad” marker on all paid ads that run on Snapchat. Our Commercial Content Policy requires all organic content posted by influencers to be marked appropriately and we now offer a “Paid Partnership” tag tool that influencers and users may use when they post commercial content to help them comply with this policy and their legal obligations. We also provide transparency on our privacy practices including ads on our Privacy Center and provide an ads library in line with DSA requirements.
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, for example we prevent inappropriate ads for Teens and advertising based on profiling. We make available robust reporting; and we offer Family Center and provide guidance to parents on the web. Our parents site provides additional guidance for parents and carers on risks and support.¹³⁰
Content Authenticity	Yes, Snap has taken steps to mitigate the risk that (i) its

¹³⁰ <https://parents.snapchat.com>.



[REDACTED]

Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat’s in-scope services.
--	---

Conclusion

Snap considers Right to Consumer Protection risks to fall within our Level 3 risk prioritization. Although there is a risk of significant harm arising from content and activity negatively impacting the Right to Consumer Protection, our evidence indicates a relatively low likelihood on Snapchat. In response, Snap takes a multipronged approach and has put in place a range of mitigation measures. These include, for example, developing and enforcing advertising policies that safeguard the Right to Consumer Protection. Our ad policies aim to prevent inappropriate and illegal advertising and our review processes were designed to enforce these policies. Through these terms and other mitigations, such as safeguards in the product design and policy enforcement, Snap has been able to effectively uphold users’ consumer protection rights.

We have concluded therefore that Snapchat’s in-scope services have reasonable, proportionate, and effective mitigation measures for the risk of Negative Effects on the Right to Consumer Protection. There is no change in this conclusion from our 2024 Report.

4.2.8 Right to Property

The property right that has a significant risk of being impacted by Snapchat’s in-scope services is the right to intellectual property. This risk stems from the disclosure of such property contrary to the intellectual property rights of a natural or legal person. This is discussed above under [Section 4.1.5](#) (IP Infringement).

In addition, we continue to consider there is a potential risk that individuals may harm someone else’s property while under pressure to create content that others find entertaining or humorous on Snapchat. [REDACTED]

[REDACTED]

[REDACTED]



██████████ This risk is discussed above under [Section 4.1.9](#) (Dissemination of content encouraging or engaging in violent or dangerous behavior).

4.3 Category 3: Negative effects on Public Security

(Article 34.1.c / DSA Recital 82)

In this part of the Report, we explain the results of our assessment on actual or foreseeable negative effects of Snapchat's in-scope services on Democratic and Electoral Processes, Civic Discourse and Public Security as required by Article 34.1.c and Recital 82 of the Digital Services Act. We have assessed in particular Negative Effects on Democratic and Electoral Processes, Civic Discourse, as well as Public Security.

Note that for all harms, where there is (1) a risk of significant damage to the physical or emotional well-being of Snapchatters, and (2) imminent, credible risk of severe harm, including threats to human life, safety, and well-being, we treat this as a severe harm and an Level 1 overall risk prioritization.

Category 3 - Negative effect on Public Security				
Category	Relative likelihood of risk occurring on Snapchat	Harm classification industry wide	Risk Prioritization	Conclusion
4.3.1 Negative Effects on Democratic and Electoral Processes	Extremely Low Likelihood	Severe harm industry wide	Level 3	Low Risk / Reasonable, proportionate and effective mitigations
4.3.2 Negative Effects on Civic Discourse	Extremely Low Likelihood	Severe harm industry wide	Level 3	Low Risk / Reasonable, proportionate and effective mitigations
4.3.3 Negative Effects on Public Security	Extremely Low Likelihood	Serious harm industry wide	Level 3	Low Risk / Reasonable, proportionate and effective mitigations

4.3.1 Negative Effects on Democratic and Electoral Processes

The role of digital platforms in helping to shape information environments establishes a significant nexus with Democratic and Electoral Processes. As digital technologies such as Snap enable expression and access to information, the impact of these platforms on the free and fair



[REDACTED]

exercise of political rights warrants careful attention, presenting risks to which Snap has long been vigilant.

[REDACTED]

Likelihood

As outlined in Section 1, a significant proportion of EU citizens use Snapchat. As at 1 January 2025, we have 93.7 million average monthly active recipients of Snapchat in the EU, and significant recipient numbers in individual Member States. There is the potential for public content on Snapchat to reach a sizable audience within the European Union (particularly within the 18-24 age bracket which accounts for the biggest share of our registered accounts).

However, Snapchat's platform architecture, combined with its commitment to responsible policy enforcement across our content surfaces, establishes unique safeguards against risks to democracy. The steps Snap has taken to mitigate threats to democracy mean that likelihood is substantially diminished.

Independent reports of electoral interference on Snapchat are vanishingly rare. In connection with a major, high-profile election in 2022, we onboarded Snap to the Election Integrity Partnership (EIP),¹³⁵ a partnership among leading research centers and civil society organizations who monitor online harms to democratic processes; as participants in the EIP threat escalation program, our teams received only one single incident report from the researchers monitoring risks on Snapchat. We participated in the Commission's stress test and multi-stakeholder roundtable dialogues ahead of the European Parliament elections in 2024 and were able to successfully navigate the test exercises. As we reported in our [election blog post](#) on 24 June 2024, Snap saw a small uptick in reported activity, but did not receive or observe any material incidents or threats. Our moderation and reporting tools worked well, and none of the reported pieces of content were verified as misinformation on Snapchat.

[REDACTED]

¹³⁵ Election Integrity Partnership (2020), [url](#).



As highlighted in [Section 4.1.10](#) (Harmful False Information), Snap’s own reporting metrics confirm the limited occurrence of content harmful to democracy:

- Our [Prevalence Testing](#) has consistently found extremely low levels of Harmful False Information. Testing conducted since July 2024 shows a further significant decline, [REDACTED]
- Our most relevant [transparency reporting](#) category on this topic is “Harmful False Information,” which our policies define as including content that “undermines the integrity of civic processes.” In the second half of 2024, false information continues to account for only 0.013% of the total of all content enforced in the EU on Snapchat, showing a decrease from 0.015% registered in the first half of the same year and even a further decrease from 0.020% in the second half of 2023.

Snap’s product design and policy practices continue to substantially reduce the likelihood of negative impacts on democracy. Our ongoing quantitative and qualitative analysis shows that the risk of potential negative impact on Democratic and Electoral Processes on Snapchat falls into our **Extremely Low likelihood category**.

Severity

Snap takes risks to democracy very seriously. Although the recent European Parliament elections unfolded overall in a positive online environment with no major threats¹³⁶, other real-world examples – like Cambridge Analytica,¹³⁷ the Brazilian elections,¹³⁸ and the US Presidential elections¹³⁹ – illustrate the devastating negative impact that social media and content platform can have on Democratic and Electoral Processes. Quite often misinformation is designed to undermine trust in the electoral process and dissuade specific groups of people from exercising their right to vote.¹⁴⁰

Accounting for the real-world examples illustrating the potential disruptive effects that digital platforms can have on Democratic and Electoral Processes, we understand that Negative Effects on Democratic and Electoral Processes have a risk of **severe harm** if not properly mitigated.

DSA Risk Factors

In accordance with [Section 3 \(Methodology\)](#), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat’s in-scope services. As set out below, while the

¹³⁶ European Commission and independent observers confirmed that they did not observe major online threats in recent multi-stakeholder dialogues relating to the European elections, [url](#)

¹³⁷ The New York Times, ‘Cambridge Analytica and Facebook: The Scandal and the Fallout So Far’, April 2018, [url](#).

¹³⁸ MisinfoReview, ‘Explaining beliefs in electoral misinformation in the 2022 Brazilian election: The role of ideology, political trust, social media, and messaging apps’, May 2023, [url](#).

¹³⁹ <https://www.hrw.org/news/2024/08/15/disinformation-about-us-elections-targets-communities-color>.

¹⁴⁰ See for example: <https://edmo.eu/publications/final-report-results-and-outcomes-of-a-community-wide-effort/> and <https://www.brennancenter.org/election-misinformation>.



following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap's existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Our recommender systems are designed not to knowingly recommend political content. Political content is not eligible for promotion in Spotlight and all Spotlight and non-professional Discover content must pass both automated and human review before wider recommendation. Detailed information regarding the mitigations relating to our recommender systems can be found in Section 5 (see the Algorithmic Systems sub-section).
(b) content moderation systems;	Snap has implemented specific proactive and reactive moderation procedures to prevent and remove electoral misinformation. Moderators are given specific escalation guidance during elections. We provide more information in Section 5 (see Content Moderation sub-section)
(c) the applicable terms and conditions and enforcement;	Our Community Guidelines (which form part of our terms of service) prohibit content that undermines the integrity of elections and civic processes. Enforcement covers all content formats, including AI-generated media. We provide more information in Section 5 (see the Terms and Enforcement sub-sections).
(d) systems for selecting and presenting advertisements; and	Our advertising systems require agreement to advertising policies and guidance that prohibit adverts from displaying information that violates the law or causes certain harms. Specifically, our political ad policies require that any political advertisements are subject to review and fact-checking before they are eligible for placement on Snapchat. We check advertisers are complying with their obligations via our advertising review process. Our advertising systems use a mix of automation and human review to review adverts before they are published. We provide easy mechanisms for users to hide and report advertisements that violate our policies or the law. We monitor ad rejection, reporting and enforcement data to monitor the effectiveness of our approach. We provide more information in Section 5 (see the Advertising Systems sub-section)
(e) our data related practices	We have strong data principles, practices and privacy, safety



	and security by design processes. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).
--	---

We have also analysed whether and how the risk of Negative Effects to Democratic and Electoral Processes is influenced by the following general factors:

General Risk Factor	How does it apply to Snapchat and this harm?
Intentional manipulation, including inauthentic use or inauthentic use or automated exploitation of the service	There are two key ways in which we consider our systems could be manipulated: (1) Users could seek to share novel illegal and violating material that is not detected by our automated systems. We are constantly working to adjust our systems and policies to address this. We provide more information on our approach in Section 5 (in particular the Content Moderation subsection) and Section 6 (Ongoing Risk Detection and Management). (2) Users could abuse our content moderation processes and report non-violating content / accounts in bad faith. We have processes to combat misuse and more information can be found in Section 5 (Content Moderation and Enforcement subsections).
Amplification and potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.	Snapchat's in-scope services have a number of features and design configurations that act to limit the amplification and potentially rapid and wide dissemination of content that is harmful to Democratic and Electoral Processes, in particular: 1. Snapchat is not an attractive platform for spreading political misinformation, in particular because it is difficult to reach a broad audience and content is deleted by default. We provide more information in Section 5 (see Snapchat Design / Function subsection) 2. Snap has implemented specific proactive and reactive moderation procedures to prevent and remove content that is harmful to electoral processes. We provide more information in Section 5 (see Content Moderation). 3. Content recommended to users on Spotlight and Discover, our video sharing platforms, is moderated using a combination of auto-moderation and human moderation, and is human moderated before being



Taking into account the real-world examples illustrating the potential disruptive effects on democracy, there is risk of severe harm if not mitigated, as illustrated by Cambridge Analytica. However, we are encouraged that our prevalence data and reporting data show a very low likelihood of harmful false misinformation on Snapchat. As a result, we assess this risk to fall within our **Level 3 overall potential risk** category. There is no change in this assessment from our 2024 Report

Taking into account the real-world examples illustrating the potential disruptive effects on democracy, there is risk of severe harm if not mitigated, as illustrated by Cambridge Analytica. However, we are encouraged that our prevalence data and reporting data show a very low likelihood of harmful false misinformation on Snapchat. As a result, we assess this risk to fall within our **Level 3 overall potential risk** category. There is no change in this assessment from our 2024 Report

Snap's Mitigations

Mitigations relating to the Guideline for providers of VLOPs and VLOSEs on the mitigation of systemic risks for electoral processes

Snapchat's architecture and its commitment to responsible policy enforcement across our content surfaces, establishes unique safeguards against risks to democracy. We understand well that online platforms may have a negative effect on the electoral processes and the exercise of



political rights by amplifying digital disinformation or deceptive content relating to political matters or processes. However, the steps Snap has taken to mitigate threats to democracy mean that likelihood is substantially diminished.

As we highlighted in our previous Reports, Snap has for some time taken a multifaceted approach to mitigating negative impacts to democracy, including policy enforcement, product design, and expert engagement. This approach aligns with Guideline for providers of VLOPs and VLOSEs on the mitigation of systemic risks for electoral processes,¹⁴¹ as follows:

Internal Processes

As we have explained in this Report, and highlighted above, our assessment of risk prioritization with regards to Negative Effects on Democratic and Electoral Processes and Snap's measures to mitigate the risk, are guided by information on elements such as the presence and activity of political actors on the service and the number of Snapchatters in the EU and evidence regarding the use of tactics, techniques and procedures for information manipulation.

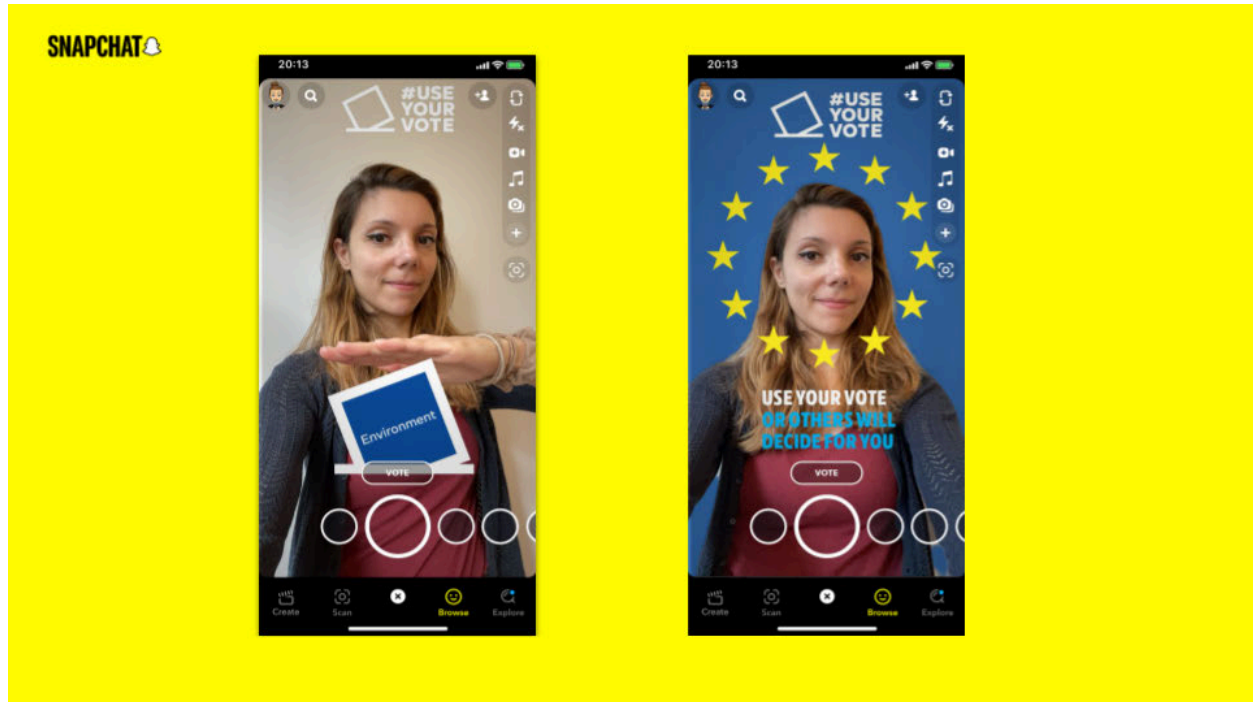
We also engage with experts in the information integrity and democracy and human rights community to inform our approach. This includes collaborations and engagement with individual experts (for example, former US Ambassador to the UN Human Rights Council Eileen Donahoe, in addition to several others), as well as think tanks (such as the Atlantic Council's Digital Forensics Research Lab) and research collaborations (such as the Election Integrity Partnership).

Access to official information on the electoral process and Media literacy initiatives

We have regularly partnered with governments around the world to inform Snapchatters about elections and invite them to go vote. We believe that civic engagement is one of the most powerful forms of self-expression and have previously worked with [election authorities in France, Netherlands](#), and Sweden to raise awareness of elections and encourage participation. A recent example was the 2023 Dutch provincial election cycle. With the Dutch Ministry of the Interior, Snap developed a lens where Snapchatters could place voting bins in their living room and answer questions about the election with 'true' or 'false'. By taking this quiz Snapchatters are increasing their knowledge about the elections and are reminded to go vote.

The recent European Parliament elections saw even more first-time voters eligible to participate – following the decision by Belgium and Germany to join Austria, Malta and Greece in lowering the voting age to 16. Ahead of these elections, we teamed up with the European Parliament on a special AR elections Lens that encourages people to get out and vote. During the election, we shared this Lens with all EU Snapchatters along with a message to remind them to vote and a link to the Parliament's election website.

¹⁴¹ Guidelines for providers of VLOPs and VLOSEs on the mitigation of systemic risks for electoral processes, April 2024, [url](#).



Snapchat partnered with the European Parliament and European Commission to promote their ‘[Use your vote](#)’ information campaign on elections, including a dedicated [Lens](#), and their [awareness campaign on the risks of disinformation](#) and deceptive content. Snap sent a push notification to over 50 million users to urge them to vote in the elections. Although out of scope of this Report, we note that we also further instructed My AI to avoid engaging on political topics. We are proud to have done our part to contribute to the highest observed turnout of the last 30 years, with 51.08% of the 357 million eligible citizens participating in the election.

Measures to provide users with more contextual information

One key way in which we mitigate the risk is through intentional product design choices. Our platform does not, for example, provide an unvetted feed of algorithmically curated political information; we disallow *all* political content from Spotlight (our broadcast platform for User Generated Content) and pre-moderate that surface to ensure that such political content is not distributed.¹⁴² This safeguard ensures that Snap is not algorithmically promoting political statements from unvetted sources, and generally reflects Spotlight’s function as an entertainment platform. (Consistent with our commitments to fundamental rights of expression and access to information, Snapchat provides other, non-algorithmically amplified spaces for users to express their views and political observations, such as Chat and My Story; users can also seek access to political information from known publishers and creators whom Snap has on-boarded for distribution on the Stories tab).

¹⁴² For these purposes, “political content” means content related to political campaigns and elections, government activities, and/or viewpoints on issues of ongoing debate or controversy. This includes content about candidates or parties for public office, ballot measures or referendums, and political action committees, as well as personal perspectives on candidate positions, government agencies/departments or the government as a whole.



Snap's policies expressly prohibit content that undermines the integrity of elections and civic processes. Drawing from expert research from the Election Integrity Partnership,¹⁴³ we orient this policy around four pillars of risk:

- *Procedural interference*: misinformation related to actual election or civic procedures, such as misrepresenting important dates and times or eligibility requirements for participation.
- *Participation interference*: content that includes intimidation to personal safety or spreads rumors to deter participation in the electoral or civic process.
- *Fraudulent or unlawful participation*: content that encourages people to misrepresent themselves to participate in the civic process or to illegally cast or destroy ballots.
- *Delegitimization of civic processes*: content aiming to delegitimize democratic institutions on the basis of false or misleading claims about election results, for example.

We take steps to explain our policy approach to safeguarding democratic information environments through our [Community Guidelines](#) and periodic [blog posts](#).

As technologies have evolved, we have updated our policies to cover all content formats – whether created by a human or generated by artificial intelligence (including deep fakes). In preparation for the recent EU elections, we also:

- Signed up to the [AI Elections Accord](#), alongside other technology firms, where we pledged to work collaboratively on tools to detect and limit the spread of AI generated content which aims to deceive voters.
- Introduced contextual symbols to help our community understand when they are interacting with Snap generated AI content.

Snap does not allow Lenses that encourage a particular political perspective. In line with this approach, politically related Lenses are disabled in Discover. Snap also rejects Lenses that perpetuate false information to elections (e.g. the wrong date). AR moderators are given strict guidance during elections to escalate misinformation.

As a result, when it comes to the in-scope content services of Snapchat, rather than taking measures to provide users with more context around disinformation and Foreign Information Manipulation and Interference (FIMI) content through labels and other indications, Snap's approach is to take steps to avoid recommending such content to a public audience in the first place (see below) and to remove such user generated content promptly when it is detected or reported. To the extent political content is distributed on in-scope services (i.e. political ads), Snap has safeguards in place, which are detailed in the political advertising section below.

¹⁴³ Election Integrity Partnership, 'Evaluating Platform Election-Related Speech Policies, October 2020, [url](#).



Recommender Systems

Content that is approved for broader audiences must comply with both our [Community Guidelines](#) and our [Content Guidelines for Recommendation Eligibility](#). All Spotlight and non-professional user generated Discover content goes through both automoderation, and often human review, for compliance with these guidelines before it is eligible for recommendation to a wide audience. As an additional safeguard, we monitor content that is achieving large-scale reach (and ensure a human reviews it) as a sort of “virality circuit breaker” and a means of checking that our pre-moderation systems are working effectively. Any content that is reported will be reviewed against the guidelines again for compliance.

Political Advertising

Political content is only eligible for broadcast (aka algorithmic distribution) on Snapchat on surfaces reserved for publishers or creators with whom Snap engages in partnership, or through advertising. Our [political ad policies](#) ensure that any political advertisements are subject to review and fact-checking *before* they are eligible for placement on Snapchat. We also prevent advertisers from manipulating small audiences with micro targeted campaigns, particularly for political ads. We do so by requiring a specific minimum audience [REDACTED] to be targeted (including Dynamic Ads on Snapchat | Snapchat for Business [REDACTED]).

In 2021 Snap joined the Dutch Code of Conduct for political ads.¹⁴⁴ Under this Code online platforms agreed to acknowledge a responsibility in maintaining the integrity of elections and avoid dissemination of misleading content and messages inciting violence or Hate Speech on their platforms, committed to making key data on online political advertising available publicly and help avoiding foreign interference in elections by banning political advertisements from outside the European Union, and putting in place a user-friendly response mechanism to answer questions or solve problems related to the Dutch elections.

In preparation for the European Parliament elections we also partnered with Logically Facts, a leading fact checking organisation and signatory of the [EU Disinformation Code of Practice](#), to help fact check political ad statements across the EU.

[REDACTED]
[REDACTED]
[REDACTED] we do not require ads to label when advertisement includes generative AI content nor do require advertisers to disclose to us the tools they used to edit or create their ad creative. Instead, our approach is to subject all of our ads to a review process, and political ads are also subject to fact checking. Deceptive ads are rejected, irrespective of whether they use AI, photoshop, or other digital editing tools. Ads that are not deceptive, and otherwise comply with

¹⁴⁴ For more details [url](#).



our Ad Policies, are approved to run (and if they are a political ad, they must include a “paid for by” disclaimer and are catalogued in Snap’s political ads library).

Influencers

Our [commercial content policy](#) requires all organic content posted by influencers to be marked appropriately. Commercial content that relates to the following is not permitted:

- Election-related content about candidates or parties for public office, ballot measures or referendums, political action committees, and content that urges people to vote or register to vote.
- Advocacy or issue content concerning issues or organisations that are the subject of debate on a local, national or global level, or of public importance. Examples include: content about abortion, immigration, the environment, education, discrimination and guns.

We now offer a “Paid Partnership” tag tool that influencers and users may use when they post commercial content to help them comply with this policy and their legal obligations. We make clear that Snap restricts the paid promotion of political messaging to traditional ad formats. This is in order to be responsible to our community and to maintain transparency.

Demonetisation of disinformation content

The policies and other mitigations highlighted in this Section ensure that the placement of advertising does not provide financial incentives for the dissemination of disinformation and FIMI with regards to electoral processes and hateful, (violent) extremist or radicalising content that can influence individuals in their electoral choices.

Integrity of services

As explained in this Section, we have appropriate procedures to ensure the timely and effective detection and disruption of manipulation of the service when this has been identified by them as a relevant systemic risk, taking into account the best available evidence. We explicitly prevent the use of “any robot, spider, crawler, scraper or other automated means or interface to access” Snapchat; use of Snapchat “in any manner that could interfere with, disrupt, negatively affect or inhibit other users from fully enjoying” Snapchat and any “attempt to circumvent any content-filtering techniques we employ” on Snapchat.

When we determine that a user has violated our Terms, we may remove the offending content, terminate the relevant account, and/or notify law enforcement. We may also briefly limit the visibility of content suspected of being illegal or otherwise violating our terms if needed to enable time for human moderators to review and provide confirmation (known as “temporary soft removal”). See Section 5.5 for more information.



Third party security and research

The severity of these risks is reflected in the resourcing Snapchat has committed to partnerships and collaborations with leading researchers and civil society organizations who are analyzing threats to democratic information environments, including the Atlantic Council's Digital Forensics Research Lab, the Center for a New American Security, the Stanford Cyber Policy Center and University of Washington, and the Poynter Institute (which is also secretariat for the International Fact-Checking Network). Also reflecting its serious approach to this risk, Snapchat has agreed on voluntary rules for the 2021 Dutch elections in a Code of Conduct, which governs transparency commitments regarding online political advertisements during election campaigns (see below for more detail).¹⁴⁵

Snap is also subject to audit under the DSA which includes a review of Snap's compliance with its risk assessment and mitigation obligations.

Fundamental rights

As set out in [Section 4.2 of this Report](#), in line with the requirements of the DSA, when assessing its risks and mitigations, Snap has paid due regard to:

- the protection of **fundamental rights** enshrined in the Charter of Fundamental Rights of the European Union, in particular the right to freedom of expression and of information; and
- the impact of measures to tackle illegal content such as public incitement to violence and hatred to the extent that such illegal content may inhibit or silence voices in the democratic debate, in particular those representing vulnerable groups or minorities.

As explained above, our platform does not, for example, widely distribute an unvetted feed of algorithmically curated political information. Under our Content Guidelines for Recommendation Eligibility, Political content is also not eligible for promotion in Spotlight, limiting the ability of any user to share political content with strangers on Snapchat, unless it's from trusted news partners and creators, and pre-moderate that surface to ensure that other such political content is not distributed.

Mitigation measures linked to generative AI

Snap maintains robust policies – applicable to both the dissemination and the creation of generative AI content – that function to mitigate risk and advance safety.

Creation

On-platform features for creating generative AI content are not part of Snap's inscope services and are out of scope of this Report (save for certain commonplace ad creation tools). Nevertheless, outside of its DSA obligations, we note that Snap has developed several internal policies relating to generative AI. In particular,

¹⁴⁵ The Dutch Code of Conduct Transparency Online Political Advertisements, [url](#).



- (1) Content and Product policies: We have developed a suite of policies that disallow the generation of harmful content (including deceptive political content). Our policy and moderation teams work in partnership with engineering and data science colleagues to ensure that our AI products are responsibly trained on these policy parameters.
- (2) Acceptable Use: We have similarly developed Acceptable Use Policies that prohibit the use of our AI tools to attempt to generate violative content at the prompt-level.

These aligned very closely with the rules for content dissemination, which are explained below.

We have also introduced contextual symbols to help our community understand when they are interacting with Snap generated AI content. We have created a [generative AI support page](#) to explain our approach to Snapchatters and other stakeholders.

Dissemination

In the context of dissemination of content on Snapchat's online platform, in scope of the DSA, we understand well that online platforms may have a negative effect on the electoral processes and the exercise of political rights by amplifying digital disinformation or deceptive content relating to political matters or processes.

Our [Community Guidelines](#) and [Terms of Service](#) set out the rules on what content is allowed on Snapchat. They are focused on preventing harm to Snapchatters and the broader community from content and behaviour, whether or not caused by generative AI or any other form of IT tools (such as Photoshop). These rules apply to all content formats across our platform, including content that is AI-generated. While the rules are agnostic to content format or creative tools, the Community Guidelines specifically note: "We implement safeguards designed to help keep generative AI content in line with our Community Guidelines, and we expect Snapchatters to use AI responsibly. We reserve the right to take appropriate enforcement action against accounts that use AI to violate our Community Guidelines, up to and including the possible termination of an account."

Our rules and internal enforcement guidance include clear provisions related to content risks for Civic Discourse and electoral processes. In particular, our Community Guidelines prohibit spreading false information that causes harm or is malicious, such as denying the existence of tragic events, unsubstantiated medical claims, undermining the integrity of civic processes, or manipulating content for false or misleading purposes (whether through generative AI or through deceptive editing).

As technologies have evolved, we have updated our policies to cover all content formats – whether created by a human or generated by artificial intelligence. Our Community Guidelines rules on false information refer to a more detailed [Explainer](#) that prohibits content that undermines the integrity of civic processes, or deep fake content or other media that is



manipulated for false or misleading purposes. The Community Guidelines further explain that these prohibitions extend to the following types of harmful content:

- Procedural interference: misinformation related to actual election or civic procedures, such as misrepresenting important dates and times or eligibility requirements for participation.
- Participation interference: content that includes intimidation to personal safety or spreads rumours to deter participation in the electoral or civic process.
- Fraudulent or unlawful participation: content that encourages people to misrepresent themselves to participate in the civic process or to illegally cast or destroy ballots.
- Delegitimization of civic processes: content aiming to delegitimize democratic institutions on the basis of false or misleading claims about election results, for example.

Sharing such content will violate Snap's Community Guidelines irrespective of whether it is AI-generated or user-generated, or whether it is generated on Snapchat or on another platform.

Snap enforces these Community Guidelines fairly and consistently, using internal policies and guidelines, and applies outcomes that are commensurate with the severity of risk. Accounts that we determine are used to perpetrate serious, high-severity harms will immediately be disabled. For other violations of our Community Guidelines, Snap generally applies a three-part enforcement process:

- Step one: the violating content is removed.
- Step two: the Snapchatter receives a notification, indicating that they have violated our Community Guidelines, that their content has been removed, and that repeated violations will result in additional enforcement actions, including their account being disabled.
- Step three: our team records a strike against the Snapchatter's account.

A strike creates a record of violations by a particular Snapchatter. Every strike is accompanied by a notice to the Snapchatter; if a Snapchatter accrues too many strikes over a defined period of time, their account will be disabled.

This strike system ensures that Snap applies its policies consistently, and in a way that provides warning and education to users who violate our Community Guidelines. The primary goal of our policies is to ensure that everyone can enjoy using Snapchat in ways that reflect our values and mission; we have developed this enforcement framework to help support that goal at scale.

Snap has a suite of internal policies and guidelines to help our content review and trust and safety teams apply the Community Guidelines to user generated content disseminated via our online platforms (such as Spotlight and Discover). They provide more granular information for our content review teams. [REDACTED]

[REDACTED]

[REDACTED]



In preparation for the 2024 European Parliament elections, we also signed up to the [AI Elections Accord](#), alongside other technology firms, where we pledged to work collaboratively on tools to detect and limit the spread of AI generated content which aims to deceive voters. However, as noted above, Snap's product design and policy practices outlined above have been demonstrated to be effective in mitigating the risks of deceptive political content, including content generated using AI tools, from achieving meaningful scale on Snapchat and substantially reducing the likelihood of negative impacts on democracy. As noted above, all of our ads are subject to review, and political ads are also subject to fact checking. Deceptive ads are rejected, irrespective of whether they use AI, photoshop, or other digital editing tools.

We continue to detect and monitor risks as outlined in [Section 6](#) of this Report (including working with our Safety Advisory Board on the intersection of safety and generative AI technology) and adapt our mitigations accordingly.

Cooperation with national authorities, independent experts and civil society organisations

Snap has closely followed the negotiations on the EU AI Act and plans to continue to actively engage and assess collaboration opportunities on the upcoming AI Act, as well as on the drawing of the related codes of practice for providers of general-purpose AI models and those regarding the detection and labelling of artificially generated or manipulated content.

More broadly, tackling risks stemming from generative AI requires (among others) broad industry-wide technical solutions which have not been clearly identified so far. This is why Snap is actively engaging with its peers and industry experts in different fora to share best practices and advance the technical debate. These partnerships, industry collaborations and efforts include:

- OpenAI: Although My AI is out of scope of this Report, the fact that My AI is powered by OpenAI's ChatGPT, has led to a good working partnership with OpenAI. This allows the companies to share best practices, including with respect to content moderation.
- Tech Coalition / Working Groups on Generative AI: Snap is a member of the Tech Coalition's Working Group on Generative AI Content, and a member of the GenAI Briefing Subgroup. The Working Group on Generative AI Content meets regularly to facilitate dialogue and information- and idea-sharing around mitigating content-level generative AI risks. The GenAI Briefing Subgroup meets periodically to plan expert briefings for Tech Coalition members on topics related to Generative AI risks; such briefings have included representatives from government, law enforcement, civil society, and the research community.
- Tech Accord to Combat Deceptive Use of AI in 2024 Elections: Snap was an initial signatory to the Tech Accord to Combat Deceptive Use of AI in 2024 Elections. This compact seeks to set expectations for how signatories will manage the risks arising from deceptive AI election content created through their publicly accessible, large-scale platforms or open foundational models, or distributed on their large-scale social or publishing platforms in line with their own policies and practices as relevant to the



commitments in the accord. The Accord was announced at the Munich Security Conference in February 2024.

- *ITI AI Futures Initiative*: Through its membership in the Information Technology Industry Council (ITI), Snap has participated alongside other private sector actors in the [AI Futures Initiative](#). Led by technical and policy experts spanning the tech ecosystem, the Initiative is a forum through which participants are developing action-oriented recommendations for AI policy and working to address emerging questions around AI. Deliverables to date have included the issuance of [Global AI Policy Recommendations](#) to help guide governments around the world as to develop responsible regulatory approaches to AI-related issues.
- *HackerOne - Red-Teaming Collaboration*: Snap partnered with HackerOne on red teaming exercises to test the strict safeguards Snap has in place around AI. Together with HackerOne, we made significant developments in the methodology for AI safety red teaming that has led to a more effective approach to surfacing previously unknown problems. We refer to the HackerOne blog for more details: <https://www.hackerone.com/ai/safety-vs-security>.
- As an active member of the [EU Internet Forum](#), Snap will support the upcoming dedicated working group on generative AI matters.
- We are also members of the [Centre for Information Policy Leadership \(CIPL\)](#) and the Future of Privacy Forum (FPF) which work with industry stakeholders (like Snap), NGOs and government agencies in each region to advance a broad array of information topics. CIPL has been a leader in AI matters for many years through its dedicated AI Project and specific Brazilian AI Project. Most recently, in Europe, CIPL has responded to the UK Information Commissioner's Office (ICO)'s consultations on Generative AI, and led various forums on Accountable Governance of AI and AI Regulation in Brussels and the UK. Similarly, FPF is working on AI Governance and other responsible Gen AI initiatives.

Further, we actively engaged in the Commission's public consultation on its proposed DSA Election guidelines, and similar consultations and queries raised by national DSCs. As shown above, we have worked to update our risk assessment to take into account the recommendations in those guidelines.

Specific Mitigations

In addition to the detailed highlights above, in the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive. As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **link in the left hand column to a full summary** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.



Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	As set out in this Section of the Report “<i>Fundamental Rights</i>”, in line with the requirements of the DSA, when evaluating design options by assessing its risks and mitigations, Snap has paid due regard to the protection of users’ Fundamental Rights. Several aspects of Snapchat’s design and function reduce this risk of negative effects on the democratic and electoral process. • Unlike many of our peers, Snapchat does not offer an open news feed where unvetted publishers or individuals have an opportunity to broadcast harmful and violative content, does not offer a broad ‘reshare’ functionality that would encourage virality, and does not allow user-generated content to be recommended to a wide audience without going through further review. • In particular, Snapchat does not, for example, widely distribute an unvetted feed of algorithmically curated political information. Under our Content Guidelines for Recommendation Eligibility, Political content is also not eligible for promotion in Spotlight, limiting the ability of any user to share political content with strangers on Snapchat, unless it’s from trusted news partners and creators, and pre-moderate that surface to ensure that other such political content is not distributed.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Snap’s policies expressly prohibit content that undermines the integrity of elections and civic processes in accordance with expert research from the Election Integrity Partnership¹⁴⁶. See “<i>Cooperation with national authorities, independent experts and civil society organisations</i>” above for more details. We take steps to explain our policy approach to safeguarding democratic information environments through our Community Guidelines and periodic blog posts. As technologies have evolved, we have updated our policies to cover all content formats – whether created by a human or generated by artificial intelligence (including deep fakes).

¹⁴⁶ Election Integrity Partnership, ‘*Evaluating Platform Election-Related Speech Policies*, October 2020, [url](#).



Mitigation Category	Applies to this risk?
	Snap does not allow Lenses that encourage a particular political perspective. In line with this approach, politically related Lenses are disabled in Discover. Snap also rejects Lenses that perpetuate false information to elections (e.g. the wrong date). AR moderators are given strict guidance during elections to escalate misinformation. When we determine that a user has violated our Terms, we may remove the offending content, terminate the relevant account, and/or notify law enforcement. We may also briefly limit the visibility of content suspected of being illegal or otherwise violating our terms if needed to enable time for human moderators to review and provide confirmation (known as “temporary soft removal”). See Section 5.5 for more information.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, we have specific proactive and reactive moderation procedures to prevent and remove violative content. As explained in the Moderation section in Section 5, Snap deploys a range of automated content moderation (which include abusive language detection, other keyword-based detection, and machine-learning-based proactive detection models) to scan Stories and Spotlight submissions. Snapchatters can report violative content to us via in-app reporting options and anyone can submit a report through the Snapchat Support Site. Our in-app reporting tool also allows users to directly report violative content. Content that is approved for broader audiences must comply with both our Community Guidelines and our Content Guidelines for Recommendation Eligibility. All Spotlight and non-professional user generated Discover content goes through both automoderation, and often human review, against these guidelines before it is eligible for recommendation to a wide audience. As an additional safeguard, we monitor content that is achieving large-scale reach (and ensure a human reviews it) as a sort of “virality circuit breaker” and a means of checking that our pre-moderation systems are working effectively. Any content that is reported will be reviewed against the guidelines again for compliance.



Mitigation Category	Applies to this risk?
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	For more details on our algorithmic systems and our measures to regulate how political content is distributed on the platform see “ <i>Measures to provide users with more contextual information</i> ” and “ <i>Recommender Systems</i> ” above, in this Section of the Report.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Our political ad policies ensure that any political advertisements are subject to review and fact-checking <i>before</i> they are eligible for placement on Snapchat. For more information on our advertising policies and partnerships aimed at preventing violative advertising practices see “ <i>Political Advertising</i> ” and “ <i>Demonetisation of disinformation content</i> ” in this Section of the Report.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, for example, Snap runs specific prevalence testing and transparency reporting which we use to help detect and manage violations of our Community Guidelines and policies. Snap is also subject to audit under the DSA which includes a review of Snap’s compliance with its risk assessment and mitigation obligations.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Yes, see the information relating to the Guidelines above, under “Internal Process” and “Access to official information on the electoral process and Media literacy initiatives” and “Cooperation with national authorities, independent experts and civil society organisations” and “Third party security and research”
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	We engage with experts in the information integrity and democracy and human rights community to inform our approach. This includes collaborations and engagement with individual experts (such as former US Ambassador to the UN Human Rights Council Eileen Donahoe, global democracy scholar and Stanford Professor Larry Diamond, and several others), as well as think tanks (such as the Atlantic Council’s Digital Forensics Research Lab) and research collaborations (such as the Election Integrity Partnership). Additionally, we partner with governments around the world to inform Snapchatters about elections and invite them to go vote. More information is also available above, under “ <i>Internal Process</i> ” and “



Mitigation Category	Applies to this risk?
	<i>Access to official information on the electoral process and Media literacy initiatives” and “Cooperation with national authorities, independent experts and civil society organisations” and “Third party security and research”.</i>
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on our terms, harms, moderation and enforcement practices (see the Annex), as well as how to get help in our Safety Center and via in-app resources (Here For You and Safety Snapshot). We make available robust reporting tools; and we provide guidance to parents on the web (see below). We proactively encourage our users to go to vote through interactive campaigns Access to official information on the electoral process and Media literacy initiatives" section, Snap regularly partners with governments and election authorities to encourage voter participation, using creative tools like AR Lenses and quizzes. Past collaborations include France, the Netherlands, and Sweden, with recent examples in the 2023 Dutch provincial elections and the 2024 European Parliament elections, where Snap reached young voters across the EU with interactive reminders to vote.
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	We limit exposure to political content to Teens, but do educate Teens with trusted new sources on current events and inform users how they can participate in a democratic society. We offer Family Center; we make available robust reporting; and we provide guidance to parents on the web. Our parents site provides additional guidance for parents and carers on risks and support.¹⁴⁷
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their	We recognise the risk that generative AI could be used to generate harmful false misinformation, including deep fakes. Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat’s inscope services. More information is also available in this Section of the Report under “<i>Mitigation measures linked to generative AI</i>”.

¹⁴⁷ <https://parents.snapchat.com>.



Mitigation Category	Applies to this risk?
online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	

Conclusion

Snap considers the overall risk potential of negative impact on democratic or electoral processes to be in the Level 3 category, given severity of potential harm. However, as described above, Snap has put in place numerous specific mitigations, such as algorithmically preventing the promotion of political content in Spotlight, enforcing political ad policies, and disallowing Lenses encouraging political perspectives. Further, the design and function of Snapchat is such that it is not conducive for the widespread distribution of viral content and we provide robust in-app reporting, which further mitigates this harm. Snap recognizes the importance of Democratic and Electoral Processes, and in fact has created interactive campaigns to raise awareness and encourage users to vote. Our prevalence data and our continuing monitoring efforts cited above show that our safeguards are effective at mitigating these risks on Snapchat. We have taken into account the Commission’s recommendations set out in the Guideline for providers of VLOPs and VLOSEs on the mitigation of systemic risks for electoral processes when carrying out our assessment.

We have concluded therefore that Snapchat’s in-scope services have reasonable, proportionate, and effective mitigation measures for the risk of Negative Effects on Democratic and Electoral Processes. There is no change in this conclusion from our 2024 Report.

4.3.2 Negative Effect on Civic Discourse

We recognize that without adequate mitigations, digital content platforms like Snapchat can contribute to Negative Effects on Civic Discourse. Across Snap’s various products, these risks could include:

- The potential for personalized content and algorithmic biases lock users into echo chambers, reinforcing existing beliefs and potentially leading to polarized communities, which hinders open dialogue.
- The risk of amplified dis- and misinformation negatively impacting public opinion on important civic issues.
- The possibility of amplification of extreme or sensational content to retain user attention leading to heightened polarization and a hostile online environment.



[REDACTED]

Likelihood

All of the risks we track on Snapchat have a low prevalence compared to the prevalence of these issues elsewhere online and offline. To aid our prioritization, our methodology seeks to assess the relative likelihood between the risks we track (even though all are low in absolute terms).

While it is difficult to quantify the likelihood of this social issue, we can draw on the reporting data available to us. Although we have not encountered any specific reports identifying this exact risk on Snapchat, we relate it to existing risks linked to harmful false information (see [Section 4.1.10](#)) and Hate Speech (See [Section 4.1.2](#)).

Our own reporting data indicates that violations in these categories - which we consider potential sources of harm to civil discourse - are rare. In particular, in our 2024 Report, the [prevalence](#) of harmful false information [REDACTED] and illegal hate speech were measured at an extremely low percentage. As of April 2025, we have subsequently observed a further substantial decrease in the prevalence of content falling within these categories. [REDACTED]

Moreover, when content related to Hate Speech or False Information was detected (whether proactively through automated tools or reactively following a user report), our Safety teams acted swiftly. According to our latest European Union Transparency Report (covering the second half of 2024), the median turnaround time for enforcement action was 36 minutes for Hate Speech and 2 minutes for False Information.

[REDACTED]

[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

Consequently, we continue to conclude that this risk still falls within our **Extremely Low likelihood category**.



Severity

At Snap, we take these risks seriously and classify it as such, despite the apparent lack of prevalence on our platform. Credible studies have shown these risks to fall within our **severe**. For example, a study from the European Parliament¹⁴⁸ concludes that the consequences of disinformation are extensive and disruptive to democratic societies. Disinformation - particularly when widespread on social media - undermines public trust in institutions, distorts political discourse, and exacerbates social divisions. The study emphasizes that disinformation doesn't merely mislead individuals; it actively reshapes public opinion and decision-making by manipulating emotions, reinforcing bias, and eroding the shared foundation of facts that democratic debate depends on.

Moreover, according to the study, beyond its effects on individual users, disinformation also compromises the integrity of elections and weakens democratic institutions. It can be weaponized by both domestic and foreign actors to destabilize political systems, influence voter behavior, and sow discord within and across communities.¹⁴⁹

Accounting for the documented risks of these harms across other platforms, we understand that the digital platforms play a significant role in potentially negatively impacting civil discourse. However, we assess that our platform architecture and approach to information integrity combine to provide unique safeguards against these threats to civil discourse.

DSA Risk Factors

In accordance with Article 34(2), our risk assessment also addresses our recommender systems, content moderation systems, applicable terms and conditions, systems for the selection and presenting of advertisements and any of our data-related practices. These factors have been listed and explained in Section 3 on the [Risk Assessment Methodology](#) and applied throughout [Section 4](#). We also considered the risk factors in the context of Negative Effects on Civic Discourse. As many of these factors pose similar risks, and are mitigated in a horizontal manner (i.e. by measures that cut across all systemic risks), we did not include them in this section again but refer to our explanation in [Section 4.1.2](#) on on Illegal Hate Speech, [Section 4.1.10](#) on Harmful False Information, and [Section 4.3.1](#) on Negative Effects on Democratic and Electoral Processes.

Overall potential risk prioritization

Snap considers the dissemination of information with actual or foreseeable Negative Effects on Civic Discourse to fall within our severe harm category. [REDACTED]

[REDACTED]

¹⁴⁸ [REDACTED]

[REDACTED].

¹⁴⁹ [REDACTED]

[REDACTED]



Snap's Mitigations

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive.

As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **link in the left hand column to a full summary** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, as outlined in the section on Democracy our platform does not, for example, provide an unvetted feed of algorithmically curated political information; we disallow <i>all</i> political content from Spotlight (our broadcast platform for User Generated Content) and pre-moderate that surface to ensure that such political content is not distributed. In addition, many of our surfaces are not ideal vehicles to cause risks to civil discourse. For example, unless saved to your Public Profile, Public Stories and Snaps on the Map are only available for a maximum of seven (7) days (and often much shorter), which limits their arc of influence. Similarly, there is considerable technical expertise required to create a Lens, making it a difficult surface



	(compared to other third party platforms) to navigate for the purpose of broadly distributed harm.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes, we take steps to explain our policy approach to safeguarding civil discourse information environments through our Community Guidelines and periodic blog posts. Snap's policies prohibit the spread of "Harmful False Information," which we define as false content that may result in broadly distributed harm, or is malicious. Referencing our internal policy guidance, Snap enforces content as Harmful False Information if both of the following elements are present: • Information is determined to be false • The false information could cause "broadly distributed harm". "Broadly distributed harm" refers to harms that undermine societal- or community-level safety or security; harms that undermine public health; harms that undermine civic processes or the exercise of political rights; and harms that denigrate the memory or history of peoples and tragic events. In addition to our internal policies, Snap's Community Guidelines also note that Harmful False Information is prohibited and includes denying the existence of tragic events, unsubstantiated medical claims, or undermining the integrity of civic processes – all of which could contribute to negative impacts on Civic Discourse. Snap policies also prohibit the use of Hate Speech, hate symbols, and/or content that valorizes the perpetrators of, or denigrates the victims of, human atrocities such as genocide. We define Hate Speech as content that demeans, or promotes discrimination towards, an individual or group of individuals on the basis of their race, color, caste, ethnicity, national origin, religion, sexual orientation, gender identity, disability, veteran status, immigration status, socio-economic status, age, weight, or pregnancy status. Our policies note that Hate Speech may include references to people that are dehumanizing or that compare humans to animals on the basis of these traits and categories. Hate Speech also includes the valorization of perpetrators—or the denigration of the victims—of hateful atrocities (e.g., genocide, apartheid, slavery, etc.), as well as the promotion of hate symbols. Under Snap's policies, hate symbols include imagery that is intended to represent hatred or discrimination toward others,



	including those featured in the hate symbols database maintained by the Anti-Defamation League (ADL). ¹⁵⁰
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, all Spotlight content goes through both automoderation and human review before it is eligible for distribution to a wide audience. Content that is approved for broader audiences must comply with our Community Guidelines and our Content Guidelines for Recommendation Eligibility. Any content that is reported will be reviewed against these guidelines again for compliance.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, our algorithmic systems mitigate risks to Civic Discourse; this includes the absence of algorithmically promoted groups, which have been shown to contribute to echo chambers and to be vectors for misinformation, with negative consequences for civil discourse. ¹⁵¹
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, our political ad policies ensure that any political advertisements are subject to review and fact-checking <i>before</i> they are eligible for placement on Snapchat. We prevent advertisers from manipulating small audiences with micro targeted campaigns, particularly for political ads, by requiring an audience of at least [REDACTED]
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, for example we have specific prevalence testing and monitoring moderation and enforcement data which we use to help detect and manage harmful false misinformation. To remain vigilant against threats to civil discourse, Snap engages with experts from across civil society and the research community who study information integrity and resilience to online harms. These engagements include consultations and collaborations with online safety experts (including those represented on Snap's Safety Advisory Board), with organizations combating online hate (such as the Anti-Defamation League), and engagement with research organizations, including the Atlantic Council Digital Forensics Research Lab and the Digital Wellbeing.

¹⁵⁰ The ADL database is available at: [url](#).

¹⁵¹ The Verge, 'Facebook will stop recommending health groups', September 2020, [url](#).



Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Yes, we cooperate with trusted flaggers, our trusted flaggers may also report misinformation, but this rarely happens because of the limited amount of misinformation on the platform.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes. We engage with experts in the information integrity and democracy and human rights community to inform our approach. This includes collaborations and engagement with individual experts (such as former US Ambassador to the UN Human Rights Council Eileen Donahoe, global democracy scholar and Stanford Professor Larry Diamond, and several others), as well as think tanks (such as the Atlantic Council's Digital Forensics Research Lab) and research collaborations (such as the Election Integrity Partnership). Snap also works closely with French regulator Arcom, which monitors industry action against misinformation. We have also worked closely with the Commission and other stakeholders during the recent EU elections.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	We proactively encourage our users to go to vote through interactive campaigns Access to official information on the electoral process and Media literacy initiatives" section, Snap regularly partners with governments and election authorities to encourage voter participation, using creative tools like AR Lenses and quizzes. Past collaborations include France, the Netherlands, and Sweden, with recent examples in the 2023 Dutch provincial elections and the 2024 European Parliament elections, where Snap reached young voters across the EU with interactive reminders to vote.
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, we limit exposure to political content to Teens, but do educate Teens with trusted new sources on current events and inform users how they can participate in a democratic society. We offer Family Center; we make available robust reporting; and we provide guidance to parents on the web. Our parents site provides additional guidance for parents and carers on risks and support.¹⁵²
Content Authenticity	We recognise the risk that generative AI could be used to generate harmful false misinformation, including deep fakes. Snap has taken

¹⁵² <https://parents.snapchat.com>.



Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services. We also label political advertisements, and maintain a political ads library.
--	---

Conclusion

Snap categorizes the risk of negative impacts on civil discourse as **Level 3**, given the potential harm from dis/misinformation and online echo chambers. Mitigations, many overlapping with those for democracy and elections, include proactive content moderation, enforcement of Community Guidelines and Terms, restricting political content, and engagement with outside experts.

We also take proactive steps, such as encouraging voting and participation in civil discourse, and setting audience minimums to prevent ad microtargeting. We monitor and act on Harmful False Information (Section 4.1.10) and Illegal Hate Speech (Section 4.1.2), provide in-app tools to report content, and hold advertisers to standards that prohibit false or misleading ads. Our goal is to prevent harmful content from reaching wide audiences while supporting positive, respectful engagement on Snapchat's in-scope services.

As a result, we have concluded that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures against Negative Effects on Civic Discourse. There is no change in this conclusion from our 2024 Report.

4.3.3. Negative Effect on Public Security

Without appropriate mitigations, we recognise that digital platforms may present risks to Public Security, particularly in the form of harmful, dangerous, or inciteful content; these risks may become compounded when such content may be amplified at great scale and distributed with high velocity. The design of Snap's products and platform architecture scrupulously accounts for these risks; accordingly, we've implemented a number of key safeguards that help to advance both the safety of Snapchatters and the interests of Public Security across our services.



Likelihood

All of the risks we track on Snapchat have a low prevalence compared to the prevalence of these issues elsewhere online and offline. To aid our prioritization, our methodology seeks to assess the relative likelihood between the risks we track (even though all are low in absolute terms).

To assess the likelihood of Snapchat's inscope services having a Negative Effect on Public Security, we have reviewed the sources of data relating to the following illegal and otherwise violating content categories considered to have a particular impact on undermining Public Security:

<i>Category</i>	<i>Relative likelihood of risk occurring on Snapchat</i>
4.1.2 Dissemination of Illegal Hate Speech	Lowest Relative Likelihood
4.1.4 Dissemination of Terrorist Content	Lowest Relative Likelihood
4.1.9 Dissemination of content encouraging or engaging in Violent or Dangerous Behavior	Lowest Relative Likelihood
4.1.10 Dissemination of harmful false misinformation	Lowest Relative Likelihood

The low prevalence rate of these harms supports our continued assessment that the volume of content presenting risks to Public Security is quite low on Snapchat, and, consequently, it is uncommon to encounter these harms on Snapchat. In terms of likelihood, this risk would fall within our **Extremely Lowlikelihood category**.

Severity

We have assessed that digital platforms can have consequential negative impacts on Public Security. For example, harmful and dangerous content such as Hate Speech, violent or gore videos and dis- and misinformation can be shared on the platform, contributing to (the incitement of) real-world harm. In light of these risks, we have undertaken important steps to minimize the likelihood of such impacts for Snap's public surfaces.

Snap considers the dissemination of information with actual or foreseeable Negative Effects on Public Security to be a serious risk. For this reason, our policies and practices reflect a commitment to safety, and identify acute areas of risk, such as imminent threats to public places and content related to terrorism and violent extremism.

Security and safety, as well as the dignity of victims, are seriously threatened by the proliferation



of terrorist materials online and the rising accessibility of such material.¹⁵³ The European Commission has outlined that online extremist propaganda has been crucial in radicalizing and motivating so-called 'lone wolves' to carry out recent terrorist strikes in Europe.¹⁵⁴ Research suggests that - while not surpassing offline radicalization - there is an increase in online radicalisation among young people and women.¹⁵⁵ The European Commission has pointed out that Terrorist Content not only has profound detrimental effects on individuals and society as a whole, but that it also erodes internet users' trust and harms the business plans and reputations of those companies that are affected.¹⁵⁶

The Global Internet Forum to Counter Terrorism (GIFCT)'s 2024 transparency report flagged how the GIFCT shared 120 reports, 62 of which highlighted attempts by terrorists or violent extremists to exploit GIFCT member platforms.¹⁵⁷

Non-terrorist threats of violence are especially high risk as well. Data from the US National Center for Education Statistics shows an increase in school shootings with casualties across public and private schools in the US.¹⁵⁸ In 2020-21, there were a total of 93 school shootings, the highest number in two decades. Victims of school-related violent deaths may include not only students and staff, but also others like students' parents and community members.¹⁵⁹ For example, the experience of school shootings is shown to have severe negative effects on surviving youth, with long lasting consequences to their mental health, educational and economic trajectories.¹⁶⁰

Real-world examples show the potential impact of the dissemination of harmful content (e.g. Hate Speech) on Public Security, e.g. amplifying hate-crimes against Rohingya muslims in Myanmar,¹⁶¹ fueling anger of Trump supporters later on contributing to the Capitol Riots,¹⁶² and riots in London fuelled by misinformation and far right supporters.¹⁶³

¹⁵³ European Commission, *Tackling Illegal Content Online Towards an enhanced responsibility of online platforms* (COM(2017) 555 final), [url](#), 28 September 2017.

¹⁵⁴ European Commission, 'Proposal for a Regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online' (COM(2018)640 final), [url](#), 12 September 2018.

¹⁵⁵ N. Hamid & C. Ariza, 'Global Network on Extremism and Technology. Offline Versus Online Radicalisation: Which is the Bigger Threat?', [url](#), 21 February 2022.

¹⁵⁶ European Commission, 'Proposal for a Regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online' (COM(2018)640 final), [url](#), 12 September 2018.

¹⁵⁷ GIFCT, 'Transparency Report 2024', [url](#), 2024.

¹⁵⁸ V. Irwin, K. Wang, J. Cui, A. Thompson, 'National Center for Education Statistics. Report on indicators of School Crime and Safety 2021', [url](#), June 2022.

¹⁵⁹ V. Irwin, K. Wang, J. Cui, A. Thompson, 'National Center for Education Statistics. Report on indicators of School Crime and Safety 2021', [url](#), June 2022.

¹⁶⁰ Stanford Institute for Economic Policy Research (SIEPR), *Surviving a school shooting: Impacts on the mental health, education, and earnings of American youth*, [url](#), June 2022.

¹⁶¹ Amnesty International, 'Myanmar: The social atrocity: Meta and the right to remedy for the Rohingya', September 2022, [url](#).

¹⁶² The Washington Post, 'Inside Facebook, Jan. 6 violence fueled anger, regret over missed warning signs', October 2021, [url](#).

¹⁶³ BBC, The real story of the news website accused of fuelling riots, [url](#).



To assess the likelihood of Snapchat's inscope services having a Negative Effect on Public Security, we have reviewed the sources of data relating to the following illegal and otherwise violating content categories considered to have a particular impact on undermining Public Security:

Category	Relative likelihood of risk occurring on Snapchat
4.1.2 Dissemination of Illegal Hate Speech	Significant harm industry wide
4.1.4 Dissemination of Terrorist Content	Serious harm industry wide
4.1.9 Dissemination of content encouraging or engaging in Violent or Dangerous Behavior	Significant harm industry wide
4.1.10 Dissemination of Harmful False Misinformation	Significant harm industry wide

As these range from significant to serious, and given the context outlined above showing serious consequences of a Negative Effect on Public Security, we assess that this category would fall within a **'serious' harm category**.

DSA Risk Factors

In accordance with Article 34(2), our risk assessment also addresses our recommender systems, content moderation systems, applicable terms and conditions, systems for the selection and presenting of advertisements and any of our data-related practices. These factors have been listed and explained in Section 3 on the [Risk Assessment Methodology](#) and applied throughout [Section 4](#). We also considered the risk factors in the context of Negative Effects on Public Security. As many of these factors pose similar risks, and are mitigated in a horizontal manner (i.e. by measures that cut across all systemic risks), we did not include them in this section again but refer to our explanation in [Section 4.1.2](#) on Hate Speech, [Section 4.1.4](#) on Terrorist Content, [Section 4.1.9](#) on Violent and Dangerous Behaviour, and [Section 4.1.10](#) on Harmful False Information.

Overall potential risk prioritization

Taking into account the real-world examples illustrating the potential disruptive effects on Public Security, this is a severe risk if not mitigated. However, we are encouraged—based on relevant prevalence data cited above—that our safeguards are substantially effective at mitigating these risks on Snapchat. The combination of low prevalence and severe nature results in a **Level 3** overall potential risk prioritization categorisation. There is no change in this assessment from our 2024 Report.



[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]		[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]				

Snap's Mitigations

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive.

As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **link in the left hand column to a full summary** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Snapchat is not an attractive platform for spreading content that may have a negative impact on Public Security, including harmful, dangerous, and inciteful content, in particular because it is difficult to reach a broad audience, and Snapchat has made conscious design decisions to restrict the ability for content to go viral, including not offering a reshare functionality and applying short retention to content. On surfaces where a broader audience can potentially be reached, our proactive detection makes it difficult for content that may have a negative impact on Public Security to reach a large audience. Moreover, our content platform, Discover, features content from approved media publishers and content creators. Our entertainment platform for user-generated content, Spotlight, is proactively and a priori moderated before content can reach a wide audience.



Mitigation Category	Applies to this risk?
	In addition, as noted in the Civic Discourse section, many of our surfaces are not ideal vehicles to cause risks to Public Security. For example, unless saved to your Public Profile, Public Stories are only available for a maximum of seven (7) days (and often much shorter), which limits their arc of influence. Similarly, there is considerable technical expertise required to create a Lens, making it a difficult surface (compared to other third party platforms) to navigate for the purpose of broadly distributed harm.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Snap's terms prohibit content that may have a negative impact on Public Security, including harmful, dangerous, and inciteful content, and they are strictly enforced. We take several steps to ensure that we are addressing this risk across Snap's products and services, including enforcement of several relevant platform policies and internal crisis protocols for managing high-risk scenarios. Snap's policies include several prohibitions that are enforced vigorously and equitably to support the interests of Public Security. These policies include a prohibition against spreading Harmful False Information. Internal policy guidance instructs that violations of these policies include risks to Public Security such as Snaps denying the Holocaust or a school shooting, or information obtained illegally that is being shared to embarrass the person from whom the information was stolen. Snap's policies also include prohibitions on content promoting terrorism or violent extremism, as well as "content that attempts to incite, glorify, or depict real violence that results in personal injury or death," and "depictions of human violence, child abuse, animal abuse, or gore." We may also consider off-platform behavior when assessing risks to Public Security. Our Community Guidelines state expressly that "Snap reserves the right to remove or restrict account access for users whom we have reason to believe, in our sole discretion, pose a danger to others, on or off of Snapchat. These include leaders of hate groups and terrorist organizations, individuals with a reputation for inciting violence, perpetrating severe harms against others, or behavior that we believe poses a threat to human life."



Mitigation Category	Applies to this risk?
	Taken together, these several policy provisions provide a basis for appropriately actioning any content that poses an acute risk to Public Security. In addition, we have internal operational protocols for responding to public crises (see Section 5.12). These protocols include the following steps: • Our vendor teams carefully apply the Community Guidelines and Content Guidelines for Recommendation Eligibility to ensure the content is assessed appropriately against our rules (for example, routinely distinguishing between <i>documenting</i> violence and <i>advocating for</i> violence). • When breaking news happens, such as ongoing violent protests, the vendor teams connect with our full-time content review team to summarize the kind of content they are encountering (e.g., violence, property damage, fires, expressions of criticism or support for various political positions), and summarize how they are currently actioning that type of content against our existing guidelines. • That summary list comes to our Platform Policy team for review. Almost all of the time, Policy's answer is that they're actioning content correctly. (To cite a recent example, in the case of French protests over the course of summer 2023, our team determined that existing policies and procedures were working as intended.) • In the event that the Platform Policy team determines that the policies are not being applied appropriately, the team will expeditiously draft clarifying guidance for vendors and content review teams. The draft guidance will be shared among relevant internal leaders for review before being distributed to operational teams. Separately, we maintain tight internal protocols for escalating Terrorist Content or other imminent threats to the appropriate legal or emergency authorities. In such cases, vendors and review teams are trained to preserve relevant information and immediately send a report to Snap's Law Enforcement Operations team, who are professionally trained to appropriately engage with legal and emergency authorities.



Mitigation Category	Applies to this risk?
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, we have specific proactive and reactive moderation procedures to prevent and remove violative content. As explained in the Content Moderation section in Section 5 of this Report, Snap deploys a range of automated content moderation (which include abusive language detection, other keyword-based detection, and machine-learning-based proactive detection models) to scan Stories and Spotlight submissions. Snapchatters can report violative content to us via in-app reporting options and anyone can submit a report through the Snapchat Support Site. Our in-app reporting tool also allows users to directly report violative content. Furthermore, all Media Partners are vetted prior to being permitted to distribute their content broadly on Snapchat by a team of editors. Media Partners go through an editorial review of their content, a reputational search (to evaluate if a Media Partner has a history damaging press, legal actions, etc.), and compliance review before they're able to distribute content.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, unlike many of our peers, Snapchat does not offer an open news feed where unvetted publishers or individuals have an opportunity to broadcast violative content, does not offer a broad 'reshare' functionality that would encourage virality, and does not allow user-generated content to be recommended to a wide audience without going through review. Our algorithmic systems do not knowingly recommend violative content content, in particular content that may have a negative impact on Public Security, including harmful, dangerous, and inciteful content. As explained in the Content Moderation section in Section 5 of this Report, on our high-reach surfaces, like Spotlight and Discover, we take a proactive approach to moderating any content that may violate these rules prior to the content being recommended to a wide audience.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of	Yes, other mitigations listed here also apply to our Advertising Systems.



Mitigation Category	Applies to this risk?
advertisements in association with the service they provide.	
<u>Ongoing Risk Detection and Management</u> Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Snap runs specific prevalence testing and transparency reporting for content that may have a negative impact on Public Security.
<u>Trusted Flaggers</u> Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Yes, we cooperate with trusted flaggers, our trusted flaggers may also report content that may have a negative impact on Public Security, but this rarely happens because of the limited amount of this type of content on Snapchat.
<u>Codes and Crisis Protocols</u> Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various industry groups.
<u>Transparency</u> Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on harms and how to get help in our Safety Center. We make available robust reporting tools; and we provide guidance to parents on the web (see below).
<u>Protection of Minors</u> Taking targeted measures to protect the rights of the child, including age verification and	Yes, we have protective measures to limit Teen contact with strangers; we offer <u>Family Center</u> , reporting, and guidance. Our new parents site provides additional guidance for parents and caregivers on risks and support. ¹⁶⁴

¹⁶⁴ <https://parents.snapchat.com>.



Mitigation Category	Applies to this risk?
parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	Yes, Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services.

Conclusion

Snap considers the negative impact to Public Security to have a Level 3 overall potential risk given the potential disruptive effects of content that can, among other things, harm, put in danger, and incite the public at large. That being said, Snap ██████████ ██████████ has put in place a range of mitigation measures to bring the likelihood of this risk from coming to fruition into the lowest category. These measures include our proactive content moderation which is designed to detect and prevent hateful, dangerous, and inciteful content from reaching a broad audience on Snapchat's in-scope services. As noted in other sections, we continue to invest in measures that prevent this type of content from reaching a broad audience on Snapchat, as well as provide our users with tools to report content to Snapchat and law enforcement, and support our community via online and in-app support tools.

We have concluded therefore that Snapchat's in-scope services have reasonable, proportionate, and effective mitigation measures for the risk of Negative Effects on Public Security. There is no change in this conclusion from our 2024 Report.



4.4 Category 4: Negative Effects on Public Health

(Article 34.1.d / DSA Recital 83)

In this part of the Report, we explain the results of our assessment on actual or foreseeable negative effects of Snapchat's in-scope services on Public Health as required by Article 34.1.d and Recital 83 of the Digital Services Act. We have assessed in particular Negative Effects on Public Health, Gender-Based Violence, Minors, as well as serious negative consequences to a person's Physical and Mental Well-Being. We have considered risks relating to the design, functioning or use, including through manipulation such as by coordinated disinformation campaigns related to Public Health or from online interface design that may stimulate behavioral addictions of recipients of the service.

Note that for all harms, where there is (1) a risk of significant damage to the physical or emotional well-being of Snapchatters, and (2) imminent, credible risk of severe harm, including threats to human life, safety, and well-being, we treat this as a severe harm and an Level 1 overall risk prioritization.

Category 4 - Negative Effects on Public Health				
Category	Relative likelihood of risk occurring on Snapchat	Harm classification industry wide	Risk Prioritization	ConclusionSnap's Mitigations
4.4.1 Negative Effects on Public Health	Extremely Low Likelihood	Severe harm industry wide	Level 3	Low Risk / Reasonable, proportionate and effective mitigations
4.4.2 Negative Effects on Gender-Based Violence	Extremely Low Likelihood	Serious harm industry wide	Level 2	Low Risk / Reasonable, proportionate and effective mitigations
4.4.3 Negative Effects on Minors	Varies	Severe harm industry wide	Level 1	Reasonable, proportionate mitigation measures have been implemented to safeguard minors from identified harms, with the Commission's guidance on Article 28 taken



[REDACTED]

				into account and continuing to inform this approach.
4.4.4 Serious Negative Consequences on Physical and Mental Well-Being	Extremely Low Likelihood	Severe harm industry wide	Level 1	Low Risk / Reasonable, proportionate and effective mitigations

4.4.1 Negative Effects on Public Health

We recognize that without adequate mitigations, digital content platforms like Snapchat could contribute to Negative Effects on Public Health. We believe the health and wellness of the public and our users is paramount to our goal to be a platform of fun and freedom of expression. [REDACTED]

[REDACTED]

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

While we believe these risks to be probable in the absence of mitigations, we assess that Snap's mitigations appreciably reduce the likelihood of encountering these harms on our in scope services.

Likelihood

All of the risks we track on Snapchat have a low prevalence compared to the prevalence of these issues elsewhere online and offline. To aid our prioritization, our methodology seeks to assess the relative likelihood between the risks we track (even though all are low in absolute terms).

With regards to coordinated disinformation campaigns related to Public Health, as well as any dissemination of content that promotes harmful/unhealthy behavior (e.g., eating disorders or other Self-Harm content), we are encouraged that available data suggests prevalence of content on Snapchat related to these risks is quite low:



- In our 2024 Report, we observed that we had a low prevalence of Harmful False Information (including health misinformation) to be extremely low and Self-Harm and Suicide content is extremely low based [prevalence testing](#).
- We are pleased to have subsequently observed a further fall in prevalence of this content,

Moreover, our testing indicates a continued year-on-year decline in the prevalence of information related to the sale of prohibited products or services on Snapchat, which covers both legal or illegal products that could potentially lead to health-related issues. At present, content associated with the sale of illegal goods has a prevalence rate of 0.0037%, while content related to regulated (but not illegal) activities stands at 0.0175%.

Therefore, we continue to assess there is an **extremely low likelihood** of Negative Effects on Public Health arising from Snapchat's in-scope services.

Severity

Accounting for real-world examples of negative impacts to Public Health, we assess this to present a severe risk for digital content platforms that must be effectively mitigated. The severity of these risks was perilously illustrated by the spread of misinformation related to the Covid-19 pandemic, which promulgated false claims about remedies; resulted in harmful behavior; and jeopardized the containment of the virus.

Reports have also shown the impact some social media might have on the mental health of Teens.¹⁶⁵ Commenting on the impact of social media on eating disorders amongst users, the US National Eating Disorders Associations clarified that social media does not cause eating disorders, however it can contribute to eating disorders because of potentially triggering content.¹⁶⁶ The European Parliament's CULT Committee also acknowledges the risk of young people encountering content that promotes eating disorders, body image dissatisfaction and distorted values and attitudes.¹⁶⁷

Moreover, and as outlined in [Section 4.1.3](#), we are conscious of the serious concerns raised by various organizations — including Dutch health experts and youth organizations — regarding the alleged illegal sale and promotion of vape-related products, particularly flavoured e-cigarettes, to minors on the platform. We continue to engage with NGOs and law enforcement authorities to adopt measures aimed at preventing such illegal sales.

¹⁶⁵ The Guardian, 'Facebook aware of Instagram's harmful effect on teenage girls, leak reveals', [url](#), 14 September 2021; E. Bozola e.a., 'The Use of Social Media in Children and Adolescents: Scoping Review on the Potential Risks', [url](#), August 2022.; A.M. Memom e.a., 'The role of online social networking on deliberate self-harm and suicidality in adolescents: A systematized review of literature', [url](#), October 2018.

¹⁶⁶ The New York Times, 'Eating Disorders and Social Media Prove Difficult to Untangle', [url](#), October 2021.

¹⁶⁷ European Parliament, requested by the CULT Committee, 'The influence of social media on the development of children and young people', [url](#), 2023.



In light of this evidence, Snapchat continues to assess the severity of digital content platforms’ risks to Public Health to be high and classifies this as falling within our ‘**severe** harm’ category

DSA Risk Factors

In accordance with Article 34(2), our risk assessment also addresses our recommender systems, content moderation systems, applicable terms and conditions, systems for the selection and presenting of advertisements and any of our data-related practices. These factors have been listed and explained in Section 3 on the Risk Assessment Methodology and applied throughout Section 4. We also considered the risk factors in the context of serious Negative Effects on Public Health. As many of these factors pose similar risks, and are mitigated in a horizontal manner (i.e. by measures that cut across all systemic risks), we did not include them in this section again but refer to our explanation in [Section 4.1.3](#) on the Sale of Prohibited Goods and Services, [Section 4.1.8](#) on Self-Harm and Suicide and [Section 4.1.9](#) on Violent and Dangerous Behaviour, and [Section 4.1.10](#) on Harmful False Information.

Overall potential risk prioritization

Snap assesses negative impacts to Public Health to present a systemic, severe risk that must be appropriately mitigated. Given the severe nature but the extremely low prevalence, this would classify as a **Level 3 risk** in terms of Snap’s overall potential risk prioritization matrix. There is no change in this assessment from our 2024 Report.

Snap's Mitigations

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat’s in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive.



As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **link in the left hand column to a full summary** of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Snapchat's in-scope services have been adapted to include proactive moderation to reduce the spread of Harmful False Information, including unsubstantiated medical claims, and the glorification of self-harm, including the promotion of self-injury, suicide or eating disorders. Snapchat offers well-being features designed to educate and empower members of the Snapchat community to support friends who might be struggling with their social or emotional well-being. These features include "Here for You" content Snap has developed with the intention of educating Snapchatters about the importance of mental health, and ways to seek support. In response to troubling search inquiries or content indicating mental or emotional distress, our products and teams intervene to surface mental health resources and support (either automatically, or at the discretion of Trust and Safety personnel). These resources are tailored to a user's geographic region. Moreover, Snapchat does not offer a marketplace for the sale of goods.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Our Terms prohibit the spread of Harmful False Information, including unsubstantiated medical claims, and the glorification of self-harm, including the promotion of self-injury, suicide or eating disorders. Our policies elaborate that such prohibited content includes any content that, for example, recommends untested therapies for preventing the spread of Covid-19; or that features unfounded conspiracy theories about vaccines. We have strict enforcement policies. However, since Snapchat is used for communication with friends and family, it is important to us that our enforcement actions do not deprive users' friends and family of important distress signals and an opportunity to intervene. Accordingly, we instruct reviewing agents that: ○ Reported depictions of suicide or self-harm that reflect an emergency situation should be removed and possibly escalated to law enforcement or emergency authorities.



	○ Content glorifying or inciting self-harm must be removed and is subject to an enforcement “strike.” ○ Depictions of self-harm or suicidal ideation that do not reflect an emergency situation are permitted so that the community of people around this person can offer help and support. Our Terms also explicitly prohibit the sale of prohibited products or services.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Specific proactive and reactive moderation processes are in place to prevent the sale of prohibited products and services and remove Harmful False Information, including unsubstantiated medical claims, and the glorification of self-harm, including the promotion of self-injury, suicide or eating disorders.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, our algorithmic systems do not categorize or recommend content concerning the sale of prohibited products or services, Harmful False Information or content that glorifies self-harm.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, we have specific prevalence testing and transparency reporting related to the sales of prohibited products and services, Harmful False Information and self-harm content.
Trusted Flaggers	Yes, we cooperate with trusted flaggers in relation to illegal content



Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	that harms Public Health. We also cooperate with trusted flaggers in relation to the sale of prohibited products or services, in particular the National Crime Agency (NCA - particularly in relation to the sale of drugs).
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various industry groups e.g. EUIF. To inform a responsible approach to mitigating these risks to Public Health, Snap regularly engages with experts from across the field of online safety, health, and wellbeing. Our Safety Advisory Board includes several such experts (including, for example, Dr. Michael Rich, pediatrician, founder and director of the Digital Wellness Lab & Clinic for Interactive Media and Internet Disorders, with affiliations at Boston Children's Hospital and Harvard Medical School). These experts have been consulted specifically on Snap's approach to wellness and mitigating risks related to mental and emotional duress, eating disorders, and other forms of self-harm. Snap has also been partnering with third party providers, NGOs and law enforcements at local level to identify appropriate measures to prevent the illegal sale of vape related products. Although this work has focused on the private services of Snapchat, which are out of scope of this Report, the work has also resulted in improvements to mitigations on the in-scope services on Snapchat.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on our terms, harms, moderation and enforcement practices (see the Annex), as well as how to and how to get help in our Safety Center.
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	We seek to protect all users from these harms. We offer Family Center; we make available robust reporting; and we provide guidance to parents on the web. Our parents site provides additional guidance for parents and carers on risks and support.¹⁶⁸

¹⁶⁸ <https://parents.snapchat.com>.



<u>Content Authenticity</u> Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	Yes, Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat’s inscope services. This would include Harmful False Information, such as deep fakes.
--	---

Conclusion

Snap recognises that, without adequate safeguards, digital content platforms can negatively impact Public Health. To address this, we have implemented a range of measures, notably proactive content moderation to detect and prevent harmful material from reaching a wide audience. Prevalence rates for Harmful False Information and Self-Harm or Suicide content have fallen substantially and now remain at extremely low levels, indicating our mitigations are effective. We continue to invest in preventing such content from spreading and in providing our community with online and in-app support resources.

We have concluded therefore that Snapchat’s in-scope services have reasonable, proportionate, and effective mitigation measures for the risk of Negative Effects on Public Health. There is no change in this conclusion from our 2024 Report.

4.4.2 Negative Effects on Gender-Based Violence

We strongly oppose content that promotes Gender-Based Violence. We recognise that without mitigations, a recipient of an online platform’s services could seek to publish this type of content. We take this issue very seriously and have put in place mitigation measures to address this risk.

Likelihood

There have been well documented examples of individuals, like Andrew Tate, using online platforms to promote content considered to be violent and misogynistic¹⁶⁹. Based on surveys carried out by the EU’s Fundamental Rights Agency in 2012 and 2019,¹⁷⁰ it estimates that between

¹⁶⁹ See example at: [url](#).
¹⁷⁰ European Union Agency for Fundamental Rights, survey results of 2012 ([url](#)) and 2019 ([url](#)).



4 and 7 % of women aged over 18 in the EU-27 experienced cyber-harassment in the previous 12 months, and between 1 and 3 % experienced cyberstalking. The study finds that prevalence has risen with greater use of the internet and social media, and is likely to increase further.

Snap does not track Gender-Based Violence as a specific, separate category in its [Transparency Reports](#) and as part of its prevalence testing. However, this type of content is captured within the scope of each of the following broader categories tracked by Snapchat including content relating to Hate, Harassment and Bullying, Harmful Information, Adult Sexual Content (in particular sextortion) and Violent and Dangerous Behaviour. We have observed low levels of prevalence of this type of content on Snap's in- scope services:

- In our 2024 Report, we were encouraged by data (second half of 2024) suggesting the likelihood of encountering such risks on Snapchat is within the lowest level. [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
- As at 30 April 2025, we are pleased to have subsequently observed a fall in prevalence across all of these content types, [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
- We received 132,611 reports of Hate Speech content in the EU in the second half of 2024, which led to enforcement action being taken against 12,625 unique content items and 55,891 accounts. This suggests users are able and willing to report content and accounts as needed on Snapchat (note these numbers relate to the whole of Snapchat, including private spaces that are out of scope of the DSA). We are encouraged by evidence that our approach has contributed to a low prevalence of Hate Speech content on the in- scope services of Snapchat.

From this, we continue to conclude that content promoting gender- based violence falls within the **Extremely low likelihood category** relative to other risks we have assessed.

Severity

The risks of Negative Effects on Gender-Based Violence stemming from online platforms is well-documented.

In our 2024 Report we highlighted one example, a study from the National Democratic Institute of International Affairs documents the ways in which online misogyny and harassment has chilling effects on the exercise of political rights—both online and offline—that disproportionately impact women.¹⁷¹ Based on this research, such harms can take various forms across platforms, including

¹⁷¹ National Democratic Institute, 'Tweets That Chill: Analyzing Online Violence Against Women in Politics', 2019, [url](#).



gender-based disinformation, Hate Speech, sexual harassment, non-consensual intimate imagery (NCII), sextortion, and human trafficking. We also highlighted that the European Parliament has proposed a legislative-initiative report calling for EU legislation to fight gender-based cyber-violence.¹⁷² A March 2021 EPRS European added value assessment (EAVA) on gender-based cyber-violence stresses the need for better data at EU and national levels.¹⁷³

More recently, ‘technology facilitated gender-based violence’ has been recognised as a serious global-scale problem that has increased with improvements and availability of generative AI technologies.¹⁷⁴ Teenage boys are believed to be among the most vulnerable to gangs using real and fake images to blackmail victims and there are reports that teachers in the UK have been asked to help spot signs their pupils are becoming victims of sextortion.¹⁷⁵

We take these risks seriously. The harms stemming from online Gender-Based Violence are not limited to the digital realm. Studies have suggested for example that gender-based disinformation not only presents a risk to women online, but also poses threats to national security.¹⁷⁶ Moreover, the U.S.-based National Center for Victims of Crime has noted that “the sharing of Nonconsensual Intimate Images can have devastating effects on victims and survivors. When images are shared online without consent, many victims feel ashamed, fearful, guilty, and the actions of sharing non consensual intimate images can lead to real-world implications and dangers.”¹⁷⁷ Given its harmful effects and high prevalence in society, Snap continues to qualify this risk as **serious**.

DSA Risk Factors

In accordance with Article 34(2), our risk assessment also addresses our recommender systems, content moderation systems, applicable terms and conditions, systems for the selection and presenting of advertisements and any of our data-related practices. These factors have been listed and explained in Section 3 on the [Risk Assessment Methodology](#) and applied throughout [Section 4](#). We also considered the risk factors in the context of Gender Based Violence. As many of these factors pose similar risks, and are mitigated in a horizontal manner (i.e. by measures that cut across all systemic risks), we did not include them in this section again but refer to our explanation in [Sections 4.1.2](#) on Hate Speech, [Section 4.1.6](#) on Adult Sexual Content, [Section 4.1.7](#) on Harassment and Bullying, [Section 4.1.9](#) on Violent and Dangerous Behaviour, and [Section 4.1.10](#) on the Harmful False Information.

¹⁷² European Parliament, ‘Combating gender-based cyber-violence’, December 2021, [url](#).

¹⁷³ European Parliament, ‘Combating gender-based violence: cyber violence - European added value assessment’, March 2021, [url](#).

¹⁷⁴ Exposing Technology-Facilitated Gender-Based Violence in an Era of Generative AI, 2023, [url](#).

¹⁷⁵ Teachers warned to be on lookout for victims of sextortion in UK schools, The Guardian - Kevin Rawlinson, April 29, 2024, [url](#).

¹⁷⁶ Brookings, ‘Gendered disinformation is a national security problem’, March 2021, [url](#).

¹⁷⁷ StopNCII.org, via [url](#).



Although the prevalence of content within the scope of this potential risk on Snapchat is considered to be at a lower level, due to the potential for serious harm to be caused by this content, Snap considers this to be a **Level 2 overall potential risk** for Snapchat's in-scope services. There is no change in this assessment from our 2024 Report.

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of	Several aspects of Snapchat's design and function reduce the risk of violative and harmful content being shared on the platform. • Snap makes it difficult for unvetted content to reach a large audience on Snapchat, and Snap proactively moderates



Mitigation Category	Applies to this risk?
their services, including their online interfaces.	Snapchat's in-scope services that provide an opportunity to reach a larger audience. • Snapchat has been designed to limit the prevalence of sexually suggestive content. • Snapchat's in-scope services have been adapted to include proactive moderation for illegal hate speech and violence, which includes Gender-Based Violence. • Snapchat is not an attractive platform for spreading misinformation, in particular because it is difficult to reach a broad audience and content is deleted by default. • Snap has made conscious design decisions to restrict the ability for content to go viral and limiting the remix functionality to specific content types and applying short retention to content. As part of our effort to take Gender-Based Violence into consideration across the design of our services, for example, we have further reviewed the design of our Lenses. In particular, Snap has removed tips to "Try with Friends" to some Lenses where there is a risk for bullying or harassment, including in relation to Gender-Based Violence. In risky cases, Snap won't encourage users to try a Lens with friends or Snap disables the Lens for being used with the rear camera (e.g. disabling this for the Pride Lens limits the ability to out someone else). These restrictions only apply to Lenses created by Snapchat. For Lenses submitted to Snapchat, we reject harmful Lenses to reduce the likelihood that they are distributed on Snapchat.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Snap takes a multifaceted approach to mitigating risks that may negatively impact Gender-Based Violence. Our policies include several prohibitions against content that may contribute to such risks, including sextortion, sexual harassment, NCII, Harmful False Information (which may include gender-based disinformation campaigns), Hate Speech, and human trafficking. In particular, Snap's Terms and Community Guidelines expressly prohibit Violent and Dangerous Behaviour, including Gender-Based Violence, and they are strictly enforced. Our terms and community guidelines also prohibit misinformation, Hate Speech, as well as Harassment and Bullying. We have a specific Harmful False or Deceptive Information explainer which explains our approach to enforcement and a Harassment and Bullying explainer with guidance on how we apply this policy. Arda Gerken, president of Offlimits, the Dutch hotline formerly known as EOKM, noted <i>"For young males who have experienced a sextortion incident — and the majority are males — they regularly tell us that when they share the situation with their parents, they feel relieved (...). We</i>



Mitigation Category	Applies to this risk?
	<i>advise them to report to hotlines and helplines; to report to the platforms; and to tell their parents, a friend or a trusted adult. They should not be going through this alone.</i>" Our in-app reporting tool allows users to directly report threats, violence or dangerous behaviours, including Gender-Based Violence, and anyone can submit a report through the Snapchat Support Site. We promptly enforce against accounts found to be sharing this content: • Snap removes such content for all users. • Accounts we discover engaging in prohibited activities will also be promptly disabled. • Where appropriate, accounts engaging in violation of these policies are reported to law enforcement as appropriate. Additionally, when we learn of content suggesting that there is an emergency situation involving imminent danger of death or serious bodily injury involving any person, we prioritise our response. We may also proactively escalate the report to law enforcement in certain circumstances, and we have processes for referring such content to the relevant authorities as appropriate. We note that escalations to law enforcement for this category are rare. The overwhelming majority of the cases that fall within this category are not illegal, but do violate our Community Guidelines. When hateful content is reported, our teams will remove any violating content and users who engage in repeated or egregious violations will have their account access locked. Lenses identified with Hate Speech were rejected when found during submission and disabled in Discover upon review if subsequently identified. As an additional measure, we encourage Snapchatters to block any users who make them feel unsafe or uncomfortable. Snap removes violating Hate Speech as soon as we become aware of it, and will promptly disable accounts dedicated to Hate Speech, hate symbols or groups, or the glorification of hate groups or members of a hate group. Specific rules¹⁷⁸ apply to our Lenses.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the	Yes, we have specific proactive and reactive moderation procedures to prevent and remove violative content, including content representing or encouraging gender based violence. As explained in the Content Moderation section in Section 5 of this Report, Snap deploys a range of automated content moderation (which include abusive language detection, other keyword-based detection, and machine-learning-based proactive detection models) to scan Stories and Spotlight submissions. Snapchatters can report violative content to us via

¹⁷⁸ <https://businesshelp.snapchat.com/s/article/Lens-Restrictions?language=en>



Mitigation Category	Applies to this risk?
expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	in-app reporting options and anyone can submit a report through the Snapchat Support Site. Our in-app reporting tool also allows users to directly report violative content. If Lenses are found to be violating our rules and promoting misogynist content, Snap takes enforcement action against such Lenses (for example, we have removed a few Lenses promoting Andrew Tate).
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Unlike many of our peers, Snapchat does not offer an open news feed where unvetted publishers or individuals have an opportunity to broadcast harmful and violative content, does not offer a broad ‘reshare’ functionality that would encourage virality, and does not allow user-generated content to be recommended to a wide audience without going through further review. Our algorithmic systems do not categorize or recommend Violent and Dangerous Behaviour, Harassment and Bullying, Hate Speech content or Adult Sexual Content, which would include Gender-Based Violence content, as further outlined in the previous sections. Similarly, we do not recommend content encouraging or engaging in misinformation, including potential misogynistic content, i.e. there is no ‘misinformation’ interest category. We take steps to prevent content with misleading or sensationalist headlines. In addition, we would note that sensitive content distribution is limited on both Spotlight and Discover: • In Spotlight, we limit the distribution of sensitive content based on the following rules: ◦ We do not recommend sensitive content to users under 18 by default. ◦ We do not recommend sensitive content to new users (i.e. users with less than 200 views in the past 28 days). ◦ For all other users, by default, we ensure the initial video watched in a session is not sensitive and after that we ensure that sensitive content is only shown sparingly (i.e. 1 in 7 videos). • In Discover, as in Spotlight, we limit the display of sensitive content for all users. We also do not show sensitive content to users under 18 by default and display of sensitive content can be disabled entirely in the Family Center.



Mitigation Category	Applies to this risk?
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	For example, Snap runs specific prevalence testing and transparency reporting which we use to help detect and manage Violent and Dangerous Behaviour, Adult Sexual Content (including sextortion), Harassment and Bullying, Hate as well as misinformation. Our Safety Advisor Board also has several anti-bullying experts which we call on for independent review and expertise. We also undertake intentional efforts to help all stakeholders understand these problems across the online community. As part of our Year Two Digital Well-Being study, we conducted a deeper drive into teens' and young adults' exposure to "sextortion" across platforms and services. The target countries were Australia, France, Germany, India, the UK, and the U.S, which includes three of the largest European countries, two of which are in the EU). We have continued this research during 2024 ("Year Three"), and also investigated teens' and young adults' attitudes and sentiments around reporting problematic content to platforms and services, authorities and others. More information on this research can be found in Section 6.6.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	We do not have a specific trusted flagger on Violent and Dangerous Behaviour in general but we do engage with several trusted flaggers on specific behaviors. For example, for certain threats, abusive and coercive behaviours we cooperate with Refuge. We cooperate with trusted flaggers in relation to illegal hate speech. Snap cooperates with trusted flaggers in relation to Non-Consensual Intimate Image Abuse (NCII), notably Stop Fisha in France.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online	We also undertake intentional efforts to help all stakeholders understand Gender-Based Violence across the online community. As part of our Year Two Digital Well-Being study, we conducted a deeper drive into teens' and young adults' exposure to "sextortion" across platforms and services.



Mitigation Category	Applies to this risk?
platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	The target countries were Australia, France, Germany, India, the UK, and the U.S, which includes three of the largest European countries, two of which are in the EU). We have continued this research during 2024 (“Year Three”), and also investigated teens’ and young adults’ attitudes and sentiments around reporting problematic content to platforms and services, authorities and others. More information on this research can be found in Section 6.6. Snap cooperates with other providers through various industry groups e.g. e.g. EUIF which has expanded its remit to also tackle the trafficking of human beings (which is often driven by sexual crimes or pornography). We are not working with other providers on Violent and Dangerous Behaviour specifically. However, Snap is a member of a number of EU trade associations to contribute to the policy debate to support the development of a proportionate regulatory framework to promote online safety. We cooperate with other providers through various groups in relation to Hate Speech. Snap remains a signatory of the EU Code of Conduct to counter Hate Speech online and has worked hard to ensure Snap meets the requirements (including with respect to recent revision of that Code).
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on our terms, harms, moderation and enforcement practices (see the Annex), as well as how to get help in our Safety Center and via in-app resources (Here For You and Safety Snapshot).. We make available robust reporting tools; and we provide guidance to parents on the web (see below).
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	We offer Family Center; we make available robust reporting; and we provide guidance to parents on the web. Our parents site provides additional guidance for parents and carers on risks and support.¹⁷⁹
Content Authenticity	Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for illegal or otherwise violating content and (ii) illegal or otherwise

¹⁷⁹ <https://parents.snapchat.com>.



Mitigation Category	Applies to this risk?
Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services. This includes Hate Speech and violent content.

Conclusion

Similar to the related Hate Speech category, Snap considers Gender-Based Violence to fall within our Level 2 overall potential risk [REDACTED]

[REDACTED] In response it has put in place a range of mitigation measures. This includes in particular our proactive content moderation which is designed to detect and prevent Hate Speech, including Gender-Based Violence related content from reaching a broad audience on Snapchat's in-scope services. Although we do not specifically document and report on Gender-Based Violence as a category, this type of content would primarily fall within the Hate Speech bucket, which has an extremely low prevalence (PVP) on Snapchat. We take this matter very seriously, and continue to invest in measures that prevent this type of content from reaching a broad audience on Snapchat, as well as provide our users with tools to report content to Snapchat and law enforcement, and support our community via online and in-app support tools, such as Here For You and our Safety Center resources. We believe our approach to these challenges reflects our commitment to responsibly mitigating harms that may negatively impact Gender-Based Violence. We are encouraged by evidence that our approach has contributed to a low prevalence of CSEAI and Hate Speech content on the in-scope services of Snapchat.

We have concluded therefore that Snapchat's in-scope services have reasonable, proportionate and effective mitigation measures for Gender-Based Violence. There is no change in this conclusion from our 2024 Report.



4.4.3 Negative Effects on Minors

We understand that without mitigations, online platforms could have a negative impact on minors. While studies conducted by credible health organizations like the American Psychological Association (APA) have concluded that digital platforms are not inherently harmful to young people, they suggest that adolescents' experiences online are affected by the features of the platforms they use.¹⁸⁰ Almost a year after APA issued its health advisory on social media use in adolescence, the APA has noted that society continues to wrestle with ways to maximize the benefits of these platforms while protecting youth from the potential harms associated with them.¹⁸¹ Other reports have suggested the impact social media might have on the mental health of Teens.¹⁸² The Committee on the Rights of the Child, the oversight body of the Convention on the Rights of the Child (CRC), affirms that the CRC also applies to online and digital media in its General comment no. 25 (2021) on children's rights in relation to the digital environment.¹⁸³ The Committee underscores that children and Teens need to be protected from risks of all forms of violence in the digital environment stating that "such risks include physical or mental violence, injury or abuse, neglect or maltreatment, exploitation and abuse, including sexual exploitation and abuse, child trafficking, gender-based violence, cyber aggression, cyberattacks and information warfare." The Committee of Ministers of the Council of Europe considers it as their key aim in the digital environment to offer an open, inclusive and secure internet to children and Teens, while ensuring the protection of their human rights in its Recommendation CM/Rec(2018)7.¹⁸⁴

The impact of online platforms on minors has also been a subject of consideration in regulatory reports from other authorities in Europe. For example, Ofcom's *Children Register of Risks*¹⁸⁵ explains how UK children are frequently exposed to online hate and bullying targeting protected characteristics such as gender, sexual orientation, race, religion, disability, and transgender identity. The same report shows that social media is the primary source of exposure - 78% of respondents encountered hateful or discriminatory content there, and 69% reported seeing misogynistic content.

Moreover, other research, including Wired's report from 2024,¹⁸⁶ confirms that social media continues to serve as a significant vector for illegal drug promotion and access among

¹⁸⁰ American Psychological Association, *Health advisory on social media use in adolescence*, [url](#).

¹⁸¹ American Psychological Association, *Potential risks of content, features, and functions: The science of how social media affects youth*, [url](#).

¹⁸² The Guardian, 'Facebook aware of Instagram's harmful effect on teenage girls, leak reveals', [url](#), 14 September 2021; E. Bozola e.a., 'The Use of Social Media in Children and Adolescents: Scoping Review on the Potential Risks', [url](#), August 2022.; A.M. Memom e.a., 'The role of online social networking on deliberate self-harm and suicidality in adolescents: A systematized review of literature', [url](#), October 2018.

¹⁸³ Committee on the Rights of the Child, *General comment no. 25 (2021) on children's rights in relation to the digital environment*, CRC/C/GC25, March 2021, [url](#).

¹⁸⁴ Council of Europe, *Recommendation CM/Rec(2018)7 of the Committee of Ministers to Member States on Guidelines to respect, protect and fulfill the rights of the child in the digital environment*, July 2018, [url](#).

¹⁸⁵ Ofcom, *Ofcom Register of Risks*, [url](#), (2024).

¹⁸⁶ Wired, 'Drug Dealers Have Moved on to Social Media', [url](#), (2024)



adolescents. The analysis found that 13% of posts examined contained drug-related content- predominantly on TikTok and Instagram - with cannabis and psychedelics being the most frequently featured substances. Alarmingly, 60% of young people aged 13–18 reported encountering such content, and 10% had purchased drugs via social media platforms. The study further showed that exposure to drug advertisements increased the likelihood of purchase by a factor of 17, underscoring the powerful behavioural influence of online promotion.

These indications underscore the significant responsibilities that platforms like Snapchat owe to vulnerable users such as teenagers. This is a risk we take seriously as Snap’s priority is protecting the safety and wellbeing of our users whilst ensuring they have a positive experience online. Privacy, safety and security are key values of the company and at the core of our value proposition to our users. We recognise that, *without mitigations*, Teens could become exposed to content that may impair their health, physical, mental and moral development - for example if steps are not taken to ensure Teens understand the design and functioning of Snap’s products, then they could unintentionally exploit the weaknesses and inexperience of Teens and may cause addictive behavior. This underscores the significant responsibilities that platforms like Snapchat owe to vulnerable users such as teenagers. This is a risk we take seriously as Snap’s priority is protecting the safety and wellbeing of our users whilst ensuring they have a positive experience online.

Likelihood

As explained in [Snapchat Community](#) as part of our Introduction to this Report, Snapchat is used by a wide demographic, with 18-24 years making up the highest percentage of users of Snapchat (89% in the EU by August 2025). Nevertheless, there is still a percentage of our users who are Teens (13-17) in the EU (no changes from our 2024 Report). Therefore we consider that children using Snapchat are just as likely to be exposed to the issues identified in this Report as other members of the Snapchat Community as follows:



Several studies have considered the likelihood of underage use of online platforms. For example, in 2022, Ofcom in the UK found that 60% of children aged 8 to 12 use social media with their own



profiles.¹⁸⁷ There were similar reports in Belgium.¹⁸⁸ Another example found the percentage of underage users on Snapchat to be relatively low (fewer than 4% of 0-11 year olds in the US in 2024) compared with other online platforms such as YouTube (28.6%), Netflix (17.2%) and Disney+ (15.6%).¹⁸⁹ We take a risk-based approach to age assurance at present that aligns with industry practice. We support further proportionate, reasonable and effective industry wide measures supported by device OS / app store account-level age assurance and are proactively working with the Commission and others to try to achieve an EU wide approach. We have highlighted our measures regarding minors under 13 in the specific mitigation section below (and provided more detailed information in [Section 5.8](#)).

All the risks we track on Snapchat have a low prevalence compared to the prevalence of these issues elsewhere online and offline. To aid our prioritization, our methodology seeks to assess the relative likelihood between the risks we track (even though all are low in absolute terms), we conclude that the relative likelihood of a risk of Negative Effects on Minors for Snapchat's in-scope services compared with other risks varies depending on the underlying concern. In general, the risks fall within the Extremely Low likelihood category. This includes Harassment and Bullying content and Adult Sexual Content, which, while now categorized at the lowest likelihood level, remain under ongoing supervision given their historical prevalence relative to other risks.

Severity

As with likelihood above, the severity of harm caused if a particular issue arises depends on the harm caused. However, we take the safety and wellbeing of the youngest members of our community very seriously and recognise that this group is particularly vulnerable and if a particular risk materializes, there is an increased risk that the severity of the harm they suffer is higher. [REDACTED]

¹⁸⁷ Children's Online User Ages - Quantitative Research Study, Ofcom, updated July 2022, [url](#).

¹⁸⁸ Réseaux sociaux, règles d'utilisation, intelligence artificielle : comment a évolué l'utilisation des écrans chez les jeunes ?, RTBF, May 31, 2024.

¹⁸⁹ Youth and Social Media, How US Kids and Teens Use Platforms From TikTok to Snapchat to YouTube, EMarketer, March 2024.





[REDACTED] [REDACTED]	[REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]
[REDACTED] [REDACTED] [REDACTED]	[REDACTED] [REDACTED] [REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]
[REDACTED]	[REDACTED] [REDACTED] [REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] is a risk they will suffer more harm.
[REDACTED]		
[REDACTED]	[REDACTED]	[REDACTED] [REDACTED] [REDACTED]
[REDACTED]	[REDACTED] [REDACTED] [REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]
[REDACTED]	[REDACTED]	[REDACTED] [REDACTED] [REDACTED]
[REDACTED]	[REDACTED]	[REDACTED] [REDACTED] [REDACTED]



		[Redacted]
[Redacted]	[Redacted]	[Redacted]
[Redacted]	[Redacted]	[Redacted]
[Redacted]	[Redacted]	[Redacted]
[Redacted]	[Redacted]	[Redacted]
[Redacted]	[Redacted]	[Redacted]
[Redacted]	[Redacted]	[Redacted]
[Redacted]	[Redacted]	[Redacted]
[Redacted]	[Redacted]	[Redacted]
[Redacted]	[Redacted]	[Redacted]
[Redacted]	[Redacted]	[Redacted]



In general, therefore, children and Teens suffer a risk of greater harm from the issues we have identified and we have chosen to place the severity of harm arising from an issue that negatively affects children in our ‘**severe harm**’ category.

DSA Risk Factors

In accordance with Section 3 ([Methodology](#)), we have also taken into account, in particular, whether and how the following service risk factors referenced in Article 34(2) of the DSA influence the risk of this harm on Snapchat’s in-scope services. As set out below, while the following risk factors could in principle influence this systemic risk, their overall impact is limited given Snap’s existing mitigations, which are described later in this chapter and in Chapter 5.

Service Risk Factor	How does it apply to Snapchat and this harm?
(a) the design of recommender systems and any other relevant algorithmic system;	Our recommender systems take user age into account to provide age-appropriate recommendations. Spotlight and Discover content for Teens is subject to stricter eligibility rules, moderation, and filtering to prevent sexually suggestive or otherwise age-inappropriate material from being shown. Detailed information regarding the mitigations relating to our recommender systems can be found in Section 5 (see the Algorithmic Systems subsection).
(b) content moderation systems;	Snap has implemented specific proactive and reactive moderation procedures for minors. These include age-gating of sensitive content, escalation guidance for moderators, removal of reported content, and Family Center settings that



	allow parents or guardians to adjust access to sensitive material and review their Teen's activity. We provide more information in Section 5 (see Moderation subsection)
(c) the applicable terms and conditions and enforcement;	Our Term of Service requires public content to be suitable for all Snapchatters (i.e. 13+). This helps ensure Snapchat is suitable for Teens and non-Teen users and reduces risks across all of the risk categories identified for Snapchat. We provide more information in Section 5 (see the Terms and Enforcement subsections).
(d) systems for selecting and presenting advertisements; and	Our advertising systems require agreement to advertising policies and guidance that prohibit adverts from displaying information that violates the law or causes certain harms to minors. We do not permit ads that target or appeal specifically to children, and ads unsuitable for minors are not served to Teen accounts. We check advertisers are complying with their obligations via our advertising review process. Our advertising systems use a mix of automation and human review to review adverts before they are published. We provide easy mechanisms for users to hide and report advertisements that violate our policies or the law. We monitor ad rejection, reporting and enforcement data to monitor the effectiveness of our approach. We provide more information in Section 5 (see the Advertising Systems subsection)
(e) our data related practices	We have strong data principles, practices and privacy, safety and security by design processes. Snap applies Privacy by Design and PASS reviews across products, with privacy explainers in our Privacy Center written for a 13+ audience. Teens and their parents can access "Privacy by Product" pages and videos that explain how data is used and what controls are available. We provide more information in Section 5 (see the Transparency sub-section) and Section 6 (see the Platform Principles Framework and Privacy and Safety by Design subsections).

We have also analysed whether and how the risk of Negative Effects on Minors is influenced by the following general factors:

General Risk Factor	How does it apply to Snapchat and this harm?
---------------------	--



Intentional manipulation, including inauthentic use or inauthentic use or automated exploitation of the service	There are two key ways in which we consider our systems could be manipulated: (1) Users could seek to share novel illegal and violating material that is not detected by our automated systems. We are constantly working to adjust our systems and policies to address this. We provide more information on our approach in Section 5 (in particular Moderation) and Section 6 (Ongoing Risk Detection and Management). (2) Users could abuse our content moderation processes and report non-violating content / accounts in bad faith. We have processes to combat misuse and more information can be found in Section 5 (Content Moderation and Enforcement).
Amplification and potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.	Snapchat's in-scope services have a number of features and design configurations that act to limit the amplification and potentially rapid and wide dissemination of content harmful to minors, in particular: 1. Snapchat's design limits virality for minors. Teen accounts cannot create Public Profiles by default, Spotlight and public content are moderated before wide distribution, and Teen Spotlight posts are anonymised and subject to stricter moderation and reply controls. Content flagged as sexually suggestive or otherwise sensitive is filtered from Teen feeds. We provide more information in Section 5 (see Snapchat Design / Function). 2. Snap has implemented specific proactive and reactive moderation procedures to prevent minors from accessing content that is harmful to them. We provide more information in Section 5 (see Content Moderation). 3. Content recommended to users on Spotlight and Discover, our video sharing platforms, is moderated using a combination of auto-moderation and human moderation, and is human moderated before being widely distributed. Lens and Ads are subject to review processes before submission. We provide more information in Section 5 (see Content Moderation).
Specific regional or linguistic aspects, including when specific to a Member State.	We recognise that our users may come from different Member States and content may be shared in different languages. To address this, • Snapchat is provided in multiple EU languages and



	our Terms are available in all EU languages. We provide more information in Section 5 (see Terms). • We deploy content moderation systems and moderators work across multiple EU languages. We provide more information in Section 5 (see Content Moderation). • Our Terms of Service requires public content to be suitable for all Snapchatters (i.e. 13+). This helps ensure Snapchat is suitable for Teens and non-Teen users and reduces risks across all of the risk categories identified for Snapchat.
--	---

Overall potential risk prioritization

Although the relative likelihood for the Negative Effects on Minors varies, Snapchat considers the risk of harm to fall within the severest category. Consequently, Snap considers this to be a **Level 1 overall potential risk**. There is no change in this assessment from our 2024 Report.

Snap's Mitigations

Highlights

The risk of Negative Effects on Minors falls within our highest risk prioritization level. To protect Teens on Snapchat, at a high level, we focus on three core things: 1) mitigating unwanted contact; 2) scanning for, detecting and blocking/removing content that violates our Terms (including our Community Guidelines) or the law; and 3) working with law enforcement to help bring criminals to justice. Snap has dedicated extensive resources to ensuring protections to safeguard the rights of Teens on the platform, greatly reducing the likelihood of rights infringement. These measures are set out in [Section 5.8](#) of this Report,



[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

- [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]



[REDACTED]

- 1 [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
- 2 [REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

- 1 [REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

- 1 [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

- 1 [REDACTED]
[REDACTED]
- 2 [REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

- 1 [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]



[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



□ □ □ □ □

██████████

299



[Redacted]

- [Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]
- [Redacted]

- [Redacted]
[Redacted]
- [Redacted]
[Redacted]
[Redacted]
[Redacted]
 - [Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

- [Redacted]
- [Redacted]
[Redacted]
 - [Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

A more detailed run through of our mitigations to protect Teens on Snapchat is set out in [Section 5.8](#) Protection of Minors). Taken together, these mitigations contribute to a safe and responsible environment for young Snapchatters.

Specific Mitigations

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat’s in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive. As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a **link in the left hand column to a full summary** of the specific mitigation in [Section 5](#) of this Report which explains in more detail how each mitigation operates to reduce the risk.



Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Yes, Snapchat's in-scope services have been adapted to include proactive safety by design features and content moderation for Teens. For example: • We have created a different product experience for Teens and adults. For example, we don't show sexually suggestive content to Teens. • All content on Discover has to be appropriate for 13+. • Regulated goods don't appear in ads to Teens. • Snap Map is designed to mitigate particular risks to Teens. For example, location sharing is off by default.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Yes. For example, our Terms require that all content is appropriate for 13+, we require all users on our platform to be over the age of 13, and we strictly enforce our terms. If we discover that a user is under the age of 13 we will remove their account.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, specific proactive and reactive moderation procedures to age gate and ensure age-appropriate content (for example restricting Teens access to suggestive content), adjust content settings as designated in Family Center and remove reported content from view.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Yes, our algorithmic systems take user age into account to provide age appropriate recommendations.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems. We have also launched changes to Snapchat's in-scope services so they no longer display advertisements based on profiling for our under 18 accounts in the EU.



Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	Yes, we have specific prevalence testing and transparency reporting for violations, including for example in relation to CSEAI.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21.	Yes, we cooperate with numerous trusted flaggers for child safety who are able to flag other CSEAI or other illegal and violating activities involving Teens.
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Yes, we cooperate with other providers through various industry groups e.g. Technology Coalition, WeProtect Global Alliance, EUIF, Alliance and CIPL to better protect minors online. We continue to review our measures against the recommendations in the recently released Article 28 Guidelines to assess if further reasonable, proportionate and effective measures are needed. See Protection of Minors for more information.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on our terms, harms, moderation and enforcement practices (see the Annex), as well as how to and how to get help in our Safety Center. All information on our Privacy and Safety Center or our Support Center is drafted for 13+. For example, Privacy By Product - Privacy Features Snapchat Privacy provides Teens with ample opportunity to understand the Snapchat features. We also provide Family Center as a resource to Teens and their parents or trusted adults.
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	Yes, Snapchat's in-scope services have been adapted to include proactive safety by design features and content moderation for Teens. We make available robust reporting and enforcement of our terms. Our Family Center - Parental Control For Teens Snapchat Safety provides Teens and their parents or trusted adults a suite of resources and guidance. Our parents site provides additional



	guidance for parents and carers on risks and support. ¹⁹⁰
<u>Content Authenticity</u> Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	Yes, Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for creating illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's in-scope services. We display an icon in some Lenses that manipulate an image of a Snapchat to make them look younger.

Conclusion

Although the prevalence of public content that may have negative effects on children on Snap's in-scope services is generally very low, we recognize that Teens are at risk of greater harm if exposed and we take the safety and wellbeing of our community, particularly its youngest members, very seriously. As such, we have assessed this risk to be in our higher risk prioritization category, Level 1, relative to other risk categories.

In response, Snap has put in place a range of mitigation measures. This includes general platform safeguards such as our Teen friendly terms and support pages, our moderation and enforcement processes, our parental tools – [Family Center](#), in-app reporting, and Teen specific content moderation and restrictions. Plus, additional safeguards have been put in place to help Teens understand and recognize Lenses and ensure that advertisers and advertisements on our platform comply with our requirements.

We have concluded therefore that Snapchat's in-scope services have reasonable, proportionate, and effective mitigation measures for the risk of negative effects on Teens and children under 18. We continue to review our measures against the recommendations in the recently released Article 28 Guidelines to assess if further reasonable, proportionate and effective measures are needed. We strongly advocate for further guidelines on the role of 'gatekeeper' services (such as device operating systems, app stores and web browsers).

¹⁹⁰ <https://parents.snapchat.com>.



4.4.4 Serious Negative Consequences on Physical and Mental Well-Being

Snapchat focuses on helping Snapchatters communicate with their close friends in an environment that prioritizes their safety and privacy. It is purposely designed differently from traditional social media. It doesn't open to a public news feed powered by an algorithm with likes and comments. Instead, as outlined earlier in this Report, Snapchat opens to a camera and has five tabs: Camera, Chat, Map, Stories, and Spotlight. Additionally, conversations on Snapchat delete by default to reflect real-life conversations. Before social media, our fun, spontaneous, and silly interactions with friends only lived on in our memories. Snapchat is designed to mirror that dynamic, to help people feel comfortable expressing themselves without feeling pressure or judgment. We will discuss these risks and our mitigations in more detail below.

Likelihood

All of the risks we track on Snapchat have a low prevalence compared to the prevalence of these issues elsewhere online and offline. To aid our prioritization, our methodology seeks to assess the relative likelihood between the risks we track (even though all are low in absolute terms).

Snap assesses that serious negative consequences on physical and mental wellbeing are high in likelihood in the absence of appropriate mitigations. Without mitigations, users of digital platforms may be exposed to content affecting their mental health, contributing to body dissatisfaction and low self-esteem. They may also be exposed to content inciting physically harmful activities, such as dangerous pranks or challenges.

On Snapchat, data related to relevant policy enforcements suggests a low prevalence of content associated with harm to physical well-being on Snapchat. For example:

- In our 2023 Report, we measured the [prevalence](#) (PVP) of Self-Harm content (including the promotion or glorification of unhealthy behaviors) and the prevalence of content promoting dangerous activities to be extremely low. In 2024, we have subsequently observed a significant decrease in the prevalence of all of these categories of content, [REDACTED]
- As of April 30, 2025, the prevalence of Self-Harm content has decreased further [REDACTED]
- [REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

[REDACTED]



We also monitor community support requests, which are a good indication of the well-being of our Community. We continue to receive very few user complaints from EU users related to dangerous categories that could be considered to have a serious Negative Effect on Physical and Mental Well-Being, [REDACTED]

-
- | Government | Percentage |
|---------------------|------------|
| Current government | 85% |
| Previous government | 15% |





Over 90% of our community says they feel happy, connected, and comfortable while using Snapchat.¹⁹¹ Research from University of Chicago's NORC¹⁹² shows that 2 in 3 say messaging with family and close friends makes them extremely or very happy. On the other hand, a majority of teens and young adults feel overwhelmed at the way traditional social media makes them feel pressured to post content that will get lots of likes and comments, or will make them look good to others. Perhaps most importantly, according to the NORC data, respondents who use Snapchat report higher satisfaction with the quality of friendships and relationships with family than non-Snapchatters.

Snapchat is also aware of a study from the Netherlands,¹⁹³ that, according to researchers, is the first quantitative study that compares online platforms on three key factors: impact on well-being, self-esteem and friendship closeness. The researchers conducted an extensive 100-day daily diary study among 479 adolescents (14-17 years). With respect to Snapchat specifically, the research concludes that Snapchat is the only platform that positively impacts well-being and Snapchat also has a strong positive effect on friendships and no net negative effect on self-esteem. It notes in particular:

- Snapchat positively affected friendship closeness and well-being but had no significant impact on self-esteem. Using WhatsApp had a notably strong effect on friendship closeness but no significant effect on well-being and self-esteem.

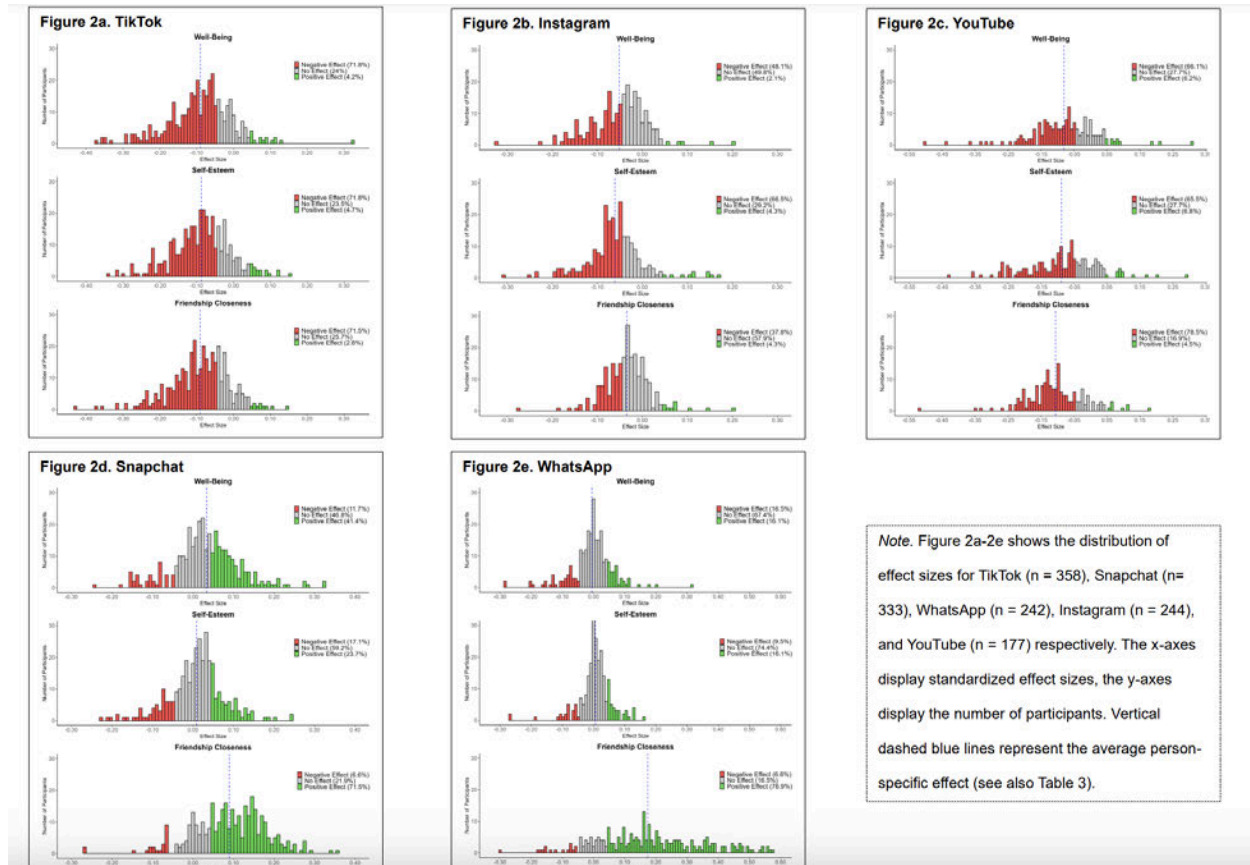
¹⁹¹ 2022 Alter Agents study commissioned by Snap Inc. [url](#).

¹⁹² [https://www.norc.org/about/who-we-are.html](https://www.norc.uchicago.edu/about/who-we-are.html).

¹⁹³ Social Media Use Leads to Negative Mental Health Outcomes for Most Adolescents, Amber van der Wal, Ine Beyens, Loes H. C. Janssen, and Patti M. Valkenburg, 2024, pre-print, [url](#)



- The majority of adolescents (60%) experienced unity in negative effects of social media, suggesting that social media use is a contributor to mental health issues. Moreover, 13.6% of adolescents experienced duality in effects, indicating that social media use simultaneously harms and benefits different dimensions of their mental health.
- The positive and null effects associated with Snapchat and WhatsApp indicate that we should avoid a blanket condemnation of all social media platforms.
- Snapchat scores a 41.4% positive effect on well-being, 23.7% on self-esteem and 71.5% on friendship closeness, as shown in the figures and table below:



**Table 3***Overall Within-Person Effects of the Top Five Social Media Platforms on Mental Health*

	Well-Being	Self-Esteem	Friendship Closeness
TikTok	-.09***	-.08***	-.09***
Instagram	-.05***	-.06***	-.04**
YouTube	-.08***	-.09***	-.11***
Snapchat	.03***	.01	.09***
WhatsApp	-.00	.01	.17***

Note. Cells marked in red indicate significantly negative overall within-person effects, cells marked in green indicate significantly positive overall within-person effects.

* $< .05$. ** $< .01$. *** $< .001$.

On Safer Internet Day, 6 February 2023, we launched our inaugural [Digital Well-Being Index \(DWBI\)](#), a measure of Generation Z's online psychological well-being across different online services and platforms. To gain insight into how teens and young adults are faring online – across all platforms and devices, not just Snapchat – and to help inform our [Family Center](#) and the broader online ecosystem, we polled more than 9,000 people across three age demographics in six countries. Not surprisingly, the research showed that social media plays a major role in Gen Z's digital well-being, with more than three-quarters (78%) of respondents saying social media had a positive influence on the quality of their lives. We have repeated this research in 2024. The [third Digital Well-Being Index](#) for the six geographies covered stands at 63, one percentage point higher than the previous two years, and still a somewhat average reading on a scale of 0 to 100 – neither particularly favorable, nor especially worrisome. 10% of participants described their experience online as extremely positive, 44% of them described their overall online experience as being overall positive, 40% rated their experience as mixed while only 6% reported to have consistently encountered negative experiences and outcomes. Further information can be found in Section 6.6 (DWBI Initiative).

Accordingly, while Snap assesses these risks across digital platforms to be high in the absence of safeguards and mitigations, we are encouraged by research and data indicating that our approach to mitigating these risks is effective at reducing the likelihood of such negative impacts on physical and emotional wellbeing on Snapchat. We continue to assess there to be a **low likelihood** of encountering this content on Snapchat.



Severity

Snap is deeply committed to the safety and well-being of its community; we recognize that negative experiences on digital platforms can have a detrimental impact on mental health and wellness, and therefore work intentionally to ensure that our policies and practices reflect our commitment to fostering joyful, creative, and expressive products that contribute positively to the lives of our users.

Researchers at Harvard's Chan School of Public Health note that the current body of scientific literature "characterizes the link between social media use and health in at least two ways. One body of literature considers social media use as a normal social behavior with positive or negative effects on health-related outcomes, while the other focuses on problematic use and associated effects."¹⁹⁴ While such studies have shown both positive and negative potential impacts on users' health and wellbeing, the risks of negative impacts are well documented (particularly for teenagers of a particular age).¹⁹⁵ Research on screen time shows that an increase in screen time reduces physical activity among children and young people substantially.¹⁹⁶ The excessive use of screen-based media among young individuals is an evident and growing public health issue that requires significant focus and ongoing research to stay abreast of the rapid and profound transformations in screen media technology and its usage patterns. Additionally, addiction to social media is a major contributing factor to the rapid increase in several mental health issues¹⁹⁷

We note also that in May 2023 the (US) Surgeon General's Advisory issued a warning regarding social media and youth mental health, in particular with regards to exposure to hate-based content and suicide/self-harm-related material¹⁹⁸ and that studies continue into the impact of digital screen media on physical and mental wellbeing.¹⁹⁹ On the other hand, other research has concluded that there is no evidence that screen time harms children's thinking abilities or wellbeing²⁰⁰ and others have presented strong criticism of existing research²⁰¹. Many have concluded that there is not yet enough evidence.²⁰²

¹⁹⁴ Bekalu e.a., 'Association of Social Media Use With Social Well-Being, Positive Mental Health, and Self-Rated Health: Disentangling Routine Use From Emotional Connection to Use', November 2019, [url](#).

¹⁹⁵ The New York Times, 'Does Social Media Make Teens Unhappy? It May Depend on Their Age', March 2022, [url](#).

¹⁹⁶ European Research Council, 'Immediate and long-term health risks of excessive screen-based media use', January 2023, [url](#).

¹⁹⁷ S. Vatsa e.a., 'Social Media and its Effects on Mental Health, March 2020, European Scientific Journal, [url](#).

¹⁹⁸ <https://www.hhs.gov/sites/default/files/sq-youth-mental-health-social-media-advisory.pdf>.

¹⁹⁹ See for example: [Impact of digital screen media activity on functional brain organization in late childhood: Evidence from the ABCD study](#), Jack Miller a, Kathryn L. Mills b, Matti Vuorre a d, Amy Orben c, Andrew K. Przybylski, 2023.

²⁰⁰ Children's brains 'not harmed by more screen time', Kat Lay, The Times, 18 Nov 2023.

²⁰¹ See for example: <https://www.peteetchells.com>.

²⁰² [Science doesn't yet support broad restrictions on teens' access to social media, experts say](#) LA Times – Corinne Purtill – December 13, 2023.



DSA Risk Factors

Overall potential risk prioritization

██████████ If we would follow our matrix, we would qualify the overall potential of this risk as Level 3, however, given the importance of this issue, especially in relation to younger users, we have decided to deviate from the matrix, and marked the potential risk as a **Level 1 priority**. There is no change in this assessment from our 2024 Report.

[illegible]

Snap enables user photo and video sharing and other user interaction across Snapchat's in scope services. This functionality has many positive benefits for users, including in particular

312



facilitating engagement with friends and family and Snap has put in place significant measures to substantially diminish the likelihood and impact of Snapchat's in-scope services having an impact on the Physical and Mental Well-Being of users.

In the table below we indicate the specific measures we have taken to mitigate this risk in respect of Snapchat's in-scope services, using the defined list of mitigations set out in Article 35 of the DSA. The primary purpose of the below table is to indicate whether each specific mitigation category applies to this risk and the descriptions are illustrative rather than exhaustive. As many of our mitigations apply to all of the risks assessed in this Report, to reduce duplication in this Report, each row in the tables provides a [link in the left hand column to a full summary](#) of the specific mitigation in Section 5 of this Report which explains in more detail how each mitigation operates to reduce the risk.

Mitigation Category	Applies to this risk?
Snapchat Design and Function Adapting the design, features or functioning of their services, including their online interfaces.	Several aspects of Snapchat's design and function reduce this risk. • Starting with the aforementioned decision to open to the camera and not a news feed. This encourages self expression, communication, and exploration through our AR Lenses. • Snapchat and third parties have created Lenses centered on movement, fitness, yoga poses, breathing activities. In addition to this, there are several partnered lenses that prompt Snapchatters to talk about wellness, mental health and their experiences. • Our user generated content feature, Spotlight, has both creator and viewer protections in place. On Spotlight, we put in place protections for both creators and views: • Creator protections ○ Users can post to Spotlight and choose to disable comments. ○ If comments are not disabled, Spotlight comments are auto-moderated for abusive language before they are viewed by the creator and all comments can be reported to human moderation. This protects the creator from seeing harmful comments. ○ Teens are protected on Spotlight by not having their usernames displayed. ○ We limit the recommendation of content from younger users to older users on Spotlight. This is to protect Teens from being contacted by older users.



	○ We provide users the ability to post content to Spotlight anonymously. ○ Creators can choose to approve comments on their Spotlight Stories prior to publication. ○ We do not show view-counts on Spotlight with less than a certain number of views. This is to prevent embarrassment over low view numbers. ○ We aim to distribute content created by Teens to Teens. This is to prevent Teens from building a following that is not their own age. ○ Creators are in control of adding hashtags / topics to their videos. This provides creators some control over how their content is categorized. ● Content on Spotlight does not auto-advance. ● We do not have public “favorites”, i.e. a user’s likes and interests are not public. ● Viewers can “hide” either content or a creator. Subsequently, the user will have a lower likelihood of seeing content of such nature or content from the creator that has been “hidden”. ● We survey a subset of our users quarterly to understand whether they find their time spent on our experience entertainment and satisfactory. We use this to track whether our product changes are improving viewers' overall perception of the app. ● We provide a diversity of perspectives. We have multiple programs to foster a more diverse content community and surface different perspectives (e.g. Black Creator Accelerator program). ● We ensure there is always a large mix of content from creators from viewers’ home country and content in the language in which they have set their device. ● We add diversity to every viewer’s feed in terms of the account they see, and the categories of content we surface to them. This prevents users from entering an echo chamber or filter bubble of seeing the same content repeatedly. We use machine learning to understand content categories and diversify it. For example, we allow users with access to a Public Story to turn off all story reply messages so they don’t see messages from users who reply to their Stories. We also give users control over Story Replies and filter out words they don’t want to see. Users can input words that they don’t want to receive in the Story replies from their subscribers. If a Story reply contains an inputted word, the user does not receive the story reply (and any other story replies) from the
--	--



	sender. Additionally, we allow creators to block repliers or report them.
Terms and Enforcement Adapting their terms and conditions and their enforcement.	Snap's Community Guidelines prohibit a range of behaviors and content that may negatively impact wellbeing, including Harassment and Bullying; content or Lenses that glorify unhealthy behaviors or promote unrealistic beauty standards; violent or disturbing content, or content that promotes dangerous activities; and content that promotes suicide or self-harm. Snap enforces such Community Guidelines along with its Terms.
Moderation Adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation.	Yes, we have specific proactive and reactive moderation procedures to prevent and remove prevent Harassment and Bullying, content or Lenses that glorify unhealthy behaviors or promote unrealistic beauty standards, violent or disturbing content, or content that promotes dangerous activities, and content that promotes suicide or self-harm. As explained in the Content Moderation section in Section 5 of this Report, Snap deploys a range of automated content moderation (which include abusive language detection, other keyword-based detection, and machine-learning-based proactive detection models) to scan Stories and Spotlight submissions. Snapchatters can report violative content to us via in-app reporting options and anyone can submit a report through the Snapchat Support Site. Our in-app reporting tool also allows users to directly report violative content. If, for example, Lenses are found to be violating our rules and glorifying unhealthy behaviours or promoting unrealistic beauty standards, Snap takes enforcement action against such Lenses.
Algorithmic Systems Testing and adapting their algorithmic systems, including their recommender systems.	Unlike many of our peers, Snapchat does not offer an open news feed where unvetted publishers or individuals have an opportunity to broadcast harmful and violative content, does not offer a broad 'reshare' functionality that would encourage virality, and does not allow user-generated content to be recommended to a wide audience without going through further review. Our algorithmic systems do not categorize or recommend content that our Community Guidelines prohibit. Our Content Guidelines for Recommendation Eligibility further describe how sensitive and disturbing content is demoted for distribution on Spotlight and Discover. For example, glorification of violence is not suggested content to users on Spotlight or Discover



	and any discussion on self-harm, including eating disorders is demoted to users based on their age, location, or personal preferences. In addition, we would note that sensitive content distribution is limited on both Spotlight and Discover: • In Spotlight, we limit the distribution of sensitive content based on the following rules: ◦ We do not recommend sensitive content to users under 18 by default. ◦ We do not recommend sensitive content to new users [REDACTED] ◦ For all other users, by default, we ensure the initial video watched in a session is not sensitive and after that we ensure that sensitive content is only shown sparingly (i.e. 1 in 7 videos). • In Discover, as in Spotlight, we limit the display of sensitive content for all users. We also do not show sensitive content to users under 18 by default and display of sensitive content can be disabled entirely in the Family Center.
Advertising Systems Adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide.	Yes, other mitigations listed here also apply to our Advertising Systems. For example, ads for diet and fitness products or services must not demean the user, or shame anyone on the basis of body shape or size.
Ongoing Risk Detection and Management Reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk.	For example Snap runs specific prevalence testing and transparency reporting which we use to help detect and manage for Harassment and Bullying and Self-Harm and Suicide and other prohibited content on Snapchat that may impact users mental wellbeing. Our Safety Advisor Board also has several anti-bullying experts which we call on for independent review and expertise.
Trusted Flaggers Initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions	Yes, we cooperate with trusted flaggers in relation to illegal hate speech and child safety.



of out-of-court dispute settlement bodies pursuant to Article 21.	
Codes and Crisis Protocols Initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively.	Snap cooperates with other providers through various industry groups e.g. EUIF. In 2017 Snap joined FSM and has signed the FSM Code of Conduct which aims to protect users from content offered on digital services that could endanger or impair their development.
Transparency Taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information.	Yes, we provide guidance on our terms, harms, moderation and enforcement practices (see the Annex), as well as how to get help in our Safety Center and via in-app resources (Here For You and Safety Snapshot). We make available robust reporting tools; and we provide guidance to parents on the web (see below).
Protection of Minors Taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate.	We offer Family Center; we make available robust reporting; and we provide guidance to parents on the web. Our new parents site provides additional guidance for parents and carers on risks and support.²⁰⁴ In addition, as we have explained in the mitigations section of this Report, in particular Section 5.8 (Protection of Minors), teenagers, parents and other responsible adults are able to set time limits for their teenagers, amongst other controls, via the device operating system's family tools (e.g. Google Family Link, Apple device parental controls and Family Sharing controls and Microsoft Family Safety). Mobile devices now also commonly provide default settings for late night usage, such as bedtime modes that turn off device and app notifications and turn the screen the black and white to encourage sleep.
Content Authenticity Ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or	Snap has taken steps to mitigate the risk that (i) its generative AI tools are used for illegal or otherwise violating content and (ii) illegal or otherwise violating content created using generative AI tools on any online platform are disseminated on Snapchat's inscope services. We are displaying an icon in some Lenses that

²⁰⁴ <https://parents.snapchat.com>.



video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.	manipulate an image of a Snapchat to make them look younger.
--	--

Conclusion

Given the heightened potential for negative consequences on Physical and Mental Well-Being inherent to online platforms, specifically social media, despite the prevalence on Snapchat being low, we consider the risk prioritization to be Level 1.

In response, Snap has made deliberate design and policy decisions to reduce the potential for harm on Snapchat. Snap has implemented numerous protections for both creators and viewers of Spotlight content and undertaken considerable efforts to understand users’ wellbeing on Snapchat and other platforms.

We have concluded therefore that Snapchat’s in-scope services have reasonable, proportionate, and effective mitigation measures for the risk of Negative Effects on Physical and Mental Well-Being. There is no change in this conclusion from our 2024 Report.



5. Specific Mitigations

Article 42(4)(b) of the Digital Services Act requires providers of Very Large Online Platforms to report on the specific mitigation measures that they have put in place pursuant to Article 35(1) of the DSA. Article 35(1) of the Digital Services Act requires providers of Very Large Online Platforms to put in place reasonable, proportionate and effective mitigation measures, tailored to the specific systemic risks identified pursuant to Article 34 of the DSA, with particular consideration to the impacts of such measures on fundamental rights, including where applicable the defined categories of measures set out in Article 35(1)(a)-(k).

In [Section 4](#) of this Report above, we reported on our: (i) assessment of the specific systemic risks applicable to Snapchat’s in-scope services; (ii) summary of the mitigation measures that Snap has in place tailored to those risks and (iii) conclusion as to whether those mitigation measures are reasonable, proportionate and effective. In this Section 5, we have provided details of the specific mitigations that Snap has put in place, as are summarized in Section 4, to comply with our obligation under Article 42(4)(b).

5.1 Snapchat Design and Function

5.1.1 Introduction

From day one, Snap has made conscious design decisions to mitigate systemic risks from occurring on its platform, including privacy and safety by design decisions such as shorter retention periods, default Story visibility to just friends, not promoting likes on a user’s Story, not having public friend lists, and maintaining proactive content policies. Snap has developed a Design Policy; prohibits the use of dark patterns & misleading nudge techniques; has created an Online Design Interface Process to guide Snap Engineering through the process to develop, test, adapt, and deploy changes to the interface that follow Snap’s Design Policy; and has established clear roles and responsibilities for review and approval of changes to the system as part of Snap’s overall Safety and Privacy by Design Process. Snap understands that these measures alone are not sufficient to address all harms, therefore Snap has implemented additional measures as further outlined in the remainder of this Report.

5.1.2 Oversight and Administration

Roles and Responsibilities

Roles	Responsibilities
Legal Teams	• Outlines roles and responsibilities for review of changes to the design, features, or functioning of Snapchat services. • Reviews proposed feature changes prior to launch for potential dark patterns on online interfaces.



Design, Product, and Engineering Teams	• Responsible for adhering to Snap’s Dark Patterns Design Policy when developing, testing, adapting, and deploying changes to Snapchat interface and functionality.
DSA Cross-Functional Governance Team	• Reviews Snap’s Dark Patterns Design Policy on an annual basis to ensure the incorporation of relevant regulatory requirements.
DSA Compliance Officer	• Ensures compliance of Snapchat’s online interfaces and organization with DSA requirements.

5.1.3 Adaptations and Mitigations

As a result of our privacy and safety by design approach (described in Section 6.3 of this Report), Snapchat was designed from the outset with core features and functionalities that mitigate the risks described in Snap’s Risk Assessment Report.

We have made these key foundational design decisions from day one. We hear from Snapchatters about the benefits of these choices all the time, as well as consulting with expert and teen stakeholders (such as those forming part of our Safety Advisory Board and Teen Council) and we believe that these foundational design decisions directly influence those results.

Although not all of the features listed below are in scope of the risk assessment, we have incorporated a summary of holistic mitigations that we have put in place to demonstrate our privacy and safety by design approach below. This is not intended to be a comprehensive list.

Friending, Chat and Private Stories

Friends

First, by default users need to accept bi-directional friend requests or already have each other in their contact book to start communicating directly with each other. This design decision adds friction and prevents users from communicating with each other prior to accepting a friend request or being in one’s contact book. This is also an important mitigation to prevent strangers from contacting users on Snapchat.

Private friend lists

Second, once users have accepted friend requests, the friend lists remain private. Snapchat does not disclose the friend lists of users to other users, nor do we expose the total number of friends that a user has. This protects the privacy of the user and their friends. On most other platforms friend lists are public by default or there is an option to share them publicly. These types of features create the ability for strangers to contact vulnerable groups (e.g. younger users) and infiltrate friend groups by going from friend list to friend list.



Open to the Camera not a feed

Third, Snapchat opens to the Camera and invites people to express themselves. At the surface, this may sound like a small design decision, but it directly impacts the user behavior on the platform. Instead of inviting users to scroll a feed of content, the invitation to users is to express themselves, live in the moment and share a moment with their close friends.

Stories are by default set to be viewable by friends, not the public

Fourth, once users decide to share a Snap via My Story, by default only friends can view it. Snapchatters can choose to share to everyone, only to friends, or to a customized few. This emphasis on sharing with friends and giving users controls over who can view their content are in line with how Snap takes into account privacy and safety when designing its features.

No focus on public vanity metrics

Fifth, once a user posts to their Story, we don't show vanity metrics, such as likes on that Story content. The goal is not to create a popularity contest around who has the most friends or likes.²⁰⁵ The design choice is to provide all users with a more authentic form to express themselves.

As a result of our privacy and safety by design approach, each of Snapchat's in-scope services has been designed with features and functionalities that mitigate the risks described in [Section 4](#) above.

Spotlight and Discover

Spotlight

Spotlight offers creators at all stages of their career a variety of opportunities and tools to help them grow their audiences, build sustainable businesses and make content creation a full-time career. The content shown in Spotlight is personalized to provide viewers with a more relevant experience, that 'spotlights' the best content on Snapchat. We have made following design decisions to protect our creators and users:

- Creator protections
 - Users can post to Spotlight and choose to manually approve or deny comments.
 - Spotlight comments are auto-moderated for abusive language before viewed by the creator and all comments can be reported to human moderation. This protects the creator from seeing harmful comments.
 - Adults cannot comment on Teen's Stories on Snapchat.
 - Teens are protected on Spotlight by not having their usernames displayed on their Spotlight posts.

²⁰⁵ See also our More Snapchat campaign <https://www.moresnapchat.com/>



- We restrict Teens' ability to reach a large audience on Spotlight to prevent older users from seeing content from younger users. This is to protect Teens from being contacted by older users.
- We aim to distribute content created by minors to minors. This is to prevent minors from building a following that is not their own age.
- Creators are in control of adding hashtags / topics to their videos. This gives creators control over how their content is categorized.
- Viewer protections
 - Content on Spotlight does not auto-advance.
 - We do not have public “favorites”, i.e. a user’s likes and interests are not public.
 - Viewers can “hide” either content or a creator. Subsequently, the user will not see more content of such nature or content from the creator that has been “hidden”.

Discover

Discover is dedicated to Creator Stories, which includes Media Partner content and some user generated content from popular users (“Creator Content”). The Creator Content that appears on Discover includes Public Stories from Snap Stars and other users who meet a follower count threshold. Similar to Spotlight, we made following design decisions to protect our creators and users:

- Creator protections
 - Viewer comments are not typically available on Discover content. Where comments are enabled, they are subject to auto-moderation. Creators and other users can report comments, which leads to human review. They can also block commenters which will prevent them from ever seeing the blocked users’ comments again on any content.
 - We do not show “views” on Discover Stories. This protects creators from feeling embarrassed or being subject to ridicule due to low number of views.
 - Content published by creators has a limited publication duration (which may be changed by creators with a Snapchat+ subscription). This protects creators by ensuring their content is not available forever.
 - Creators are free to re-publish new and saved stories at any time, provided it does not violate the law or our Terms.
- Viewer protections
 - Content on Discover does not auto-advance.
 - We do not have public “favorites”, i.e. a user’s likes and interests are not public.

Public Profile

Users with a Public Profile can post Public Stories that are publicly viewable for all Snapchatters. Additionally, Snapchatters can showcase their Public Stories and Spotlights on their profile.



Snapchatters can watch Public Stories on Discover, Snap Map, Spotlight. Users can also [Follow](#) a Public Profile. Unlike friend requests to non-Public Profile owners, Public Profile owners will not receive a notification for new followers. We have made the following design decisions to protect users with Public Profiles:

- Users can easily delete all of their public content. We allow users to delete all of their public content in a single tap. We delete any and all content they added to their Public Profile and that is publicly viewable. Our public options [are in fact options](#). If Snapchatters are not or no longer interested in being a creator and showcasing content publicly, they can simply choose to not add to their Public Profile, post to their Public Story, or share to Spotlight and to the Snap Map.
- We give users control over content that is publicly viewable by allowing users to hide or show their Spotlights on their Public Profile both at the time of submission and after submission.
- Public Profile users can turn off remixes. We allow users to decide whether their public content can be remixed by other users.
- We educate users on their public options and attribution controls. When users first tap on their public profile, public story, and spotlight/snap map posting, we show them educational modals that educate them about the public option.
- To ensure that users are aware when they become friends with another user so that they can control what data that user has access to, we send notifications to the user when they become friends with another user (bi-directional add has occurred).
- Only users with an older teenage account (16-17) or an adult account (18+) can have a Public Profile and Public Story. Viewers cannot distinguish between users without Public Profiles (under 16) and users with Public Profiles (16+) who have not edited the Profile in any way. We have additional mitigations for our older teenager with a Public Profile:
 - When they first interact with their Public Profile page, post to Spotlight or Snap Map, or share a Public Story, they are shown a dedicated notice explaining what Public Profiles are and how to use them appropriately. This notice links to the [support pages](#) providing more information on Public Profiles;
 - they can decide whether to make each piece of content public or private when posting; and
 - as with all Snapchatters, they have control over each piece of content they create with intentional posting options that let them determine where Snaps are shared, who can see them, and if they are saved to their profile;
 - their Spotlight will be publicly accessible with their username until the user deletes or hides the post, but will not be recommended to others until it passes our auto/human moderation checks;
 - their Public Stories posted by users 16-17 will be recommended only to the following categories of users but not on Discover and Spotlight online platforms:
 - Friends in the Friends Story Carousel of the 4th tab



- Followers in the Following Story Carousel of the 4th tab
- Friends of Friends (with a significant number of mutual friends) in the Friends Carousel of the 4th tab
- users replying to a Public Story are shown a notice explaining that their reply may be made public;
- they will continue not to have the option to post to Snap Map with their name shown; accepted submissions to the Snap Map will be displayed anonymously as they are now;
- older teens will have more limited access to analytics on Snapchat about the performance of their Public Stories and Spotlight videos, which creators use to grow an audience. In particular, they will not see how many people “favorited” their Stories or Spotlights, keeping the focus on creativity over pressure to collect public approval metrics.
- We give users control over their ability to be contacted. We allow users with access to a Public Story to turn off all Story reply messages so they don’t see messages from users who reply to their Stories. We also give users control over Story Replies and filter out words they don’t want to see. Users can input words that they don’t want to receive in the story replies from their fans. If a story reply contains an inputted word, the user does not receive the story reply (and any other story replies) from the sender. Additionally, we allow creators to block repliers or report them.

We have also built in protections for users who engage with a Public Profile. For example, we inform users before they send a Story reply to a creator that the creator could quote the reply and make it publicly viewable (with the replier’s first name and Bitmoji). We also limit unwarranted connections between younger users and bad actors on the platform.

Snap Map

We have safeguards for creators and viewers regarding the content on Snap Map, for example:

- Creator protections
 - We filter Stories posted from users with new accounts so they do not feature on Snap Map.
 - Content posted to Public Stories will only show on Snap Map with a clear location if there are multiple users posting in a short time nearby and a percentage of those posters are non-Teen accounts.
 - Currently there is no ability to comment on Snap Map content.
 - We do not show “views” for Stories on Snap Map. Protects creators from feeling embarrassed or being subject to ridicule due to low number of views.
 - Content published by creators has limited publication duration (which may be changed by creators with a Snapchat+ subscription. This protects creators by ensuring their content is not available forever).



- Younger teen accounts (13-15) in the EU also don't have the ability to post to Public Stories which means that their Snaps are not eligible for the Snap Map. They can post directly to Snap Map but will not attribute their username to the post to minimise discoverability.
- Older teen accounts (16-17) in the EU have additional public posting options. Before launching this new functionality Snap conducted extensive testing and evaluation, and presentation to the Commission and other regulators, prior to launch. Our monitoring of key metrics showed that the EU launch performed in line with expectations. We have summarised the additional safeguards above.
- Viewer protections
 - When users, including younger individuals, use Snap Map, Snap collects and uses precise location data for the purpose of providing the feature to the Snapchatter

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

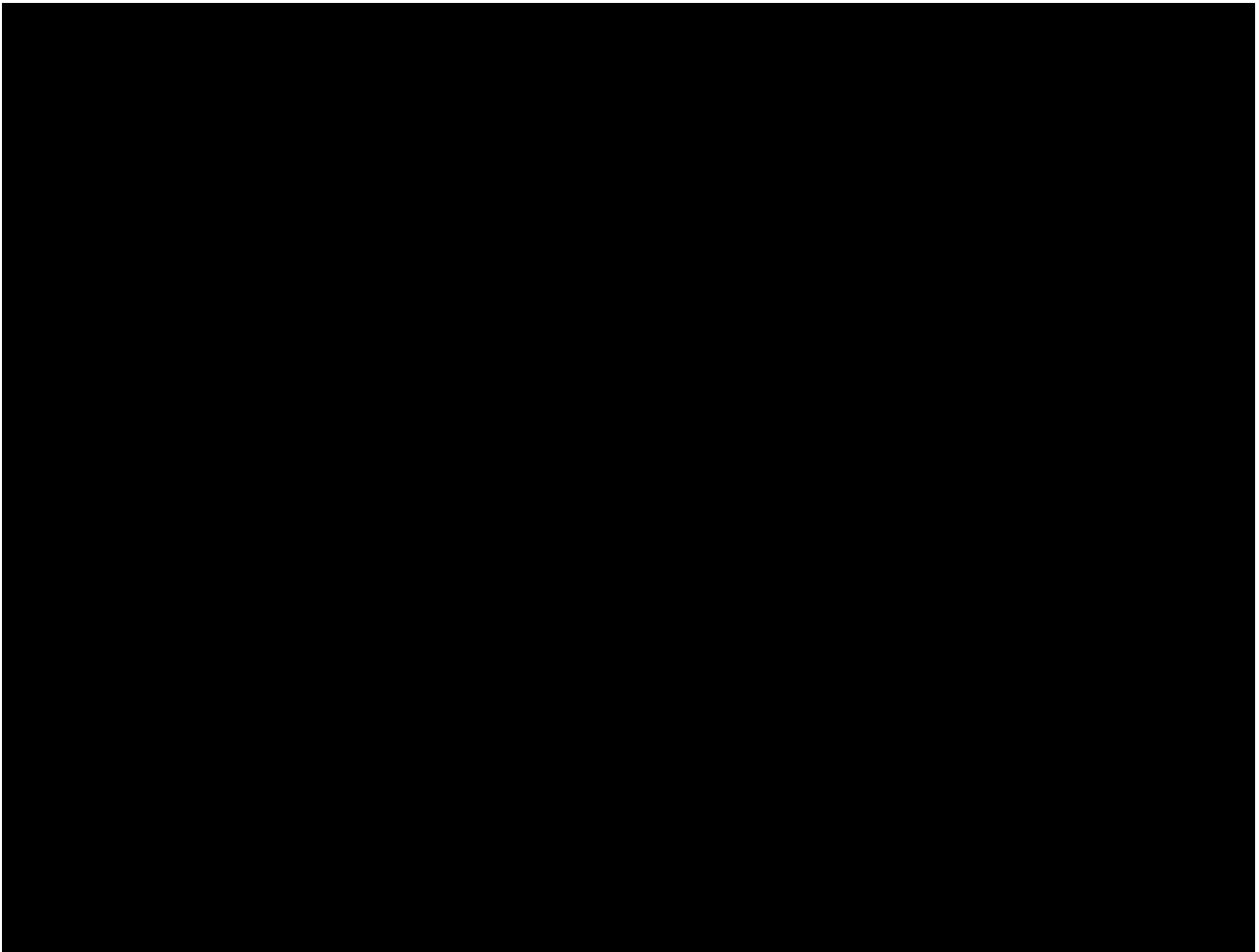
These controls prevent illegitimate use of Snap Map and protects Teens from exploitation.

With regards to location sharing on Snap Map, there are numerous additional design choices in place to make Snap Map a safer space for our community. These include:

- **Permission based and only Friend sharing.** Processing of users' precise location, including location information of children, is off by default. Given the sensitivity of geolocation data, Snap provides 'just in time' information to EU users about precise location when they access Snap Map. Clicking "Allow" on the 'just in time' information notice does not grant Snap with access to the user's precise location, but instead opens the user's device settings so that they can make the location sharing choice that they consider is most appropriate for them. This notice facilitates users' direct access to the precise location settings of their device. This permission based approach ensures that users retain control over what services they are requesting and how and when their precise location data is used.
- **No option to share location with strangers.** We want location sharing on Snap Map to be limited to engagement with friends on Snapchat. We also want to ensure user safety by not broadcasting a user's location to others who are not friends of the user. Therefore, users cannot share their location with strangers. In Settings, users can choose to share their location with their friends, or a subset of friends only. There is no option to share their location with non-friends. Friendship must be bi-directional.
- **Permission and prompts.** By default, users are not sharing their location with any friends, as all users are defaulted to "Ghost Mode". This was to ensure that location sharing would be understood by users before activation, in particular younger users so they could make informed choices about whether to use Snap Map, whether to share their location and, if



so, with whom to share it. Snapchatters can update: (1) whether they are sharing background location or only while using location; and (2) who among their friends can see their location at any time right from the settings gear in the Map.



- **Creator protections**

- Currently there is no comments on Snap Map content
- We do not show “views” for Stories on Snap Map. Protects creators from feeling embarrassed or being subject to ridicule due to low number of views.
- Content published by creators has a limited publication duration (which may be changed by creators with a Snapchat+ subscription. This protects creators by ensuring their content is not available forever).
- Creators are free to re-publish new and saved stories at any time, provided it does not violate the law or our Terms.

- **Viewer protections**

- We do not have public “favorites”, i.e. a user’s likes and interests are not public .



Lenses

Lenses (in popular language often dubbed as ‘filters’) are created by a relatively limited number of community developers, and Snap’s internal Lens Team. Our Lenses are designed with privacy-and-safety-by design principles in mind. For example, Lenses require object detection rather than facial identification. Lenses can tell what is or isn’t a face, they do not identify specific faces, limiting data processing for the use of Lenses. Snap does also not use any data collected by Lenses to customize the content that the user sees in Spotlight or Discover, nor is any data collected for advertising purposes. Besides, voice data collection of Snapchatters in the EU is off by default; it is only used to provide the service.

Snap also designs every Lens with race, gender, ethnicity and cultural norms in mind. Snap leverages its ever-growing diversity training datasets, as well as feedback from community members. If a Lens does not resonate with our community, as expressed through a high ratio of user reports, we take that feedback into consideration and will re-review the Lens with a goal to leave as-is, modify, or remove.

Advertising

We have also put in place risk mitigation measures for our advertising efforts. We prevent advertisers from manipulating small audiences with microtargeted campaigns, particularly for political ads. We do so by requiring a specific minimum audience of [REDACTED] Snapchatters to be targeted (including [Dynamic Ads on Snapchat](#) [REDACTED]). This prevents microtargeting that can influence voters politically or push targeted misinformation to certain populations. Our advertising systems also do not use ‘special category’ personal data to target ads and we require advertisers to provide additional information for political ads.

5.1.4 Integrations with other mitigations

On Snapchat we have also adapted our features to integrate with our other risk mitigations described in this Specific Mitigations section of the Report, for example:

Terms

All public content must adhere to our Terms, for example the content **must be suitable for 13+**, in order to be featured or receive broad distribution on Spotlight and Discover. This is explained in the [Terms Section](#) and [Transparency Section](#).

Content Moderation

We moderate content on Snapchat in a number of ways to mitigate the risks of users being exposed to harmful and illegal content. This is explained in the [Moderation Section](#) and [Enforcement Section](#).

Content Distribution

We have put in place risk mitigation measures to restrict the distribution of harmful content on



[REDACTED]

Snap. For example:

- Content that is Sexually Suggestive and Sensitive but otherwise allowed under our Community Guidelines is not distributed to Teens.
- Spotlight Comments with abusive language are removed.
- Ranking avoids ‘filter bubbles’ through demotion, ensuring similar content isn’t sequentially recommended to Snapchatters in Discover or Spotlight.

This is explained in [Sections 5.6](#) (Algorithmic Systems) and [5.7](#) (Advertising Systems).

5.1.5 Online Interface Design Process

Snap implemented the following process and governance around online interface design:

[REDACTED]

- 1. [REDACTED]
- 2. [REDACTED]
- 3. [REDACTED]

[REDACTED]

- 1. [REDACTED]
- 2. [REDACTED]
- 3. [REDACTED]
- 4. [REDACTED]
- 5. [REDACTED]
- 6. [REDACTED]
- 7. [REDACTED]
- 8. [REDACTED]
- 9. [REDACTED]
- 10. [REDACTED]
- 11. [REDACTED]
- 12. [REDACTED]
- 13. [REDACTED]
- 14. [REDACTED]
- 15. [REDACTED]
- 16. [REDACTED]
- 17. [REDACTED]
- 18. [REDACTED]
- 19. [REDACTED]
- 20. [REDACTED]
- 21. [REDACTED]
- 22. [REDACTED]
- 23. [REDACTED]
- 24. [REDACTED]
- 25. [REDACTED]
- 26. [REDACTED]
- 27. [REDACTED]
- 28. [REDACTED]
- 29. [REDACTED]
- 30. [REDACTED]
- 31. [REDACTED]
- 32. [REDACTED]
- 33. [REDACTED]
- 34. [REDACTED]
- 35. [REDACTED]
- 36. [REDACTED]
- 37. [REDACTED]
- 38. [REDACTED]
- 39. [REDACTED]
- 40. [REDACTED]
- 41. [REDACTED]
- 42. [REDACTED]
- 43. [REDACTED]
- 44. [REDACTED]
- 45. [REDACTED]
- 46. [REDACTED]
- 47. [REDACTED]
- 48. [REDACTED]
- 49. [REDACTED]
- 50. [REDACTED]
- 51. [REDACTED]
- 52. [REDACTED]
- 53. [REDACTED]
- 54. [REDACTED]
- 55. [REDACTED]
- 56. [REDACTED]
- 57. [REDACTED]
- 58. [REDACTED]
- 59. [REDACTED]
- 60. [REDACTED]
- 61. [REDACTED]
- 62. [REDACTED]
- 63. [REDACTED]
- 64. [REDACTED]
- 65. [REDACTED]
- 66. [REDACTED]
- 67. [REDACTED]
- 68. [REDACTED]
- 69. [REDACTED]
- 70. [REDACTED]
- 71. [REDACTED]
- 72. [REDACTED]
- 73. [REDACTED]
- 74. [REDACTED]
- 75. [REDACTED]
- 76. [REDACTED]
- 77. [REDACTED]
- 78. [REDACTED]
- 79. [REDACTED]
- 80. [REDACTED]
- 81. [REDACTED]
- 82. [REDACTED]
- 83. [REDACTED]
- 84. [REDACTED]
- 85. [REDACTED]
- 86. [REDACTED]
- 87. [REDACTED]
- 88. [REDACTED]
- 89. [REDACTED]
- 90. [REDACTED]
- 91. [REDACTED]
- 92. [REDACTED]
- 93. [REDACTED]
- 94. [REDACTED]
- 95. [REDACTED]
- 96. [REDACTED]
- 97. [REDACTED]
- 98. [REDACTED]
- 99. [REDACTED]
- 100. [REDACTED]

[REDACTED]

- 1. [REDACTED]
- 2. [REDACTED]
- 3. [REDACTED]
- 4. [REDACTED]
- 5. [REDACTED]
- 6. [REDACTED]
- 7. [REDACTED]
- 8. [REDACTED]
- 9. [REDACTED]
- 10. [REDACTED]
- 11. [REDACTED]
- 12. [REDACTED]
- 13. [REDACTED]
- 14. [REDACTED]
- 15. [REDACTED]
- 16. [REDACTED]
- 17. [REDACTED]
- 18. [REDACTED]
- 19. [REDACTED]
- 20. [REDACTED]
- 21. [REDACTED]
- 22. [REDACTED]
- 23. [REDACTED]
- 24. [REDACTED]
- 25. [REDACTED]
- 26. [REDACTED]
- 27. [REDACTED]
- 28. [REDACTED]
- 29. [REDACTED]
- 30. [REDACTED]
- 31. [REDACTED]
- 32. [REDACTED]
- 33. [REDACTED]
- 34. [REDACTED]
- 35. [REDACTED]
- 36. [REDACTED]
- 37. [REDACTED]
- 38. [REDACTED]
- 39. [REDACTED]
- 40. [REDACTED]
- 41. [REDACTED]
- 42. [REDACTED]
- 43. [REDACTED]
- 44. [REDACTED]
- 45. [REDACTED]
- 46. [REDACTED]
- 47. [REDACTED]
- 48. [REDACTED]
- 49. [REDACTED]
- 50. [REDACTED]
- 51. [REDACTED]
- 52. [REDACTED]
- 53. [REDACTED]
- 54. [REDACTED]
- 55. [REDACTED]
- 56. [REDACTED]
- 57. [REDACTED]
- 58. [REDACTED]
- 59. [REDACTED]
- 60. [REDACTED]
- 61. [REDACTED]
- 62. [REDACTED]
- 63. [REDACTED]
- 64. [REDACTED]
- 65. [REDACTED]
- 66. [REDACTED]
- 67. [REDACTED]
- 68. [REDACTED]
- 69. [REDACTED]
- 70. [REDACTED]
- 71. [REDACTED]
- 72. [REDACTED]
- 73. [REDACTED]
- 74. [REDACTED]
- 75. [REDACTED]
- 76. [REDACTED]
- 77. [REDACTED]
- 78. [REDACTED]
- 79. [REDACTED]
- 80. [REDACTED]
- 81. [REDACTED]
- 82. [REDACTED]
- 83. [REDACTED]
- 84. [REDACTED]
- 85. [REDACTED]
- 86. [REDACTED]
- 87. [REDACTED]
- 88. [REDACTED]
- 89. [REDACTED]
- 90. [REDACTED]
- 91. [REDACTED]
- 92. [REDACTED]
- 93. [REDACTED]
- 94. [REDACTED]
- 95. [REDACTED]
- 96. [REDACTED]
- 97. [REDACTED]
- 98. [REDACTED]
- 99. [REDACTED]
- 100. [REDACTED]



[Redacted]

- [Redacted]
[Redacted]
[Redacted]
- [Redacted] [Redacted]
[Redacted]
- [Redacted] [Redacted]
[Redacted]
 - [Redacted] [Redacted]
[Redacted] [Redacted]
[Redacted] [Redacted]
 - [Redacted] [Redacted]
 - [Redacted] [Redacted]
[Redacted] [Redacted]
[Redacted] [Redacted]
 - [Redacted] [Redacted]
[Redacted] [Redacted]
[Redacted] [Redacted]
- [Redacted] [Redacted]
[Redacted] [Redacted]

[Redacted]

- [Redacted] [Redacted]
[Redacted] [Redacted]
- [Redacted] [Redacted]
[Redacted] [Redacted]
 - [Redacted] [Redacted]
[Redacted] [Redacted]
 - [Redacted] [Redacted]
[Redacted] [Redacted]
[Redacted] [Redacted]
[Redacted] [Redacted]
[Redacted] [Redacted]
- [Redacted] [Redacted]
[Redacted] [Redacted]
 - [Redacted] [Redacted]
[Redacted] [Redacted]
[Redacted] [Redacted]
[Redacted] [Redacted]
- [Redacted] [Redacted]
[Redacted] [Redacted]



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Online Design Principles

Snap also established [REDACTED] Online Design Principles which prohibit the use of dark patterns & misleading nudge techniques. [REDACTED]

- [REDACTED]

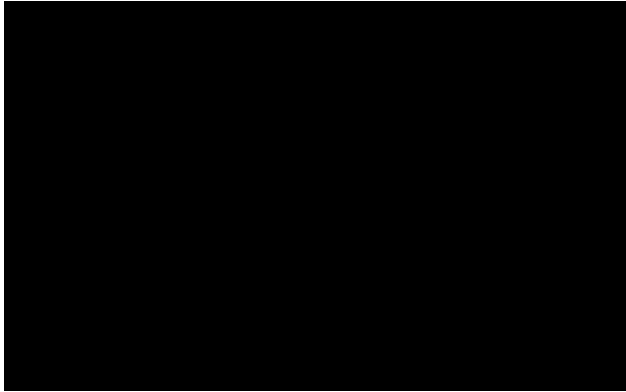
[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



5.1.6 Conclusion

From day one, Snap has made conscious design decisions to mitigate systemic risks from occurring on its platform, including privacy and safety by design decisions such as shorter retention periods, default Story visibility to just friends, not promoting likes on a user's Story, not having public friend lists, and maintaining proactive content policies. [REDACTED]

[REDACTED] Snap has implemented additional measures as further outlined in the remainder of this Report.



As explained in Section 4, we have concluded that the adaptations made by Snap to the design, features and functioning of Snapchat’s in-scope services, in combination with the other mitigations explained in this Section 5, are reasonable, proportionate and effective mitigation measures for the risks identified.

5.2 Terms

5.2.1 Introduction

This document outlines Snap’s protocol for communicating its platform Terms and Conditions to users, in compliance with the requirements of the DSA, in particular with regard to Articles 14 and 27.

Snap publishes its Terms and Conditions (which comprises the terms, guidelines, and policies defined in the section “Terms and Conditions” below) with concise summaries in clear, easily understandable, unambiguous language, in a publicly available, easily accessible, and machine-readable format. Snap’s Terms and Conditions include detail on (among other things): how use of the service may be restricted; content moderation policies and procedures; information on the use of algorithms, the parameters and criteria behind recommender system functioning, and how to adjust them; instances when user access and/or content may be restricted, suspended or terminated; and instructions on the internal complaint handling system. This information is primarily provided in Snap’s [Terms of Service](#), which are translated into all official EU member state languages. Snap’s [Terms of Service](#) prohibit any use of Snapchat to conduct illegal activities.

In addition to the [Terms of Service](#), Snap also publishes [Community Guidelines](#) (incorporated into the [Terms of Service](#) by reference), Privacy Policies, product-specific terms, a Commercial Content Policy, [Content Guidelines for Recommendation Eligibility](#), and Advertising Policies. Snap’s Community Guidelines elaborate on restrictions of the use of Snapchat related to: Sexual Content, Harassment and Bullying, Threats, Violence & Harm, Harmful False or Deceptive Information, Illegal or Regulated Activities, and Hateful Content, Terrorism, and Violent Extremism. Notably, the sections within our Community Guidelines on Illegal or Regulated Activities expressly prohibit users from using Snapchat to send or post content that’s illegal in their jurisdiction, or for any illegal activity.

5.2.2 Oversight and Administration

Change Management

On an annual basis (or sooner on an ad hoc basis should a pressing need arise), Snap’s Management body (including Legal and Public Policy team stakeholders) review and update various Snapchat Terms and Conditions (including [Community Guidelines](#), [Terms of Service](#),

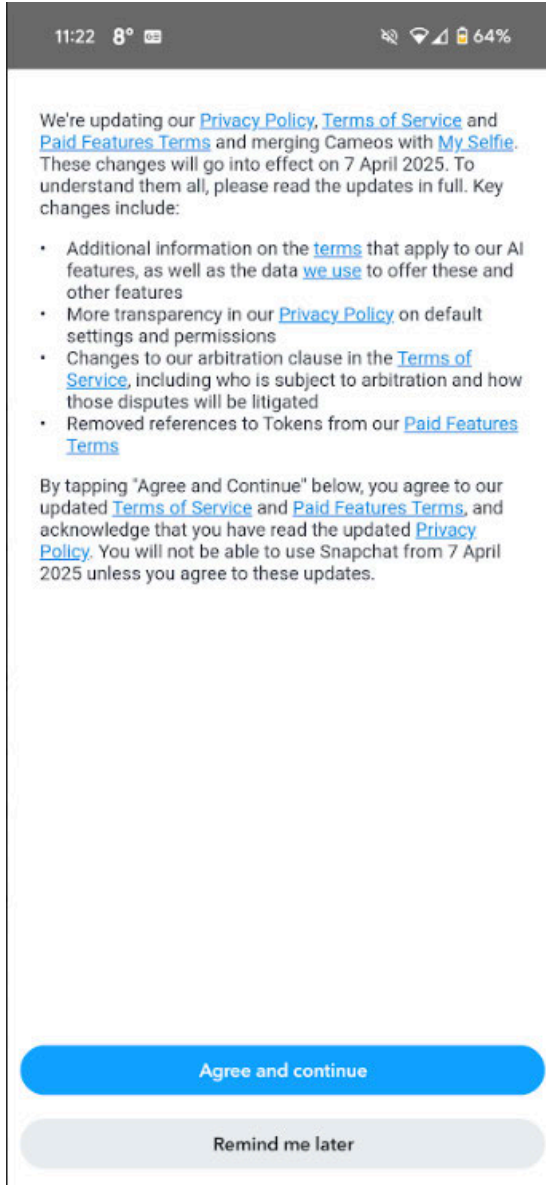


Privacy Policies, and Advertising Policies) for additional information that may result in impact on our risk and mitigation assessments, and to ensure that the Terms/Guidelines accurately reflect the contractual relationship and/or other obligations between users of the respective Services and Snap and adhere to applicable legal requirements.

Snap's Legal team has a formal process in place to make and track changes to the Terms and Conditions and to communicate key changes to stakeholders in a timely manner.

Snap's Terms and Conditions are regularly reviewed (and updated as needed) by Legal and Policy team stakeholders, including Commercial, Privacy, and Product Legal teams and Platform Policy, to ensure that they accurately reflect the contractual relationship and/or other obligations between users of the respective Services and Snap and adhere to applicable legal requirements. When updates are identified, depending on the term, policy, or guideline, proposed revisions are drafted and approved by the owner of the respective document. For certain terms, like the [Terms of Service](#), a changelog may be created to further document the update and approval process. Once the document is finalized and approved, it is then localized in all supported languages, including all official languages of the European Union as explicitly required by the DSA. To the extent the changes to the terms, policies, or guidelines are determined to be material, Snap will provide users with reasonable advance notice.

Snap provides an in-app pop-up to notify recipients of the service of material changes to the Terms and Conditions.



Roles and Responsibilities

Roles	Responsibilities
Snap Legal	• Maintains a formal process to draft, review, update, approve, and communicate changes to Snap's Terms and Conditions.
Public Policy	• Consults with industry experts, including the Safety Advisory Board, on Snap's Terms and Conditions as needed. • Reviews proposed changes to Snap's Terms and Conditions.
DSA Independent Compliance Officers	• Review Snap's Terms and Conditions to ensure they meet DSA requirements.



	• Provides oversight over Snap’s Content Moderation mechanisms to ensure the diligent, objective, and proportional enforcement of Snap’s Terms and Conditions.
--	--

5.2.3 Terms and Conditions

Snap publishes and maintains a series of legal documentation that make up our Terms and Conditions and govern use of Snap’s products and services by its users:

- [Terms of Service](#)
- [Community Guidelines](#)
- [Privacy Policy](#)
- [Snap and Ads](#)
- [How We Rank Content on Spotlight](#)
- [How We Rank Content on Discover](#)
- [How We Rank Content on Lenses](#)
- [How We Rank Content on Maps](#)

In addition, in response to the Digital Services Act we have provided short summaries in each section of our [Terms of Service](#), as well as easy to read explainers of key sections of our [Community Guidelines](#).

Each of our Terms and how they mitigate each of the DSA risk categories is explained below.

Terms of Service

All Snapchatters are required to agree to Snap’s [Terms of Service](#) before they can use Snapchat.



<

Create account

Step 1 of 5

What's your name?

First and Last name

By tapping Continue, you acknowledge that you have read the [Privacy Policy](#) and agree to the [Terms of Service](#).

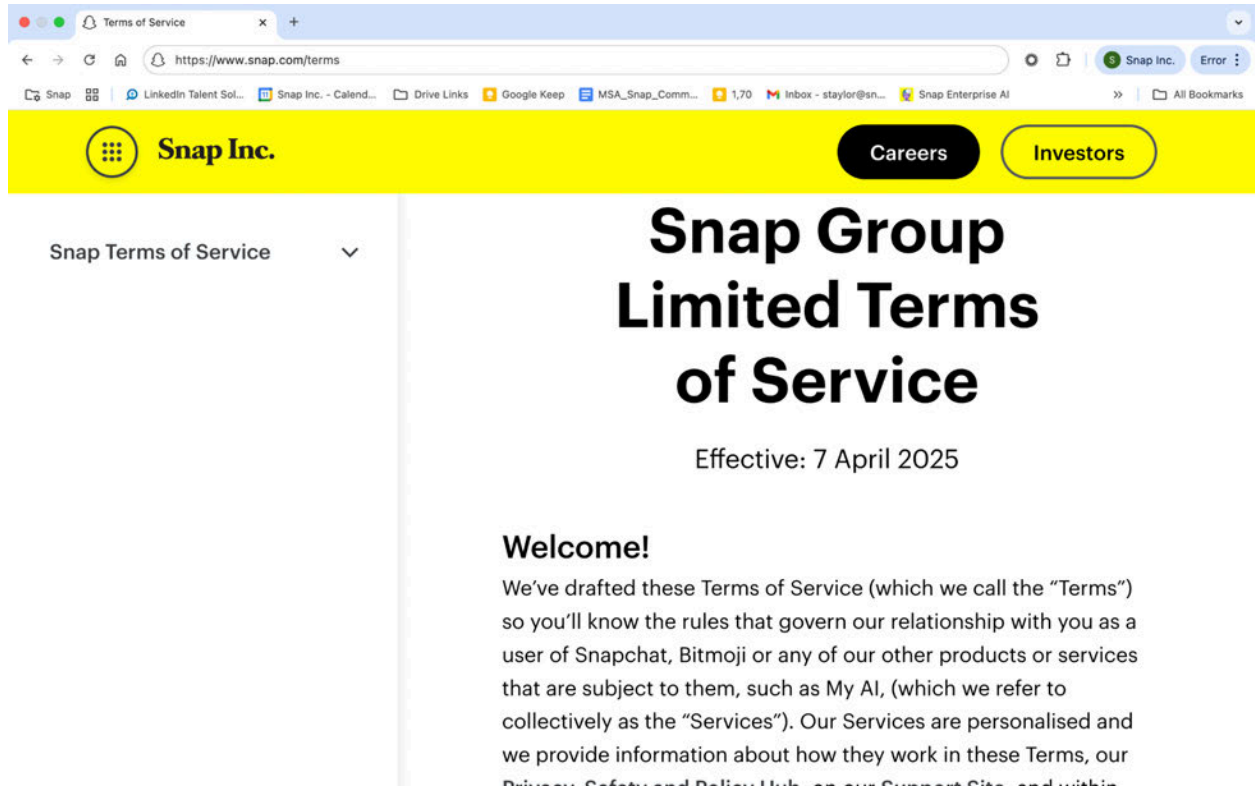
Snapchat is a [personalized, ad-funded service](#). This means we will use data about you, your friends, our community, and information provided by advertisers and partners including your activity on their websites and apps, to show you content and ads on Snapchat that we think you'll be interested in and measure the performance of those ads.

Continue

Snap publishes its [Terms of Service](#) on the website and in-app within a user's app settings. It is easily accessible via search engine and machine readable.

The [Terms of Service](#) include information on:

- Restrictions imposed on use of services (these are elaborated upon in Snap's Community Guidelines, Privacy Policy, and Advertising Policy)
- Policies, procedures, measures and tools used for the purpose of content moderation, including algorithmic decision-making and human review
- Rules of procedure of Snap's internal complaint handling system and available remedies and redress mechanisms.
- Main parameters used in Snap's recommender systems.



Restrictions Imposed

The following restrictions are included in Snap's [Terms of Service](#).

Who Can Use the Services

Our Services are not directed to children under the age of 13, and you must confirm that you are 13 years or older to create an account and use the Services

Respecting the Services and Snap's Rights

You must also respect Snap's rights and adhere to the Snapchat Brand Guidelines, Bitmoji Brand Guidelines, and any other guidelines, support pages, or FAQs published by Snap or our affiliates.

Respecting Others' Rights

You therefore may not use the Services, or enable anyone else to use the Services, in a manner that violates or infringes someone else's rights of publicity, privacy, copyright, trademark, or other intellectual property right.

Safety

By using the Services, you agree that you will at all times comply with these Terms, including our Community Guidelines and any other policies Snap makes available in order to maintain the safety of the Services.



Content Moderation

The following Content Moderation information is included in Snap's [Terms of Service](#).

Much of the content on our Services is produced by users, publishers, and other third parties. Whether that content is posted publicly or sent privately, the content is the sole responsibility of the user or entity that submitted it. Although Snap reserves the right to review, moderate, or remove all content that appears on the Services, we do not review all of it. So we cannot — and do not — guarantee that other users or the content they provide through the Services will comply with our Terms, Community Guidelines or our other terms, policies or guidelines. You can read more about Snap's approach to content moderation on our [Support Site](#).

Users can report content produced by others or others' accounts for violation of our Terms, Community Guidelines or other guidelines and policies. More information about how to report content and accounts is available on our [Support Site](#).

We may restrict, terminate, or temporarily suspend your access to the Services if you fail to comply with these Terms, our Community Guidelines or the law, for reasons outside of our control, or for any other reason. That means that we may terminate these Terms, stop providing you with all or any part of the Services, or impose new or additional limits on your ability to use our Services. For example, we may deactivate your account due to prolonged inactivity, and we may reclaim your username at any time for any reason. And while we'll try to give you reasonable notice beforehand, we can't guarantee that notice will be possible in all circumstances.

Before we restrict, terminate or suspend your access to the Services, we will take into account all relevant facts and circumstances apparent from the information available to us, depending on the underlying reason for taking that action. For example, if you violate our Community Guidelines we consider the severity, frequency, and impact of the violations as well as the intention behind the violation. This will inform our decision whether to restrict, terminate or suspend your access to the Services and, in the event of suspension, how long we suspend your access. You can find out more about how we assess and take action against misuse of our Services on our [Support Site](#).

Internal Complaint Handling

The following information on Snap's Internal Complaint Handling is included in Snap's [Terms of Service](#).

Where we restrict, terminate or suspend your access to the Services for violation of our Community Guidelines, we will notify you and provide an opportunity for you to appeal as explained in our [Moderation, Enforcement and Appeals explainer](#).



We hope you'll understand any decisions we make about content or user accounts, but if you have any complaints or concerns, you can use the submission form available here or use available in-app options. If you use this process, your complaint must be submitted within six months of the relevant decision.

Upon receiving a complaint, we will:

- *ensure the complaint is reviewed in a timely, non-discriminatory, diligent and non-arbitrary manner;*
- *reverse our decision if we determine our initial assessment was incorrect; and*
- *inform you of our decision and of any possibilities for redress promptly.*

Recommender Systems

The following information on Snap's Recommender Systems is included in Snap's [Terms of Service](#).

Our Services provide a personalized experience to make them more relevant and engaging for you. We will recommend content, advertising and other information to you based on what we know and infer about your and others' interests from use of our Services. It is necessary for us to handle your personal information for this purpose, as we explain in our [Privacy Policy](#). Personalization is also a condition of our contract with you for us to be able to do so, unless you opt to receive less personalization in the Services. You can find more information on personalized recommendations on our [Support Site](#).

Community Guidelines

In our Community Guidelines, which are explicitly incorporated into our [Terms of Service](#), we provide further guidance on the categories of illegal content, and content that Snap deems in violation of its Terms. The Community Guidelines are easily accessible via Search Engine and in Snap's Transparency Center and are machine readable and easily understandable.



Our Community Guidelines are broken up into the following sections: Sexual Content, Harassment and Bullying, Threats, Violence & Harm, Harmful False or Deceptive Information, Illegal or Regulated Activities, and Hateful Content, Terrorism, and Violent Extremism.

These categories have been fine tuned over many years of content moderation on Snapchat, and encompass the illegal content that we have encountered on Snapchat over the years. For ease of reference we have incorporated a more detailed breakdown of each category below.

Sexual Content

- *We prohibit any activity that involves sexual exploitation or abuse of a minor, including sharing child sexual exploitation or abuse imagery, grooming for sexual purposes, or sexual extortion (sextortion), or the sexualization of children.*
- *We prohibit any communication or behaviour that attempts to persuade, trick or coerce a minor with the intent of sexual abuse or exploitation, or which leverages fear or shame to keep a minor silent.*
- *We prohibit all other forms of sexual exploitation, including sex trafficking, sextortion and deceptive sexual practices, including efforts to coerce or entice users to provide nudes.*
- *We prohibit producing, sharing or threatening to share non-consensual intimate imagery (NCII) – including sexual photos or videos taken or shared without permission, as well as “revenge porn” or behaviour that threatens to share, exploit or expose individuals’ intimate images or videos without their consent.*



- *We prohibit all forms of sexual harassment. This may include making unwanted advances, sharing graphic and unsolicited content, or sending obscene requests or sexual invitations to other users.*
- *We prohibit promoting, distributing, or sharing pornographic content, including photos, videos or even highly realistic animation, drawings or other renderings of explicit sex acts, or nudity where the primary intention is sexual arousal.*
- *We prohibit offers of sexual services, including both offline services (such as, for example, erotic massage) and online experiences (such as, for example, offering sexual chat or video services).*

Additional guidance on sexual conduct and content that violates our Community Guidelines is available [here](#).

These Terms make clear to Snapchatters the extent to which sexual content is prohibited. This reduces the likelihood of several risks falling with categories identified by the DSA, including in particular: (i) the dissemination of Child Sexual Abuse Material and Adult Sexual Content in Category 1, (ii) the Right to Human Dignity and Children's Rights in Category 2 and (iii) Negative Effects on Public Health, Minors, and Gender-Based Violence in Category 4.

Harassment and Bullying

- *We prohibit bullying or harassment of any kind. This extends to all forms of sexual harassment, including sending unwanted sexually explicit, suggestive, or nude images to other users. If someone blocks you, you may not contact them from another Snapchat account.*
- *Sharing images of a person in a private space — like a bathroom, bedroom, locker room, or medical facility — without their knowledge and consent is prohibited, as is sharing another person's private information without their knowledge and consent or for the purpose of harassment (i.e., "doxxing").*
- *If someone is depicted in your Snap and asks you to remove it, please do! Respect the privacy rights of others.*
- *Please also do not harass another Snapchatter by abusing our reporting mechanisms, such as intentionally reporting content that is permissible.*
- Additional guidance on how Harassment and Bullying violate our Community Guidelines is available [here](#).

These Terms make clear to Snapchatters that the extent to which harassment and bullying is prohibited. This reduces the likelihood of several risks falling with categories identified by the DSA, including in particular: (i) the Right to Human Dignity, Private Life and Data Protection, and Children's Rights, in Category 2 and (ii) Negative Effects on Public Health, Minors, and Gender-Based Violence, as well as serious negative consequences to a person's Physical and Mental Well-Being in Category 4.



Threats, Violence and Harm

- *Encouraging or engaging in violent or dangerous behavior is prohibited. Never intimidate or threaten to harm a person, a group of people, or someone's property.*
- *Snaps of gratuitous or graphic violence, including animal abuse, are not allowed.*
- *We don't allow the glorification of self-harm, including the promotion of self-injury, suicide, or eating disorders.*
- Additional guidance on threats, violence, and harm that violate our Community Guidelines is available [here](#).

These Terms make clear to Snapchatters the extent to which threats and violence are prohibited. This reduces the likelihood of several risks falling with categories identified by the DSA, including in particular: (i) the Right to Human Dignity and Property, and Child Rights, in Category 2, (ii) Negative Effects on Civic Discourse and Public Security in Category 3 and (iii) Negative Effects on Public Health, Minors, and Gender-Based Violence, as well as serious negative consequences to a person's Physical and Mental Well-Being in Category 4.

Harmful, False, or Deceptive Information

- *We prohibit spreading false information that causes harm or is malicious, such as denying the existence of tragic events, unsubstantiated medical claims, undermining the integrity of civic processes, or manipulating content for false or misleading purposes (whether through generative AI or through deceptive editing).*
- *We prohibit pretending to be someone (or something) that you're not, or attempting to deceive people about who you are. This includes impersonating your friends, celebrities, public figures, brands, or other people or organizations for harmful, non-satirical purposes.*
- *We prohibit spam, including pay-for-follower promotions or other follower-growth schemes, the promotion of spam applications, or the promotion of multilevel marketing or pyramid schemes.*
- *We prohibit fraud and other deceptive practices, including the promotion of fraudulent goods or services or get-rich-quick schemes, or imitating Snapchat or Snap Inc.*
- Additional guidance on harmful false or deceptive content that violates our Community Guidelines is available [here](#).

These Terms make clear the extent to which harmful false or deceptive information is prohibited. This reduces the likelihood of several risks falling with categories identified by the DSA, including in particular: (i) the dissemination of Harmful False Information, Fraud and Spam in Category 1, (ii) the Right to Human Dignity, Private Life and Data Protection, and Children's Rights in Category 2, (iii) Negative Effects on Democratic and Electoral Processes, Civic Discourse and Public Security in Category 3 and (iv) Negative Effects on Public Health, Minors,



and Gender-Based Violence, as well as serious negative consequences to a person's Physical and Mental Well-Being in Category 4.

Illegal or Regulated Activities

- *Don't use Snapchat to send or post content that's illegal in your jurisdiction, or for any illegal activity. This includes promoting, facilitating, or participating in criminal activity, such as buying, selling, exchanging, or facilitating sales of illegal or regulated drugs, contraband (such as child sexual exploitation or abuse imagery), weapons, or counterfeit goods or documents. It also includes promoting or facilitating any form of exploitation, including sex trafficking, labor trafficking, or other human trafficking.*
- *We prohibit the illegal promotion of regulated goods or industries, including unauthorized promotion of gambling, tobacco or vape products, and alcohol.*
- Additional guidance on prohibited illegal or regulated activities that violate our Community Guidelines is available [here](#).

These Terms make clear the extent to which illegal or regulated activities are prohibited. This reduces the likelihood of several risks falling with categories identified by the DSA, including in particular: (i) the dissemination of illegal content, including child sexual abuse material and other types of misuse of Snapchat for criminal offences, and the conduct of illegal activities, such as the sale of products or services prohibited by European Union or Member State law, including dangerous or counterfeit products, or illegally-traded animals in Category 1, (ii) the Right to Property and Children's Rights in Category 2, (iii) Negative Effects on Public Security in Category 3 and (iv) Negative Effects on Public Health and Minors, as well as serious negative consequences to a person's Physical and Mental Well-Being in Category 4.

Hateful Content, Terrorism, or Violent Extremism

- *Terrorist organizations, violent extremists, and hate groups are prohibited from using our platform. We have no tolerance for content that advocates or advances terrorism or violent extremism.*
- *Hate Speech or content that demeans, defames, or promotes discrimination or violence on the basis of race, color, caste, ethnicity, national origin, religion, sexual orientation, gender, gender identity, disability, or veteran status, immigration status, socio-economic status, age, weight, or pregnancy status is prohibited.*
- Additional guidance on hateful content, terrorism, and violent extremism that violates our Community Guidelines is available [here](#).

These Terms make clear the extent to which Hate Speech and terrorism are prohibited. This reduces the likelihood of several risks falling with categories identified by the DSA, including in particular: (i) the dissemination of illegal Hate Speech and other types of misuse of Snapchat for criminal offenses and the conduct of illegal activities in Category 1, (ii) the Right to Human Dignity, Non-Discrimination and Children's Rights in Category 2, (iii) Negative Effects on Public

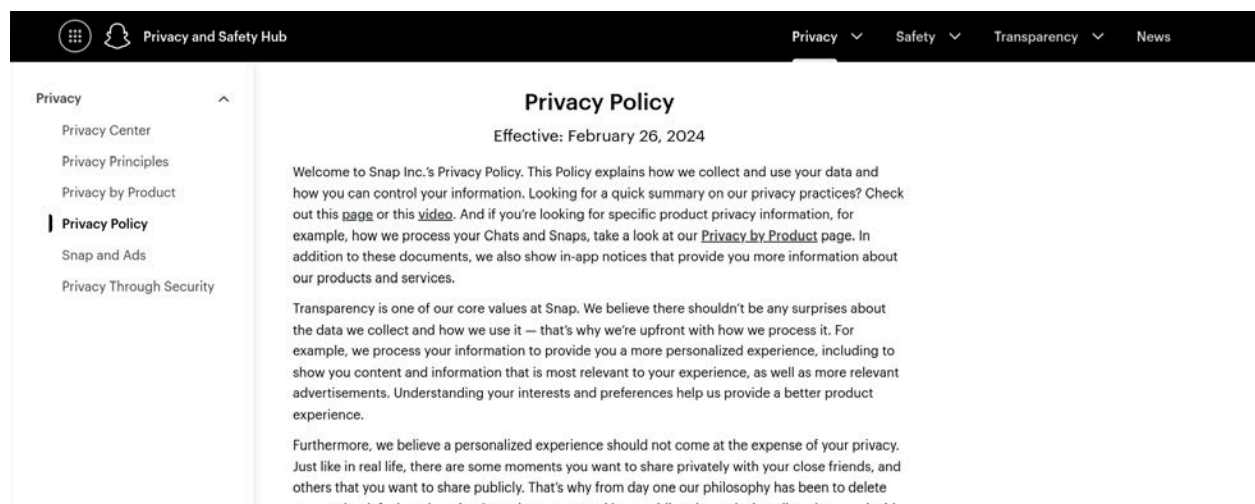


Security in Category 3 and (iv) Negative Effects on Minors, as well as serious negative consequences to a person's Physical and Mental Well-Being in Category 4.

We understand that each of the above categories can be nuanced and open to interpretation, that is why we have included explainers for each category.

Privacy Policy

Snap also publishes a Privacy Policy, which is presented when users register an account and it is easily accessible via Search Engine as well as, online, within our Safety and Privacy Hub and in the footer of Snap websites. It is also available on the App and in Apple/Android stores. It is machine readable and easily understandable.



Our Privacy Policy provides a detailed description of our privacy practices, including an explanation of how Snap collects and uses personal data and how individuals can control the processing of their information. We recognize that policies and terms can be overwhelming documents. That is why our long standing approach has been to provide additional, bite-sized information on our general practices (Privacy Center), our philosophy to privacy (Privacy Principles), what we do with user data (How we use your information), advertising (Snap and Ads), and specific products (Privacy by Product). In the introduction to our Privacy Policy, we state the following:

"We've done our best to write this Privacy Policy in a way that's easy to understand for all our users and free of difficult language and legal phrases. If you want to review something later on, you can always take a look at our [Privacy Center](#). We designed it to give you easy-to-digest summaries of our privacy practices. For example, our [Privacy by Product](#) page gives product-specific information and links to support pages with tips and tricks. Still have questions? Just [reach out](#) to us."



Privacy by Product gives users concise and easily understandable information about our products. For example, this webpage provides an overview of our approach to Snaps & Chats, as well as hyperlinks to more detailed information on specific aspects. Similarly, there's a section on Spotlight, Lenses, My AI, Stories, and many more products. In addition to these documents, we also show in-app notices that provide more information about our products and services.

Product Specific Terms

In addition to the Terms and Conditions described in detail above, we also have specific, publicly-available terms and policies that govern the use of additional aspects of Snapchat's features:

Spotlight

Snapchat users who choose to contribute content to Spotlight are required to adhere to the [Creator Monetization Policy](#). Snap also provides users who submit content to Spotlight with clear [Content Guidelines for Recommendation Eligibility](#), describing the policy, technical, and legal requirements for submissions to Spotlight, as well as reminding users of the Terms (including our [Community Guidelines](#)).

Discover

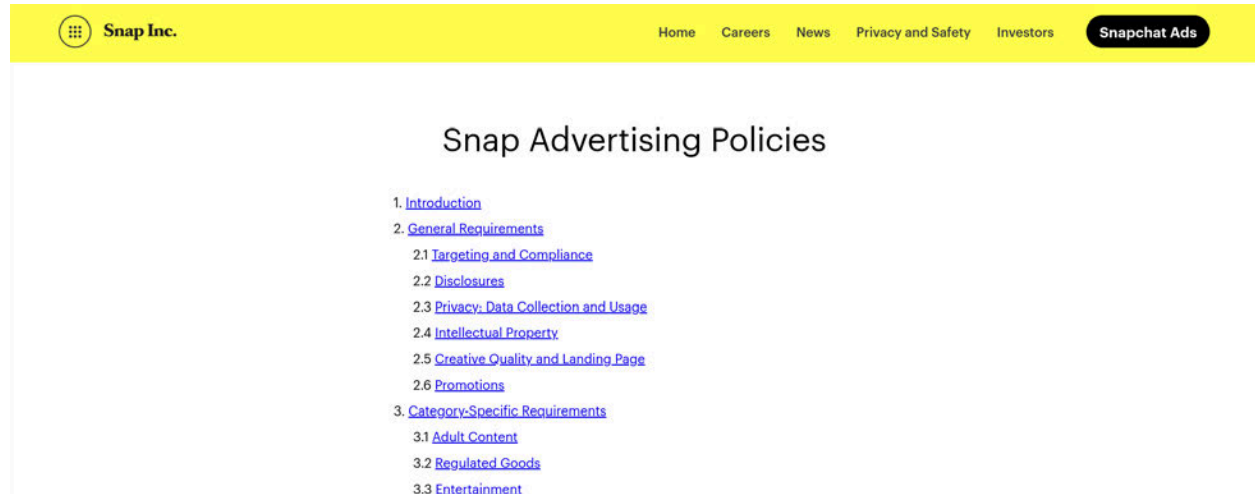
We have specific publishing agreements with our premium partners that post content on Discover, such as media organizations and Snap Stars, that require them to abide by our Terms (including our [Community Guidelines](#)).

Lenses

Snapchat users who choose to develop and submit Lenses for publication on Snapchat via Lens Studio must agree to the [Lens Studio Terms](#). Lenses must comply with our [Lens Studio Submission Guidelines](#), which also remind users of the Terms (including our [Community Guidelines](#)).

Advertising

Snap also publishes Advertising Policies, which outline the terms and conditions for use of Snap advertising services. These are easily accessible via Search Engine, machine readable, and easily understandable.



Snapchat users who choose to advertise to other users on Snapchat must agree to our [Snap Advertising Policies](#), including an obligation for advertisements to comply with applicable laws and rules in the European Union and each Member State where the update advertisements will run.

5.2.4 Accessing Terms and Conditions

Company	Community	Advertising	Legal
Snap Inc. Careers News Privacy and Safety	Snapchat Support Pixy Support Community Guidelines	Snapchat Ads Advertising Policies Political Ads Library Brand Guidelines Promotions Rules	Snap Terms Law Enforcement Cookie Policy Cookie Settings Report Infringement

Accessibility

Safety and Privacy Hub

Snap's Privacy and Safety Hub was launched last year and combines our Privacy Center, Safety Center, and Transparency Center all under one umbrella site. This is where Snap publishes formal transparency reports.

The rationale behind the integration of these three centers is that we believe there is a natural overlap between these areas, and that all the information provided in those domains contribute to providing awareness and building trust with our community and other stakeholders, such as parents, teachers, journalists, trusted flaggers, law enforcement, regulators, and NGOs.

The top navigation provides Policy, Privacy, Safety, and Transparency resources, as well as our latest News in those areas. In this section, we highlight a number of areas for illustration purposes, and refer to the [website](#) for further information.



Policy Center

Our Policy Center is a key resource for understanding the rules and policies that explain the rules and policies applicable across Snapchat. It includes links to our Community Guidelines, Content Guidelines for Recommendation Eligibility, Advertising Policies, Commercial Content Policy and Creator Monetization Policy.

Privacy Center

Our Privacy Policy provides a detailed description of our privacy practices, but we recognize that policies and terms can be overwhelming documents. That is why our long standing approach has been to provide additional, bite-sized information on our general practices (Privacy Center), our philosophy to privacy (Privacy Principles), what we do with user data (various pages, including the Privacy Policy), advertising (Snap and Ads), and specific products (Privacy by Product).

Safety Center

From the Privacy Center, users can easily navigate to the Safety Center, which provides an overview of our Safety resources, including tips on how to report a safety concern, information about our approach to safety partnerships, our Trusted Flagger Program, Safety Advisory Board, [Digital Well-Being Index](#) and more. Again, the goal here is to provide easy to navigate and process information.

Transparency Center

Our [Transparency Center](#) provides additional transparency resources to our users and to the public at large, including our Transparency Reports and EU-specific information required under the DSA.

On our [EU](#) transparency page, we publish EU-specific information required under the DSA, including the number of Average Monthly Active Recipients of our Snapchat app in the EU, and information about our legal representative in the EU, how EU law enforcement agencies can submit requests to snapchat, and the regulatory authorities that regulate us under the DSA.

News Page

Snap also frequently publishes related information on the Hub's [News](#) webpage. The purpose of these news articles is to inform the general public about recent developments on issues relevant to privacy, safety and transparency on Snapchat, and more.

5.2.5 Support Site

Our fully-searchable and thematically-organised [support site](#) contains the following terms and conditions which, again, are also accessible via links within our Terms of Service:

- [How We Rank Content on Spotlight](#)
- [How We Rank Content on Discover](#)



- [How We Rank Content on Lenses](#)
- [How We Rank Content on Maps](#)

5.2.6 Languages

Our [Terms of Service](#) have been translated into all official languages of the European Union as explicitly required by the DSA.

Bahasa Indonesia	Magyar	اردو
Bahasa Melayu	Malti	العربية
Dansk	Nederlands (Nederland)	मराठी
Deutsch (Deutschland)	Norsk (bokmål)	हिन्दी
Eesti (Estonia)	Polski	বাংলা (ভারত)
English (UK)	Português (Brasil)	বাংলা(বাংলাদেশ)
✓ English (US)	Português (Portugal)	ਪੰਜਾਬੀ
Español	Română	ગુજરાતી
Español (Argentina)	Slovenčina (Slovakia)	தமிழ்
Español (España)	Slovenščina	తెలుగు
Español (México)	Suomi	ಕನ್ನಡ (India)
Filipino (Philippines)	Svenska	മലയാളം
Français (France)	Tiếng Việt	ภาษาไทย (ประเทศไทย)
Gaeilge (Gaelic)	Türkçe	中文简体
Hrvatski	Čeština	中文繁體
Italiano	Ελληνικά	日本語
Latviešu	Български	한국어 (韩国)
Lietuvių	Русский	

5.2.7 Readability

As outlined in our previous Reports, our EU Snapchatter community consists of a diverse range of ages and genders. Snapchat services are not primarily directed at or used by minors. While Snapchat does have a young demographic, only a small percentage of European Union users fall within the 13-17 age category. The largest age category is 18-24 with 41% of European Union users.

Snap's Terms and Conditions have been designed to be a concise summary in clear, easily understandable, unambiguous language, in EU member state languages, in a publicly available, easily accessible, and machine-readable format, including summaries and explainers. This helps



all users to understand what activity is prohibited on Snapchat and the consequences, which reduces the likelihood they will engage in illegal or violating activity.

In addition, our Privacy and Safety Hub and Support pages as explained in the Transparency part of our Report have also been designed to be user friendly and easily understandable. For example, we created our Privacy and Safety Hub, with pages such as our Privacy by Product page, to give Snapchatters a high-level summary of our privacy and safety practices across each of our products and features. We also created a video to visualize our privacy practices, and use icons and other best practices as recommended by privacy and safety experts and the recognised Age Appropriate Design Code. This helps all users to understand how Snapchat works, what options they may have, how we moderate and enforce our terms and how they can get support. This reduces both the likelihood of illegal or violating activity and the severity of harm in the event they are exposed to illegal or violating activity despite our limitations. Teens reading our Privacy Center can understand how their data is being processed by Snap and find more information about relevant privacy settings which reduces the likelihood and severity of negative effects on Teens' data protection rights.

5.2.8 Conclusion

Snap provides terms and conditions for the recipients of its services, which incorporate the content and meet the language requirements of the DSA.

As explained in Section 4, we have concluded that Snap's terms and conditions, in combination with the other mitigations explained in this Section 5, are reasonable, proportionate and effective mitigation measures for the risks presented by Snapchat's in-scope services.

5.3 Transparency

Snap is focussed on providing users with the right level of information, at the right time. We understand that our community does not always have time to read multi page documents. This is why we strive to provide users with bite-sized information that is easy to access and understand, while also giving them an opportunity to review more detailed information where appropriate.

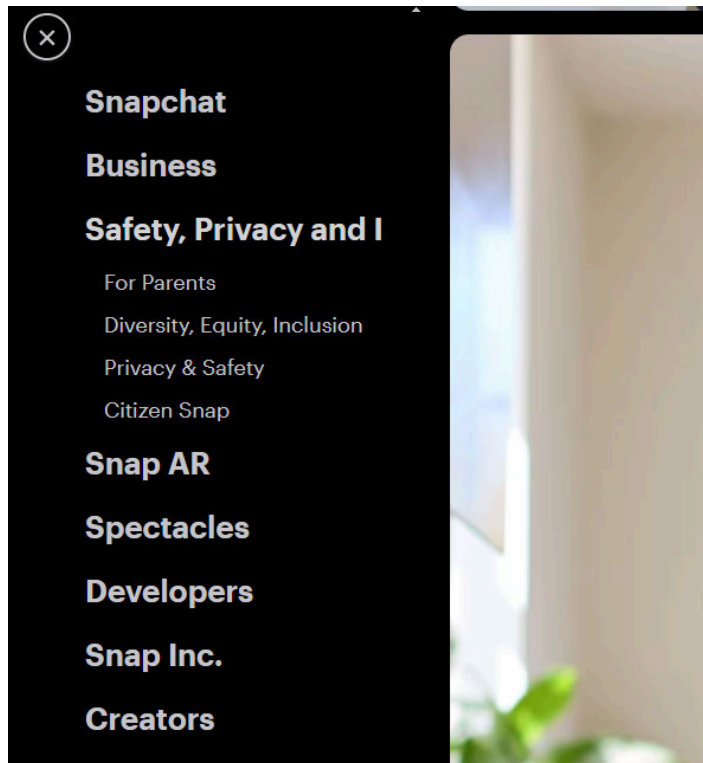
Information provided to users can be divided into three categories:

1. Information we provide on our website;
2. Information provided in app stores; and
3. Information we provide in our application.

5.3.1 Information we provide on our website

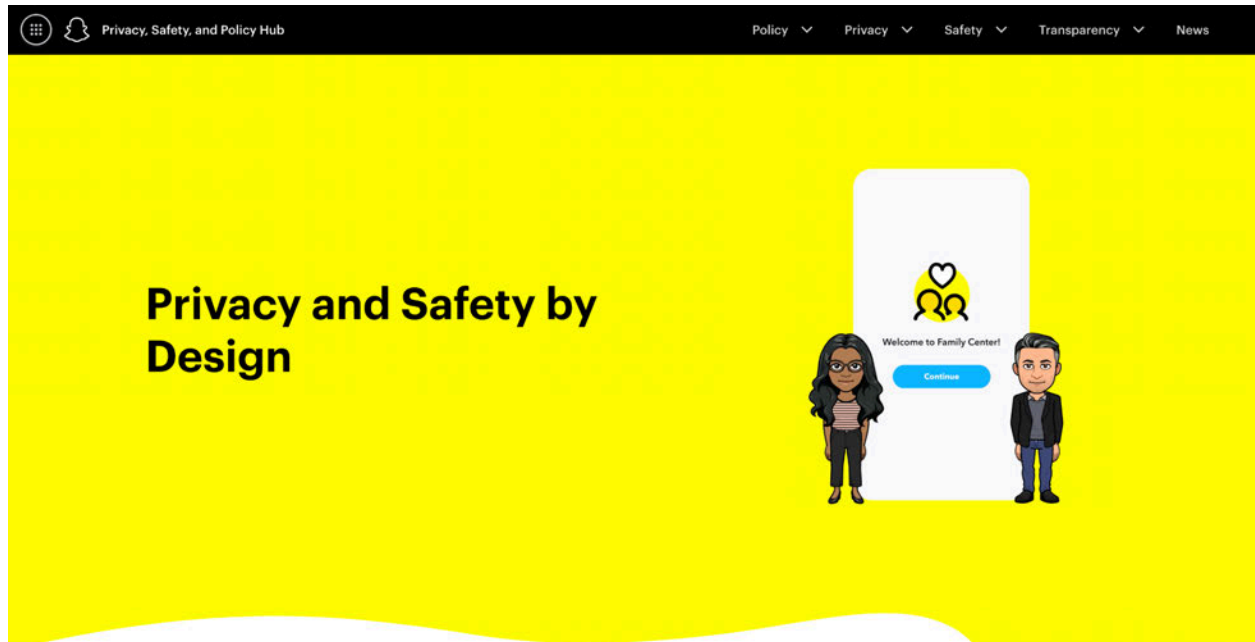


At Snap we have a number of avenues to provide information to users. The two primary sources of information outside of our application are our [Privacy, Safety, and Policy Hub](#) and our [Support Center](#).



Privacy, Safety, and Policy Hub

Snap's Privacy, Safety and Policy Hub was launched in 2022 and combined our Privacy Center, Safety Center and Transparency Center all under one umbrella, (with a dedicated Policy Hub added in 2024). The rationale behind this change is that we believe there is a natural overlap between these areas, and that all the information provided in those domains contribute to providing awareness and building trust with our community and other stakeholders, such as parents, teachers, journalists, trusted flaggers, law enforcement, regulators, and NGOs.



The top navigation provides Policy, Privacy, Safety, and Transparency resources, as well as our latest News in those areas. In this section, we highlight a number of areas for illustration purposes, and refer to the [website](#) for further information.

Policy Center

We want Snapchat to be a safe and positive experience for everyone who uses our platform or products. For this reason, we created rules and policies that explain the rights and responsibilities of all members of our community. Our Policy Center provides a central place for our Community Guidelines, Advertising Policies, Content Guidelines, and Commercial Content Policy. We have also reformatted our policies in a way that is easier and more intuitive for users to navigate, replacing our previously long, text-heavy pages with shorter, more digestible segments with clear headings and organization by subject matter.

Privacy Center

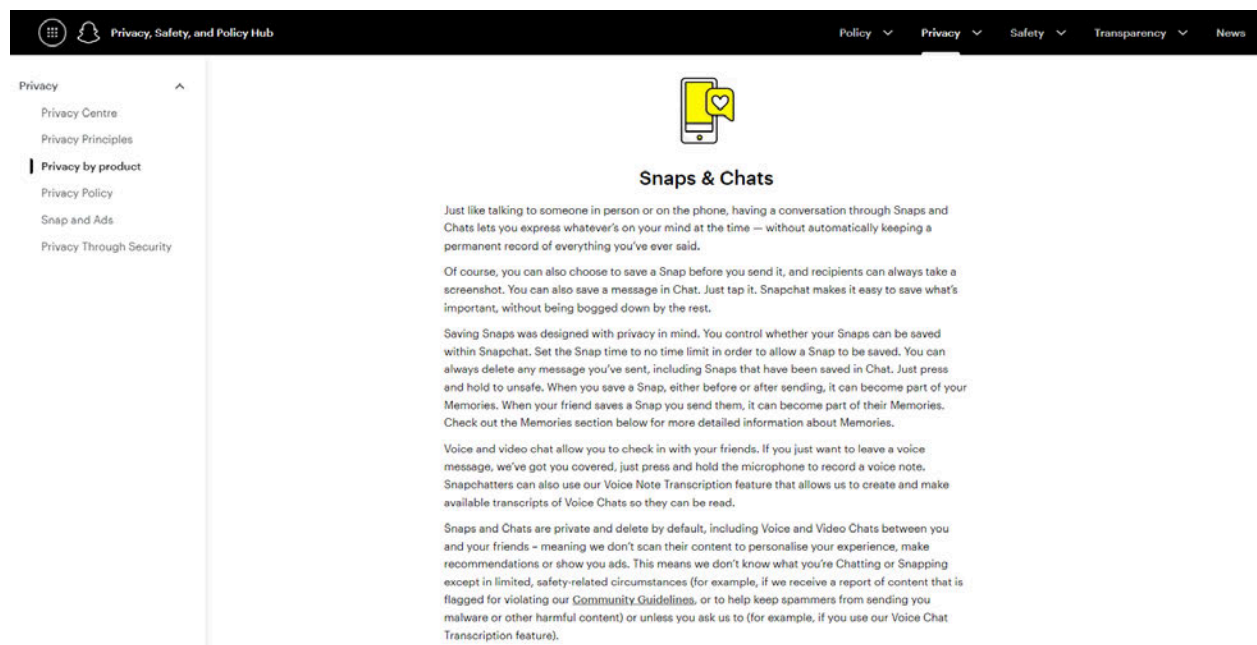
Our Privacy Policy provides a detailed description of our privacy practices, but we recognize that policies and terms can be overwhelming documents. That is why our long standing approach has been to provide additional, bite-sized information on our general practices (Privacy Center), our philosophy to privacy (Privacy Principles), what we do with user data (How we use your information), advertising (Snap and Ads), and specific products (Privacy by Product). In the introduction to our Privacy Center we state the following:

“Privacy policies tend to be pretty long – and pretty confusing. That’s why we did our best to make our [Privacy Policy](#) brief, clear, and easy-to-read!”



You should read our entire Privacy Policy, but when you only have a few minutes or want to remember something later on, you can always take a look at this summary – so you can learn or recall some of the basics in just a few minutes.”

Privacy by Product gives users concise and easily understandable information about our products. For example, this webpage provides an overview of our approach to Snaps & Chats, as well as hyperlinks to more detailed information on specific aspects. Similarly, there’s a section on Spotlight, Lenses, My AI, Stories, and many more products.



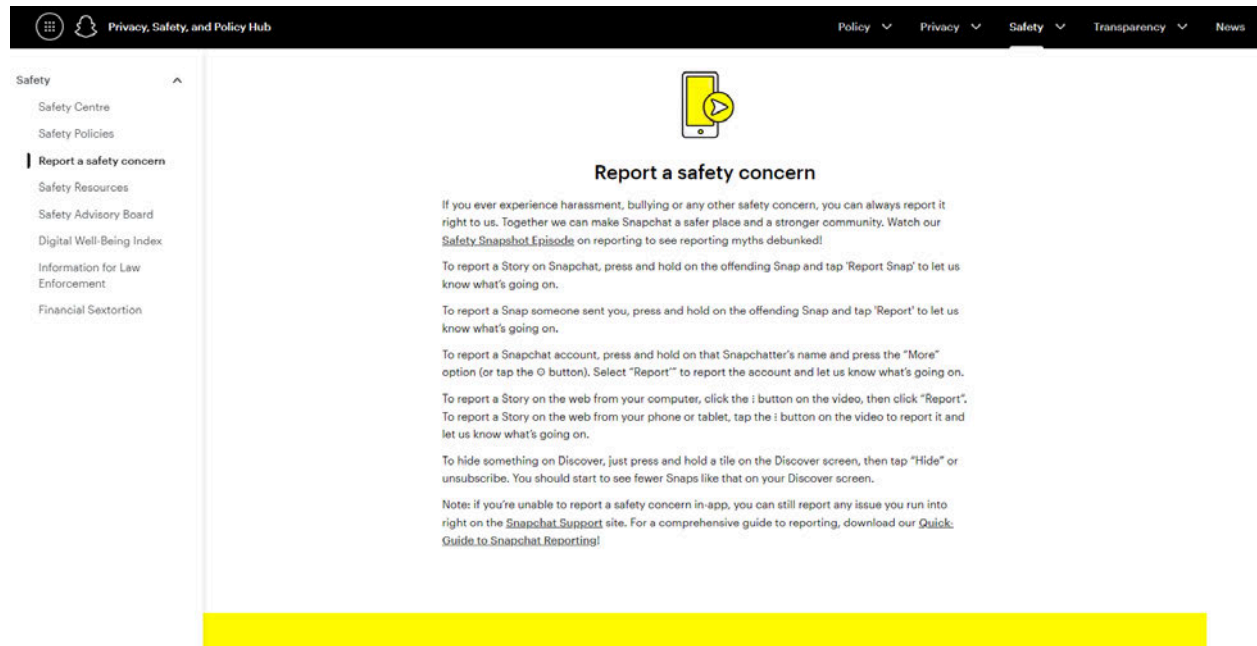
Safety Center

Our Safety Center provides an overview of our Safety resources, including tips on how to report content, the acknowledgment that safety is a shared responsibility, as well information on our Trusted Flagger Program, [Safety Advisory Board](#), [Digital Well-Being Index](#) and more. Again, the goal here is to provide information in a way that is easy to navigate and process. Since our 2023 Report, we have included direct links to our existing pages for the Safety Advisory Board, Digital Well-Being Index, our Safety Policies, Information for Parents and Information for Law Enforcement, as well as a new support page on Financial Sextortion, in the navigation bar of the Safety Center.

We have dedicated [a page](#) on our Safety Center to reporting. Our community, trusted flaggers, and other stakeholders play a vital role in the safety of our platform. A primary way they do this is



by reporting content and accounts to us when they have a safety concern. That's why we think it's crucial to raise as much awareness as possible about reporting. The dedicated page summarizes the various ways users can report content and accounts, and provides additional resources on how to report (e.g. a hyperlink to our [Safety Snapshot](#) episode on reporting). The page also links to our [Reporting Quick Guide](#) and contains a hyperlink to our web [reporting form](#).



Another important component of the Safety Center is our [Safety Resources and Support page](#). The goal of this page is to provide users with additional resources, such as a hyperlink to [MindUp](#), information about our [Here For You](#) tool, and country specific information. Since our 2023 Report, we have also included additional resources including:

- [a page](#) dedicated to explaining Financial Sextortion; and
- [a page](#) dedicated to sexual risks and harms, in an effort to support those in distress.



Privacy, Safety, and Policy Hub

Policy

Privacy

Safety

Transparency

News

Safety

Safety Centre

Safety Policies

Report a safety concern

Safety Resources

Safety Advisory Board

Digital Well-Being Index

Information for Law Enforcement

Financial Sextortion

Safety resources and support

We work with industry experts and non-governmental agencies to provide resources and support to Snapchatters in need. Here are some resources that can help if you or someone you know needs support or just wants to chat!

You can also explore our [Here For You](#) search tool which shows resources from expert localised partners when you search for certain topics related to mental health, anxiety, depression, stress, suicidal thoughts, grief and bullying.

We have also developed a page dedicated to sexual risks and harms, in an effort to support those in distress. There, you can find a list of global support resources.

Global

MindUp (global): main offices in the US, UK and CA) MindUp supports children ages 3 to 14 by providing them with the tools and knowledge to manage stress and thrive in school all while maintaining optimism, resilience and compassion.

Resources for Northern America

United States (US)

Canada (CA)

Resources for Europe

Austria (AT)

Belgium (BE)

MindUp is a non-profit organization that supports children ages 3 to 14 by providing them with the tools and knowledge to manage stress and thrive in school all while maintaining optimism, resilience, and compassion. This is to assist our Teens users by providing them with MindUp resources relevant to them.

Our Here for You search tool, which is accessible within the Snapchat app, shows resources from expert localized partners when users search for certain topics related to mental health, anxiety, depression, stress, suicidal thoughts, grief and bullying.

Our country-specific resources provide users with additional information about resources that are available to them in their country, such as children's helplines, suicide prevention hotlines, and more. See for example the below, for France:



Estonia (EE) 🇪🇪	▼
Finland (FI) 🇫🇮	▼
France (FR) 🇫🇷	^
<u>E-Enfance</u> Call 3018 The new national number against digital violence, free for children and adolescents facing problems related to their digital use-- 100% anonymous free and confidential.	
<u>Suicide Écoute</u> Call 01 45 39 40 00 Suicide Ecoute helps those who are thinking about ending their lives or have decided to do so. Suicide Ecoute allows everyone, in complete anonymity, to express their suffering.	
<u>SOS Suicide Phénix</u> Call 01 40 44 46 45 The SOS Suicide Phoenix France Federation aims to PREVENTION of suicide and PROMOTION of preventive actions in complementarity with the actors of the medico-social field.	
Germany (DE) 🇩🇪	▼
Greece (GR) 🇬🇷	▼

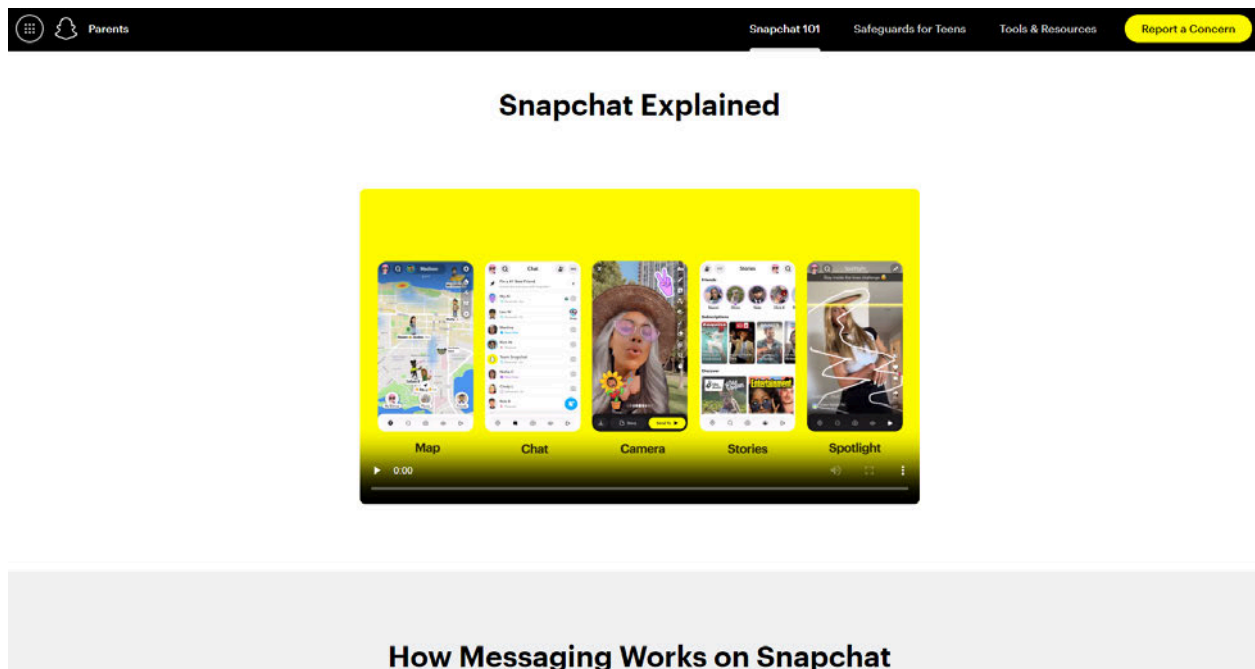
Parents

We have a dedicated [Family Center](#), our in-app tool for parents and caregivers. To help develop Family Center, we worked with families to understand the needs of both parents and teens, knowing that everyone's approach to parenting and privacy is different. We also consulted with experts in online safety and wellbeing to incorporate their feedback and insights. Our goal was to create a set of tools designed to reflect the dynamics of real-world relationships and foster collaboration and trust between parents and teens. In the coming weeks, we will add a new feature that will allow parents to easily view new friends their teens have added.

In September 2023, we launched a dedicated microsite: parents.snapchat.com to provide even more information for parents. We recognize that not all caregivers use Snapchat. Their lack of familiarity may create questions, and may also make it difficult for them to have a conversation with younger users. To address this concern, the dedicated microsite contains: an updated Snapchat 101, a specific page on Safeguards for Teens, resources about Snapchat's Family Center, Snap Map Safety, and other resources.



The Snapchat 101 page incorporates our previous '[Parent's Guide to Snapchat](#)' but lays it out in an accessible manner (including a short video).



The Safeguards for Teens page summarises the key protections for teens.

Transparency Center

Our [Transparency Center](#) provides additional transparency resources to our users and to the public at large, including our Community Guidelines (see [Terms Section](#)), Transparency Reports and EU-specific information required under the DSA.

On our [EU Transparency Page](#), we publish EU-specific information required under the DSA, including the number of Average Monthly Active Recipients of our Snapchat app in the EU, and information about our legal representative in the EU, how EU law enforcement agencies can submit requests to snapchat, and the regulatory authorities that regulate us under the DSA.

Since 2015, we have also been publishing Transparency Reports twice a year, to provide insight into Snap's safety efforts and the nature and volume of content reported on our platform, as well as the manner in which we enforce our Community Guidelines and respond to law enforcement requests, as well as how we respond to notices of copyright and trademark infringements. We are committed to continuing to make these reports more comprehensive and informative to the many stakeholders who care deeply about our content moderation and law enforcement practices, as well as the well-being of our community. As part of our DSA compliance, Snap will be adding new



metrics and information to its Transparency Report. Copies of our most recent and previous Transparency reports can be found on our [Transparency Report](#) and [Previous Reports](#) webpages.

Snap Values

Policy ▾ Privacy ▾ Safety ▾ **Transparency ▾** News

Transparency ^

Transparency Report

About

Glossary

Previous Reports

European Union

Ads Gallery

Australia

California

India

Researchers

Transparency Report

January 1, 2024 – June 30, 2024

Released:

December 05, 2024

Updated:

December 05, 2024

We publish this transparency report twice a year to provide insight into Snap's safety efforts. We are committed to these efforts and continually strive to make these reports more comprehensive and informative for the many stakeholders who care deeply about our content moderation, law enforcement practices, and the safety and wellbeing of the Snapchat community.

This Transparency Report covers the first half of 2024 (January 1 – June 30). As with our previous reports, we share data about the global volume of in-app content and account-level reports our Trust & Safety teams received and enforced across specific categories of Community Guidelines violations; how we responded to requests from law enforcement and governments; and how we responded to notices of copyright and trademark infringement. We also provide country-specific insights in the files linked at the bottom of this page.

News Page

Snap also frequently publishes Privacy and Safety related information on the Hub's [News](#) webpage. The purpose of these news articles is to inform the general public about recent developments on issues relevant to privacy, safety and transparency on Snapchat. For example, a recent article introducing Snap's Inaugural Council for Digital Well-Being, or an [article](#) informing the public on Snap's approach to keeping its community safe during the 2024 Paris Olympics.



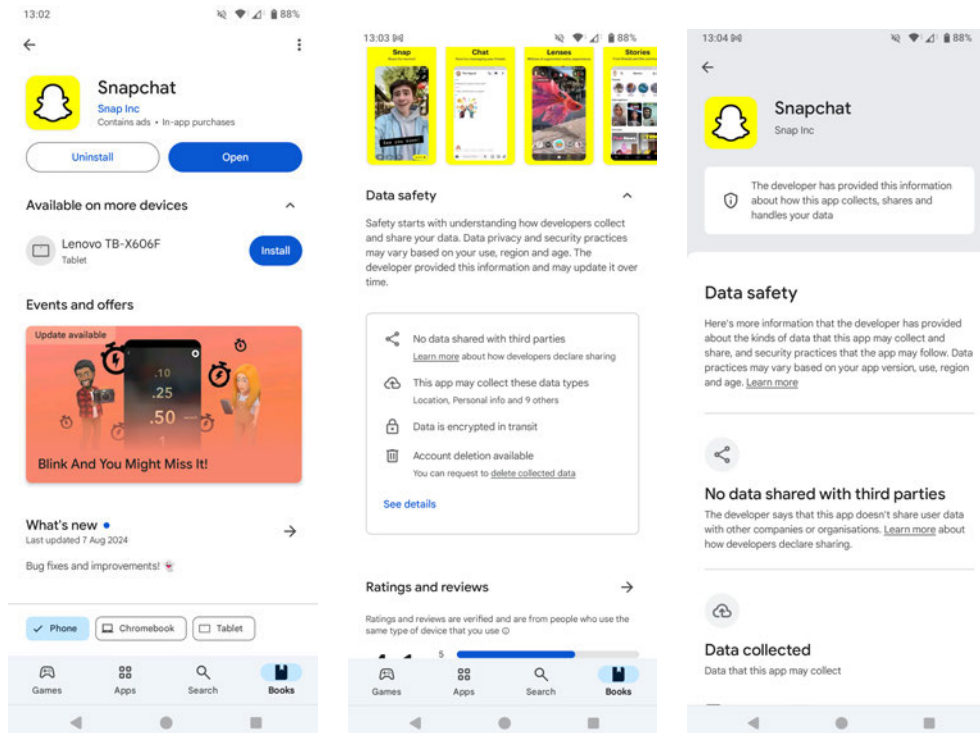
Introducing Snap's Inaugural Council for Digital Well-Being

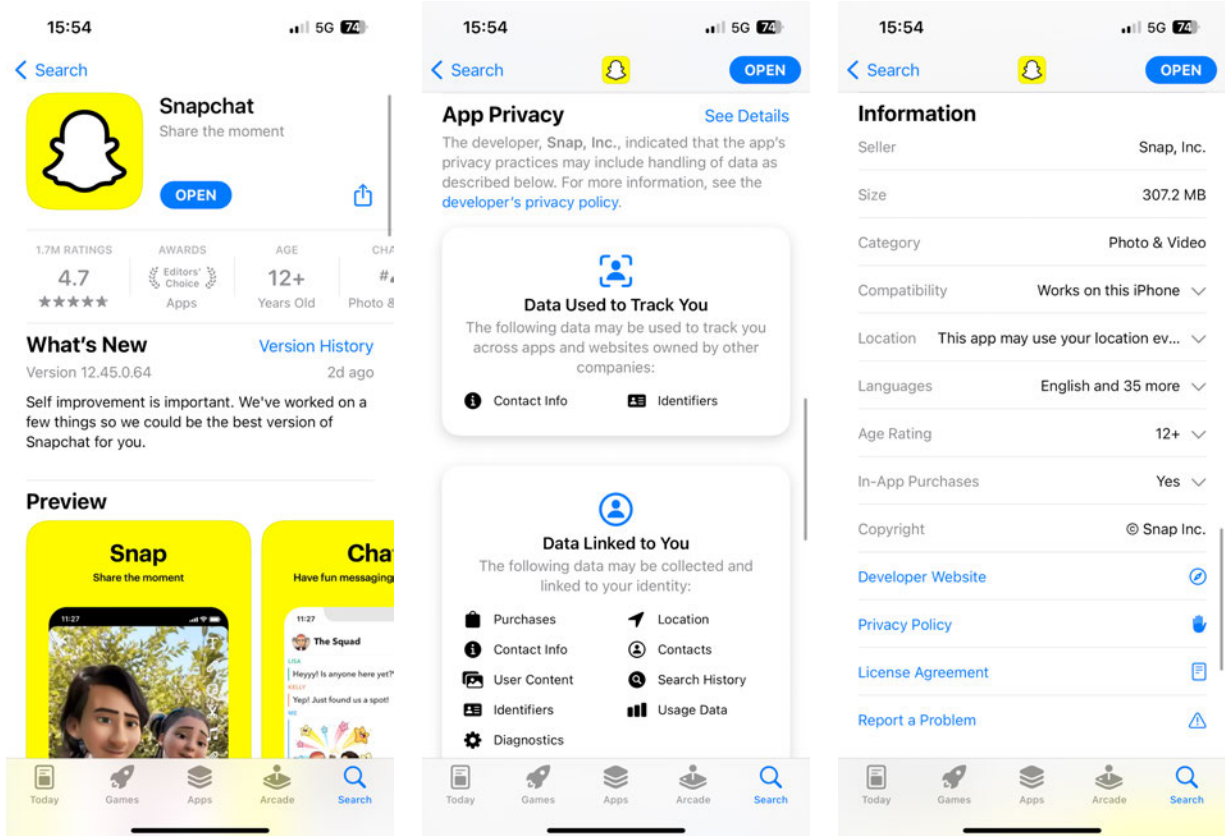
August 8, 2024



5.3.2 Information provided in app stores

Prior to downloading Snapchat, we provide users with information about the Snapchat app in the Apple and Google Play Stores. This includes general information on the functionalities of the app, as well as information on our data collection practices, and links to our website, Privacy Policy and Terms. This way users are able to get a better understanding of the application ahead of using Snapchat.





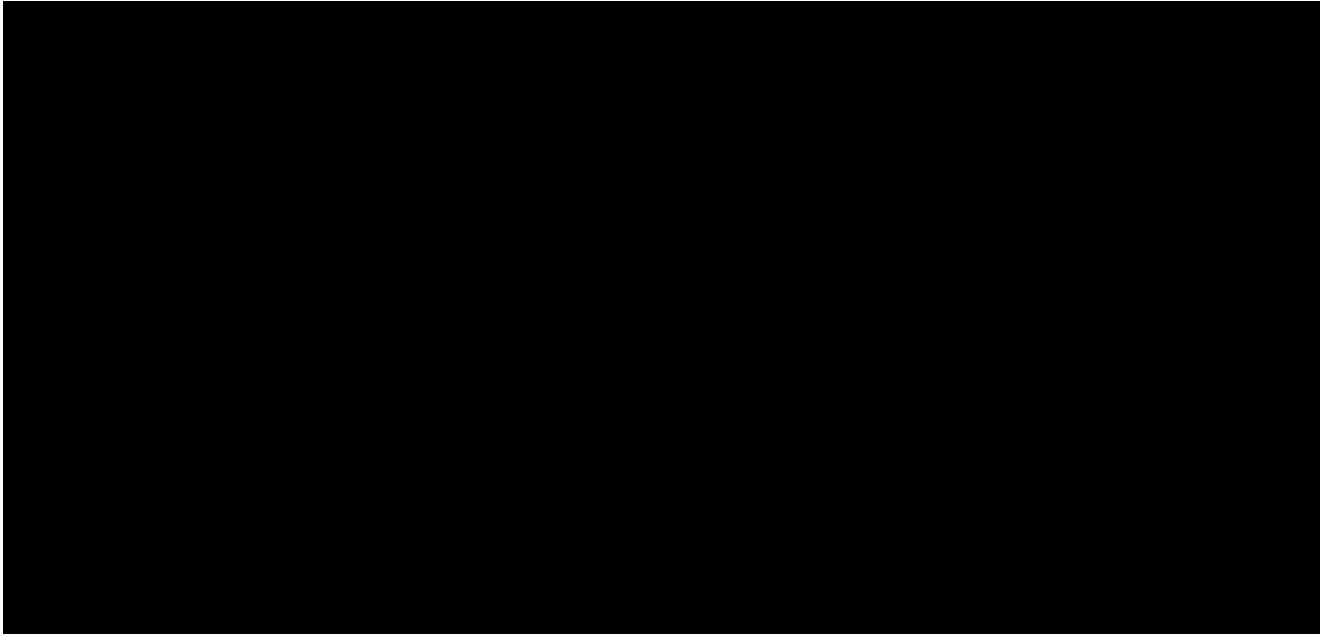
5.3.3 Information we provide in our application

Once users have downloaded Snapchat, they are invited to create an account. At Snap, our philosophy is to provide timely notifications and generate awareness at points in time where we believe they will be most effective. We provide a high level overview of our onboarding process and highlight examples of our “just-in-time” in-app notifications in this section.

Onboarding process

Step 1.

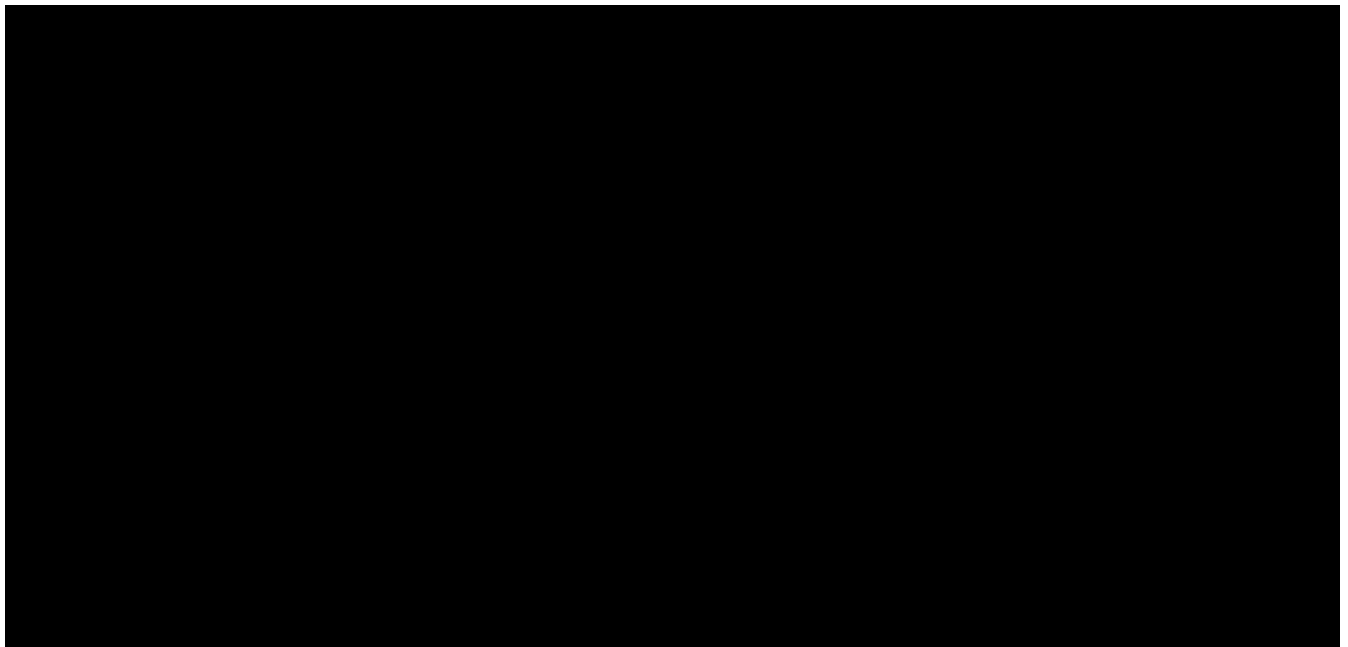
When users open Snapchat, they are invited to log in (if they have an existing account) or create a new account. The first set of notices users receive relates to notification settings, and the ability to connect their device’s contacts to find friends.



Both steps are optional. The reason we prompt users to turn on notifications is that Snap is primarily a messaging service and notifications provide an essential utility when using the service. Snap is intended for real friends and family, and requires users to accept friend requests or be already existing contacts before they can start communicating with each other. Typically, users already have their close friends and family stored in their device contact book, so the “Find Your Friends” prompt is intended to make it easier for users to send friend requests to other users and to communicate with one another.

Step 2.

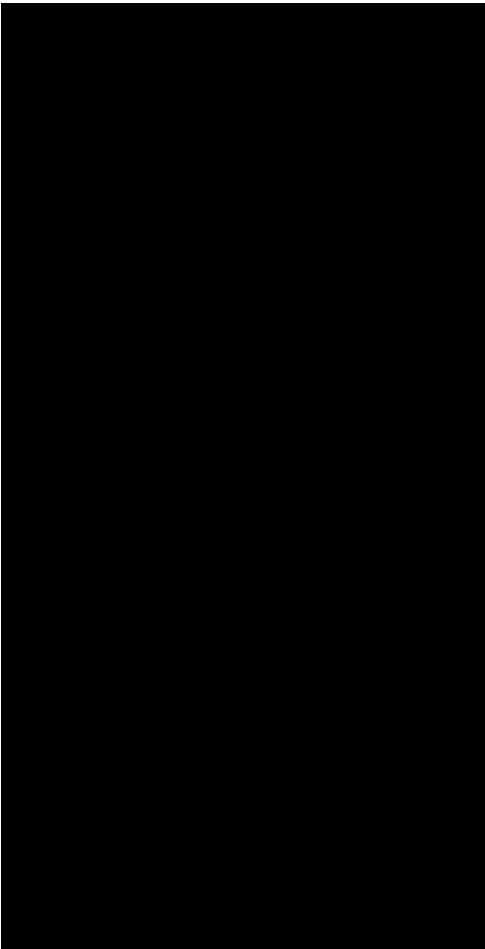
The second step of the onboarding flow requests basic account information such as the user’s first name, last name (optional), birthday and username.





When asking for their birthday, we show users a neutral age screen, and if a user selects an age under 13, they are prevented from creating an account. We don't notify the users the reason for a failure to create an account.

We have drawn on guidance from the UN Convention on the Rights of the Child²⁰⁶ and UK Age-Appropriate Design Code²⁰⁷ to adopt a risk-based approach to age verification in our age gating process. We considered the risks of the platform as well as the rights of younger user's right to privacy, freedom to access information and freedom of expression under the Convention and balanced them against safety risks. We believe more invasive age gates come at a privacy cost for all users, and also disproportionately impact marginalized groups who may not have access to government IDs.²⁰⁸ We have supported the UK Online Safety Bill amendment to require App Stores to play a more active role in sharing age signals to all app stores. We believe this is the better upstream solution to address any systemic risks associated with underage users accessing platforms.



²⁰⁶ UN OHCHR, *Convention on the Rights of the Child*, [url](#).

²⁰⁷ UK Information Commissioner's Office, Introduction to the Children's code, [url](#).

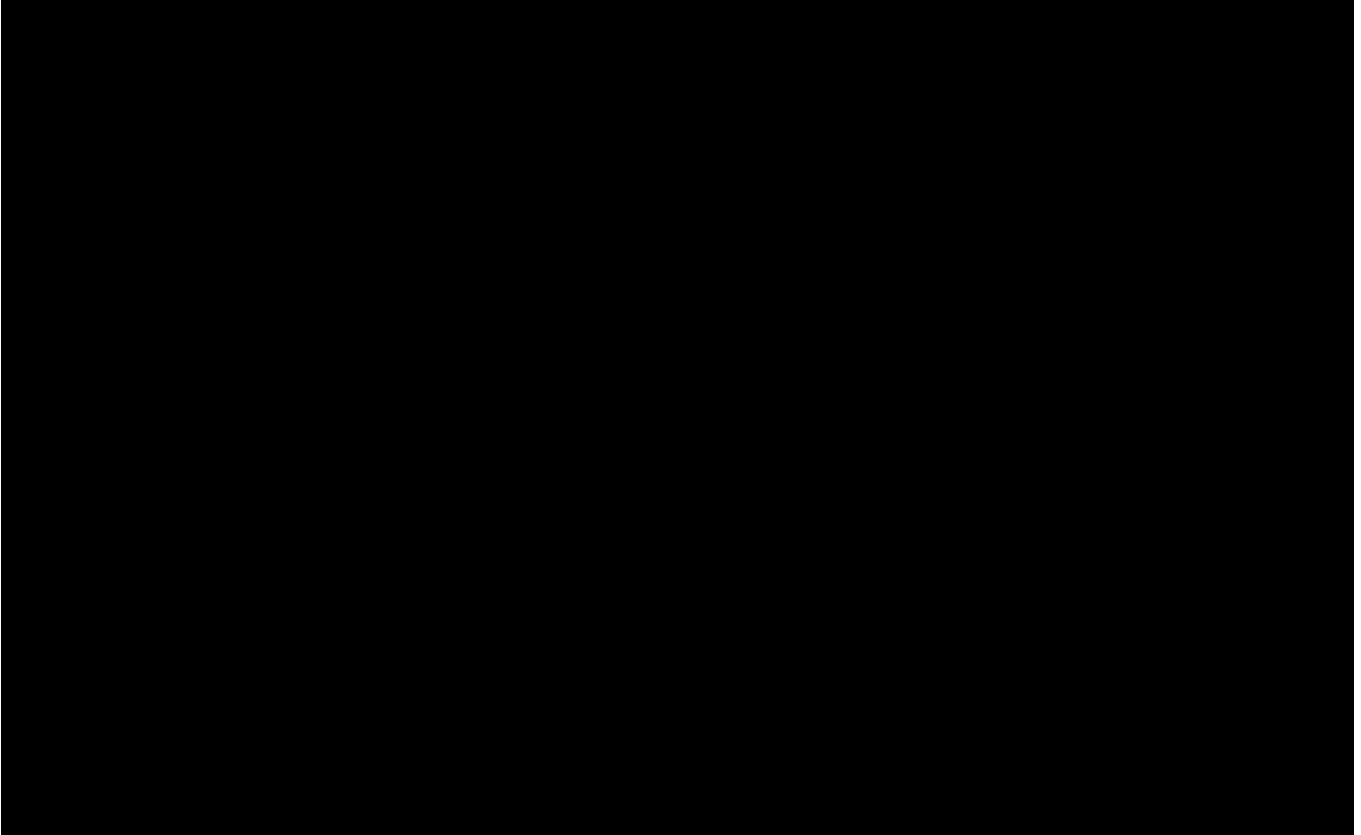
²⁰⁸ See for example the report on age verification issued by the Australian eSafety Commissioner, [url](#).



If a user has inputted an age of 13 or older, they are prompted to provide a username. We check usernames against our Abusive Language Detection (ALD) models. If users type in an abusive username (i.e., one that does not comport with our Terms), they are prevented from creating an account and are asked to enter a username that adheres to our Terms.

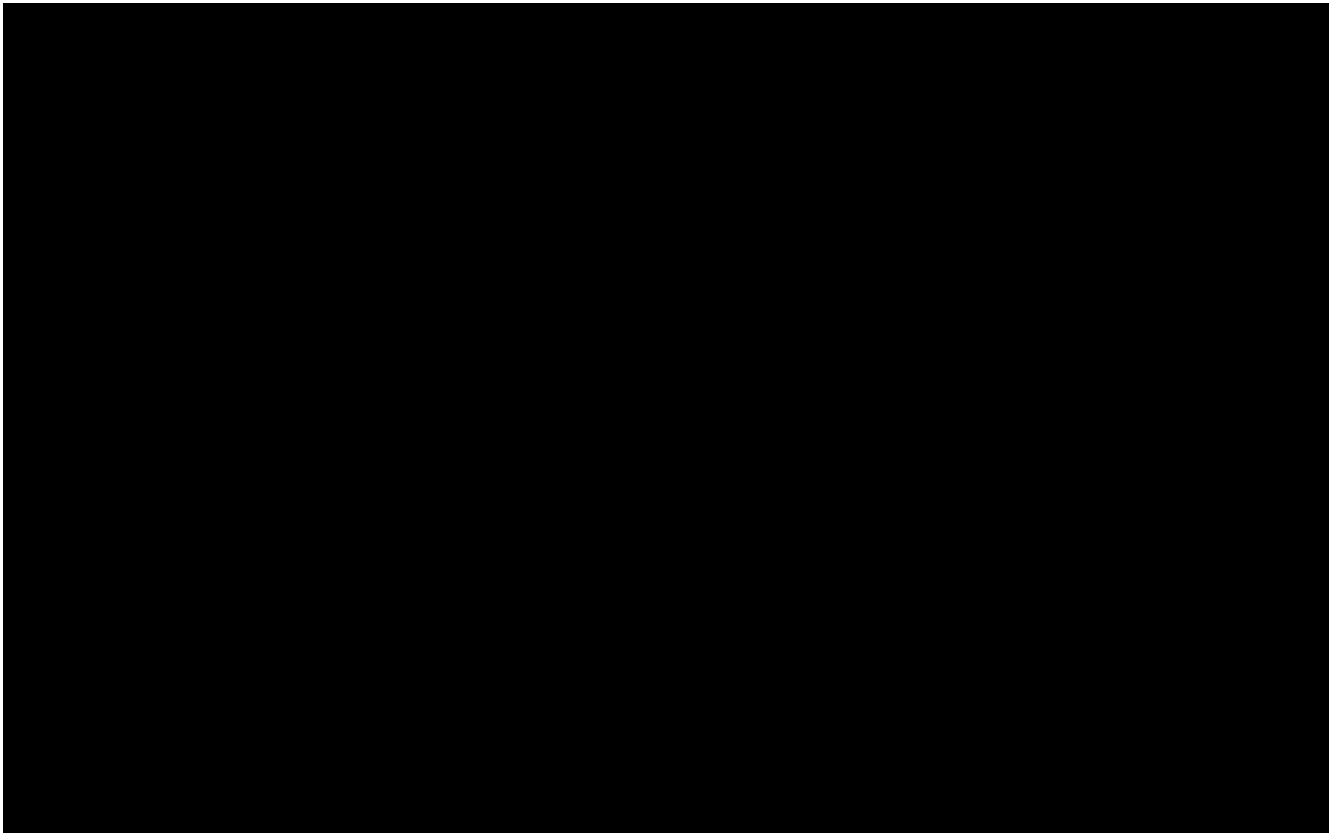
Step 3.

The third step of the onboarding process is focussed on password creation and providing a phone number and / or an email address. These are standard steps to improve account security and provide Snap the ability to communicate with users.



Step 4.

Lastly, we offer users the ability to start finding friends on Snapchat, and the option to create a Bitmoji. Snapchat shows Bitmojis instead of profile pictures. Bitmojis protect the identity of users, and prevent abuse from predators who may use profile pictures as signals to reach out to their target victims.



Just-in-time notifications

Once a user has created an account, we create awareness at a feature-specific level, typically using Just-in-Time notices or “JITs”. We conduct user research and sentiment studies, and feedback we receive from users is that JITs or icons are more effective to inform users than long text. Below we provide some examples of JITs that create feature-specific awareness.

Snap Map

Snapchat’s access to the users’ precise location, including for features like Snap Map, is off by default. Users choose who can see their location.

Since our 2023 Report, we have simplified location sharing on Snap Map. Location sharing still requires a two step permission to enable location sharing:

1. Users will choose location permissions in device settings (e.g. for iOS these are “Never”, “Ask Next Time or When I Share”, “While Using the App”, “Always”).
2. Users must also select sharing with “My friends”, “My Friends, Except” or “Only these Friends”. Users cannot share location with non-Friends.

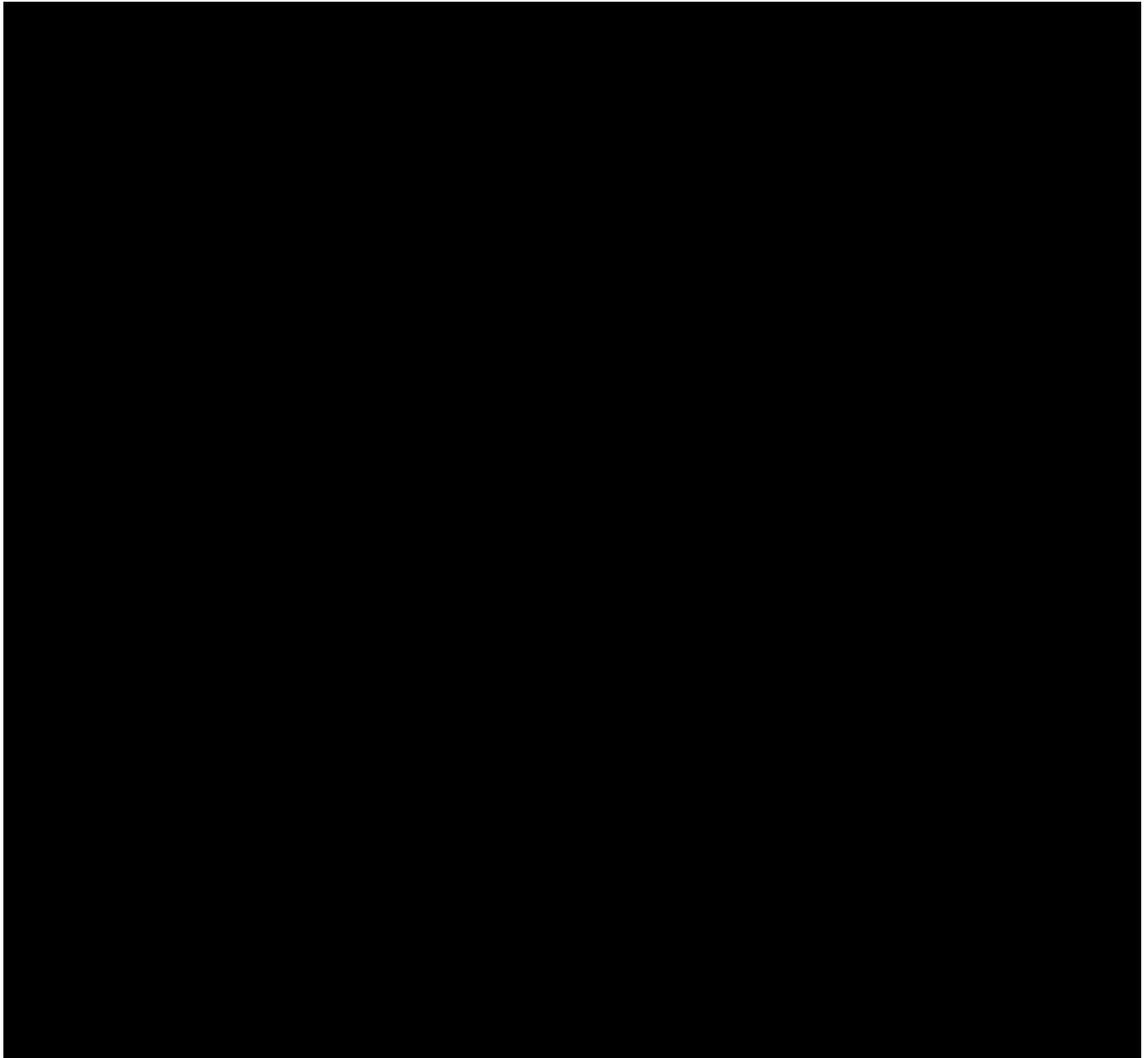
The selected Friends will see “live location” if the user chose “Always” in device settings or “last active” if the user chose “While Using the App” in device settings. Users can still decide to turn

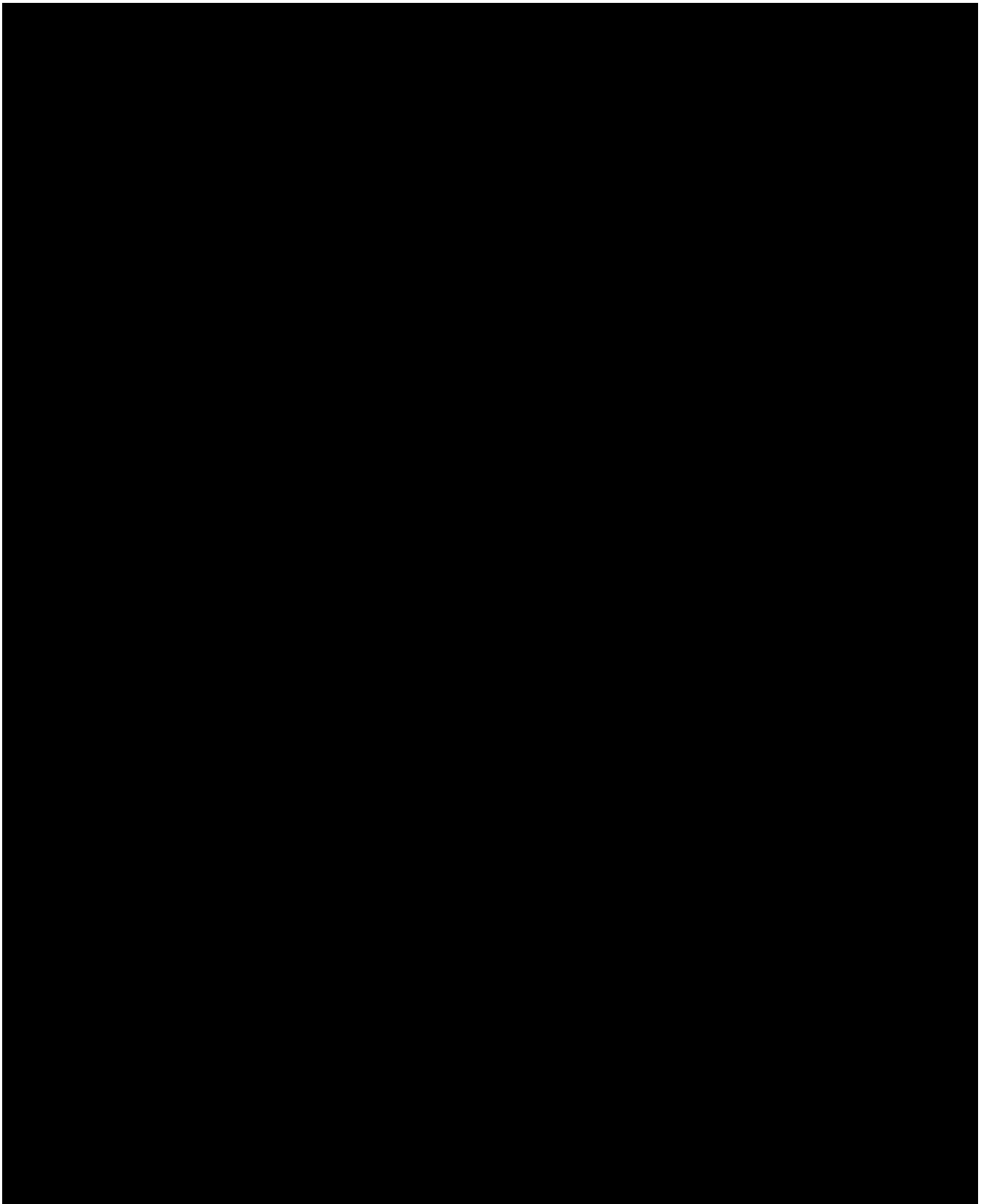


on [Ghost Mode](#) at any time when they want to disable location sharing (either for a specific time or until the user wishes to turn off Ghost Mode).

[Redacted line]

[Redacted line]:

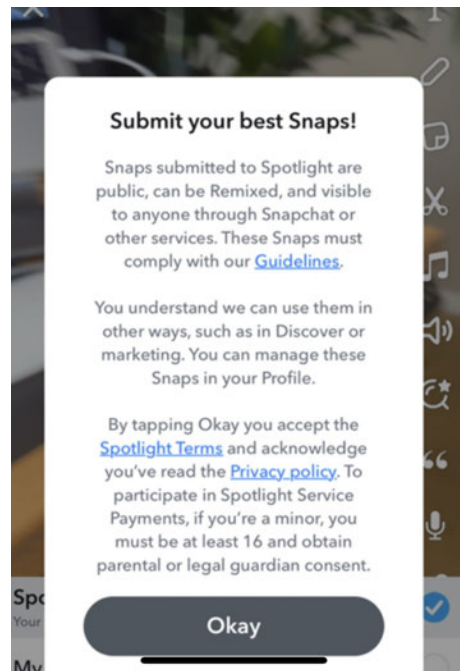




Spotlight



Before a user submits a Snap to Spotlight they are presented with a JIT informing them that Spotlight submissions are public. This is to create awareness that Spotlight is different from My Story submissions, which can be shared with friends only, unless the user actively chooses to share them with “Everyone”.

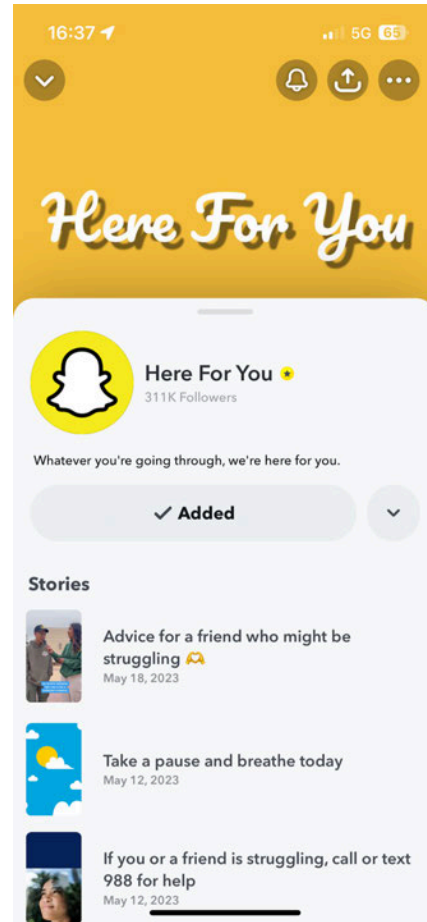
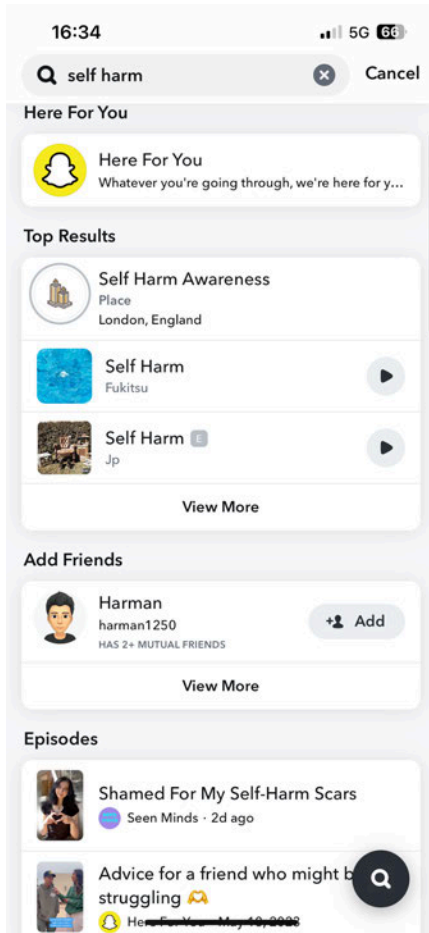


Thematic awareness and notices

Across Snapchat, we offer a number of resources to users to raise awareness on safety topics and protect them. For example:

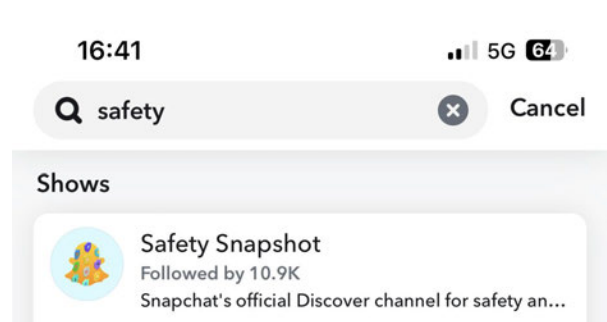
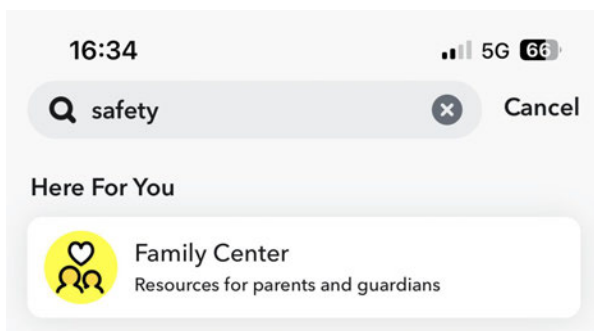
1. Here for you

If a user types in “selfharm” or related terms in our Search functionality, we try to prominently show them relevant ‘Here For You’ resources among the search results.



2. Safety

Search terms like “safety” will direct users to our relevant Here For You resources, such as information on our Family Center, and to our Safety Snapshots, our official channel for safety and privacy tips and tricks.

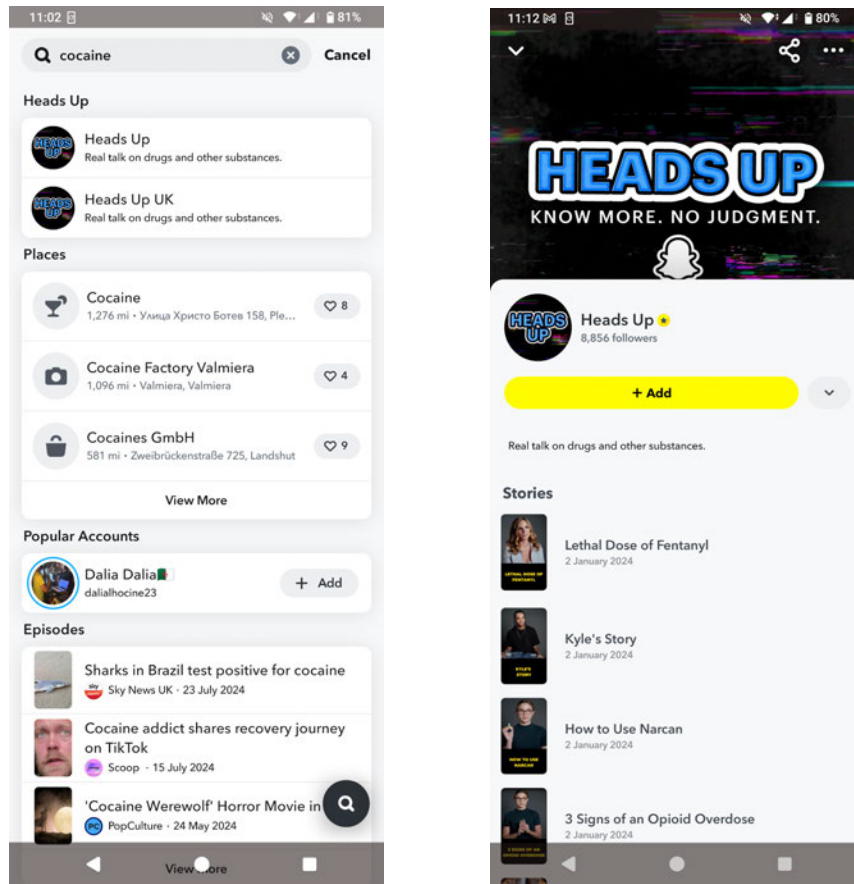


3. Heads up

If a user types in terms related to illicit drugs in our Search functionality, we try to prominently show them our ‘Heads Up’ resources among the search results. Heads up is



our in-app tool that surfaces educational content from experts to Snapchatters if they try to search for drug-related content. Our expert partners include the Centers for Disease Control and Prevention (CDC), the Substance Abuse and Mental Health Services Administration (SAMHSA), Community Anti-Drug Coalitions of America (CADCA), Shatterproof, Truth Initiative, and the SAFE Project.



We also run campaigns on Snapchat to raise awareness about certain themes. For example, on [Global Data Privacy Day](#) 2024, we informed the general public about our new Privacy Policy, announced the updated [parents guide to Snapchat](#), and launched a dedicated page on [Privacy through Security](#), and relaunched our interactive Lenses with tips on how to stay safe online.

Similarly, on [Safer Internet Day](#) 2024, we raised awareness around parents' options to participate and monitor their child's online activities through Snapchat's Family Center. We publish updates on [efforts](#) and [campaigns](#) to raise awareness around the dangers of fentanyl, and continue to partner with organizations like [Song For Charlie](#) to combat illicit drugs on Snapchat.



5.3.4 Languages

As explained above, our [Terms of Service](#) have been translated into all official languages of the European Union as explicitly required by the Digital Services Act. However, Snapchat itself is only available in certain official languages of the European Union, not all. As a result, our in-app and publicly accessible information is also only available in certain official languages of the European Union. [REDACTED]

[REDACTED] we consider it reasonable and proportionate and effective to offer our mitigation measures in the same languages as Snapchat as we anticipate recipients only using Snapchat if they understand one of the available languages.

5.3.5 Conclusion

Snap offers a wide range of in-app and publicly accessible information to raise awareness around privacy, safety and security to its community and external stakeholders. Our approach is that these tools should be easily accessible, easy to use and understand, and provided in a timely manner. [REDACTED]

[REDACTED] we believe that the awareness measures we have in place provide reasonable, proportionate and effective mitigations.

As explained in Section 4, we have concluded that Snap’s awareness raising information, in combination with the other mitigations explained in this Section 5, is a reasonable, proportionate and effective mitigation measure for the risks presented by Snapchat’s in-scope services.

5.4 Content Moderation

5.4.1 Approach

Across Snapchat, we’re committed to advancing safety while respecting the privacy and freedom of expression of our community. We take a balanced, risk-based approach to combating harms — combining transparent content moderation practices, consistent and equitable enforcement, and clear communication to hold ourselves accountable for applying our policies fairly.

Safety and privacy is a priority across Snapchat, and we use a combination of in-app reporting, automation tools, and human review to combat harms on the platform. All content must adhere to our Terms, including our [Community Guidelines](#) and [Terms of Service](#), and some content must also adhere to our [Content Guidelines for Recommendation Eligibility](#). We strive to be transparent and consistent in our practices and enforcement, while striking the right balance between privacy and safety.



Snapchat Design and Function

As a reminder, we have also designed Snapchat with privacy and safety in mind, and this design is key in helping to prevent the spread of harmful and illegal content. Snapchat does not offer an unmoderated, permanent news feed where unvetted publishers or individuals have an opportunity to broadcast hate, misinformation, or violent content to a wide audience with algorithmic amplification. We think about content on our platform in three categories:

1. “Public Content” is made available to the public and potentially accessible to anyone on Snapchat and/or the Internet. This includes Public Stories, Map Stories, Lenses, Public Profiles, advertisements and content recommended for broad distribution on Spotlight and the Discover section of the Stories tab.
2. “Limited Broadcast Content” is content broadcast to a selected audience and includes Private Stories and Community Stories.
3. “Private Content” is private messaging content sent to Friends. Private Content includes Chat and Group Messaging.

While only Public Content services are in-scope of this Report, we have also provided information relating to Limited Broadcast and Private Content services for reference.

More information about our mitigations relating to Snapchat’s design and function can be found in Section 5.1 ([Snapchat Design and Function](#)).

Community Guidelines and Terms of Service

When considering our Content Moderation approach, it is also important to bear in mind that all content everywhere on Snapchat must adhere to our [Community Guidelines](#) and [Terms of Service](#). Then, in order to be eligible for algorithmic recommendation beyond the creator’s friends or followers, content must meet the additional, higher standards described in our [Content Guidelines for Recommendation Eligibility](#). More information about our mitigations relating to Terms and user awareness can be found in the [Terms Section](#) and [Transparency Section](#).

Content Moderation

Our content moderation processes assess each piece of content against the above Terms, policies and guidelines to determine if that content is compliant. [REDACTED]

We assess content reactively (upon receipt of a report) and proactively (when flagged by our automated detection tools) using a combination of automated review tools and/or human review.

[illegible]

Enforcement



	• [Redacted]
[Redacted] [Redacted]	■ [Redacted] ■ [Redacted] ■ [Redacted] ■ [Redacted]
[Redacted]	■ [Redacted] ■ [Redacted] ■ [Redacted]
[Redacted] [Redacted]	■ [Redacted] ■ [Redacted]
[Redacted]	■ [Redacted]

[Redacted]

[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]	[Redacted]
[Redacted]	[Redacted]
[Redacted]	[Redacted]



[Redacted]

[Redacted]	[Redacted]
[Redacted]	[Redacted]
[Redacted]	[Redacted]
[Redacted]	[Redacted]
[Redacted]	[Redacted]
[Redacted]	[Redacted]
[Redacted]	[Redacted]
[Redacted]	[Redacted]
[Redacted]	[Redacted]
[Redacted]	[Redacted]
[Redacted]	[Redacted]
[Redacted]	[Redacted]

[Redacted]
[Redacted]
[Redacted]

[REDACTED]



[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

[REDACTED]

- [REDACTED]
[REDACTED]
- [REDACTED]
[REDACTED]
- [REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]



[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]



[Redacted]

[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

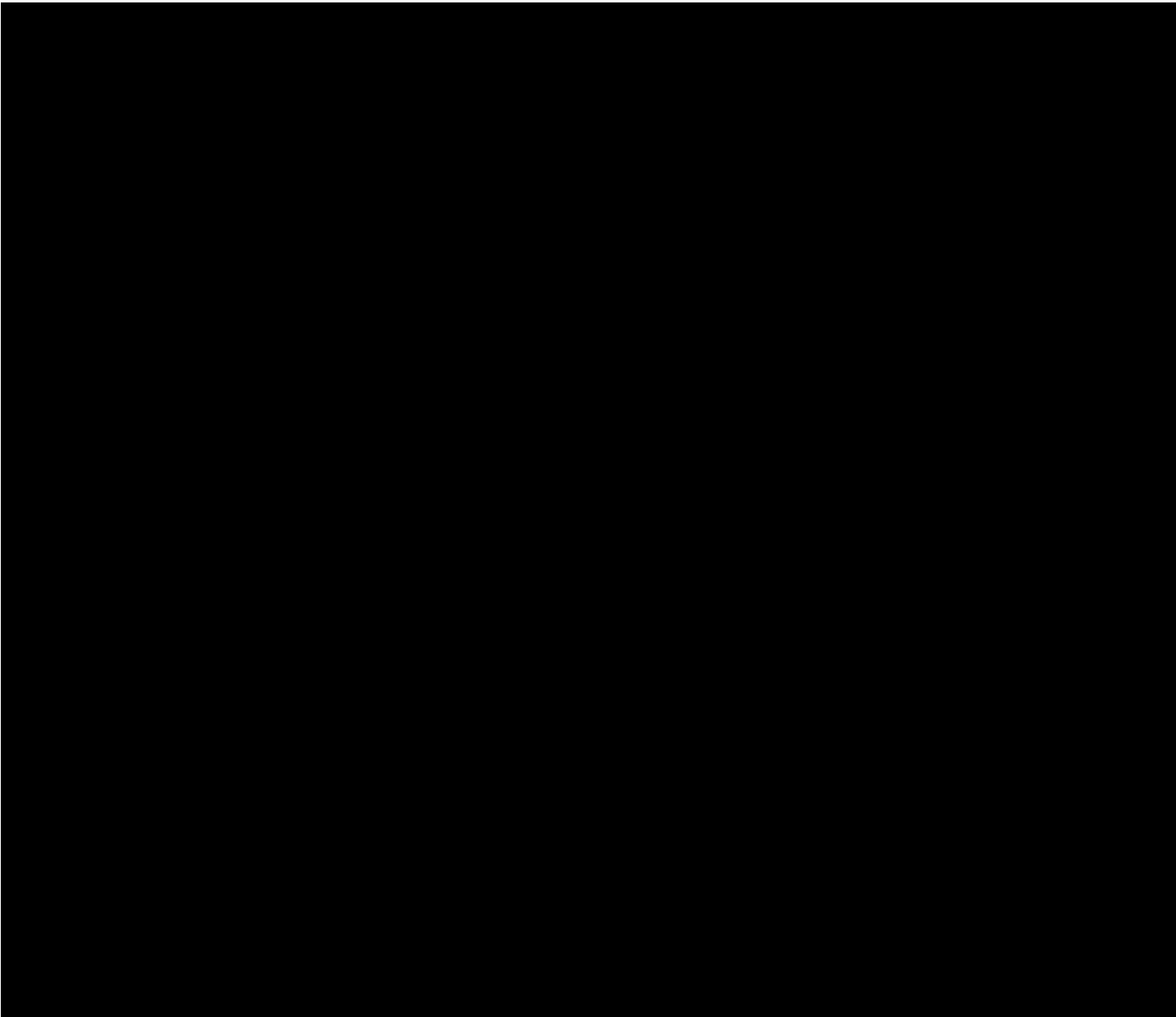
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted] [Redacted] [Redacted] [Redacted] [Redacted] [Redacted] [Redacted] [Redacted] [Redacted] [Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]



[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]



[Redacted]

[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]

- [Redacted]
[Redacted]
- [Redacted]
[Redacted]
- [Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]
- [Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]



[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

5.4.3 Broadcast Content - Proactive Moderation

We use a combination of automated tools and human review to proactively moderate public broadcast content across Snapchat, i.e., content recommended for broad distribution on Snapchat, such as content in Spotlight, Discover, Lenses and Advertisements and non-public broadcast content such as stories. Proactive moderation considers compliance with our [Content Guidelines for Recommendations Eligibility](#).

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]



[REDACTED]

[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]



[REDACTED]

[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

5.4.4 Product-Specific Moderation

Snap products leverage varied processes to determine whether content is eligible for distribution within each unique product. Products deploy varying business rules, which result in automated content moderation, based on the nature of the content presented to users on each product.

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



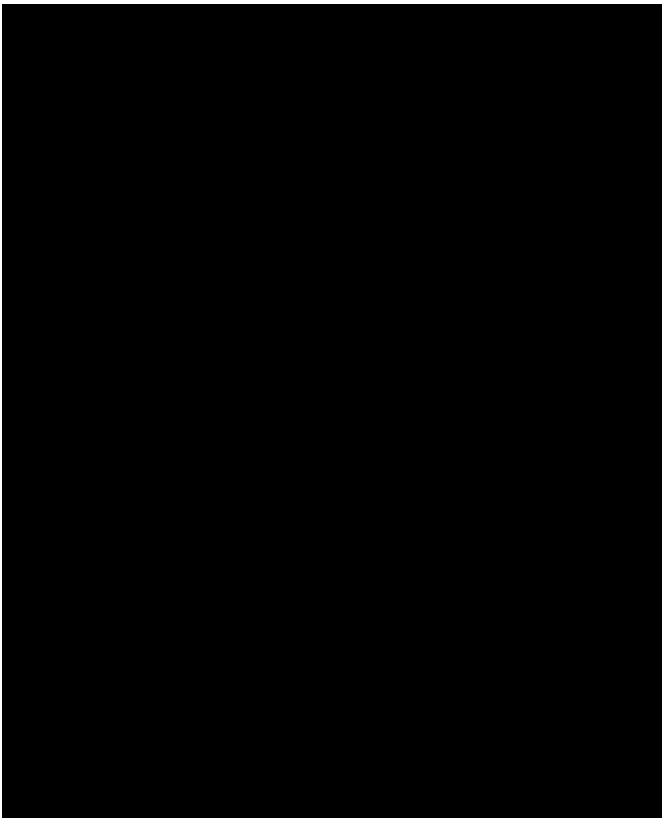
[Redacted text line]

[Redacted text block consisting of four lines]

[Redacted text block consisting of six lines]

[Redacted text block consisting of six lines]

[Redacted text block consisting of three lines]





[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]



[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]



[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

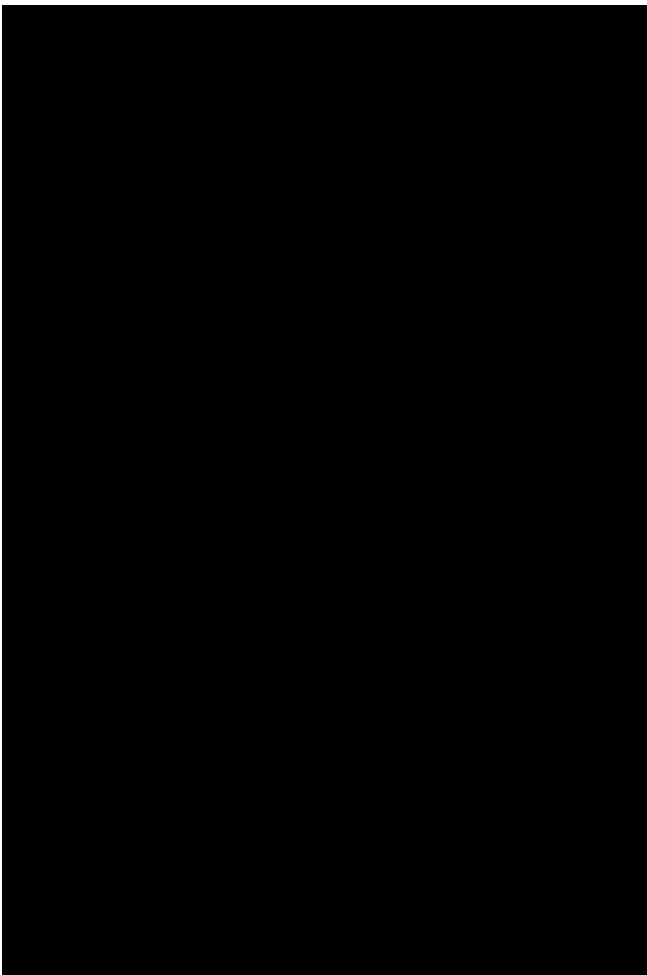
[Redacted text]

[Redacted text]

[Redacted text]



[Redacted text line]



[Redacted text block consisting of seven lines]

[Redacted text block consisting of three lines]

[Redacted text block consisting of four lines]



[Redacted text line]

[Redacted text block containing 11 lines]

[Redacted text block containing 3 lines]

[Redacted text block containing 6 lines]

[Redacted text block containing 5 lines]

[Redacted text block containing 7 lines]

[Redacted text block containing 2 lines]



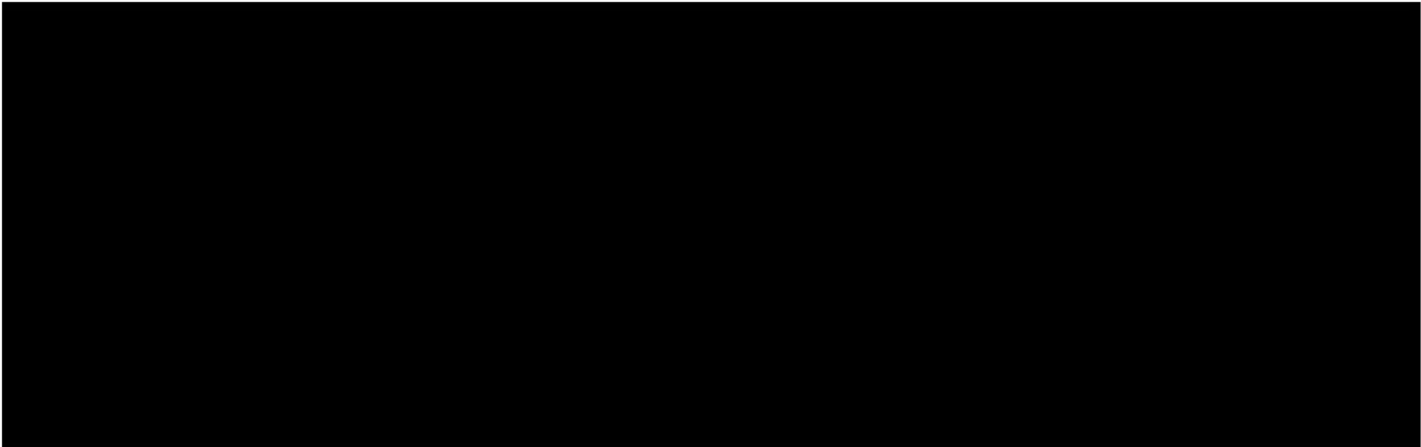
5.4.5 Reactive Moderation (Reporting)

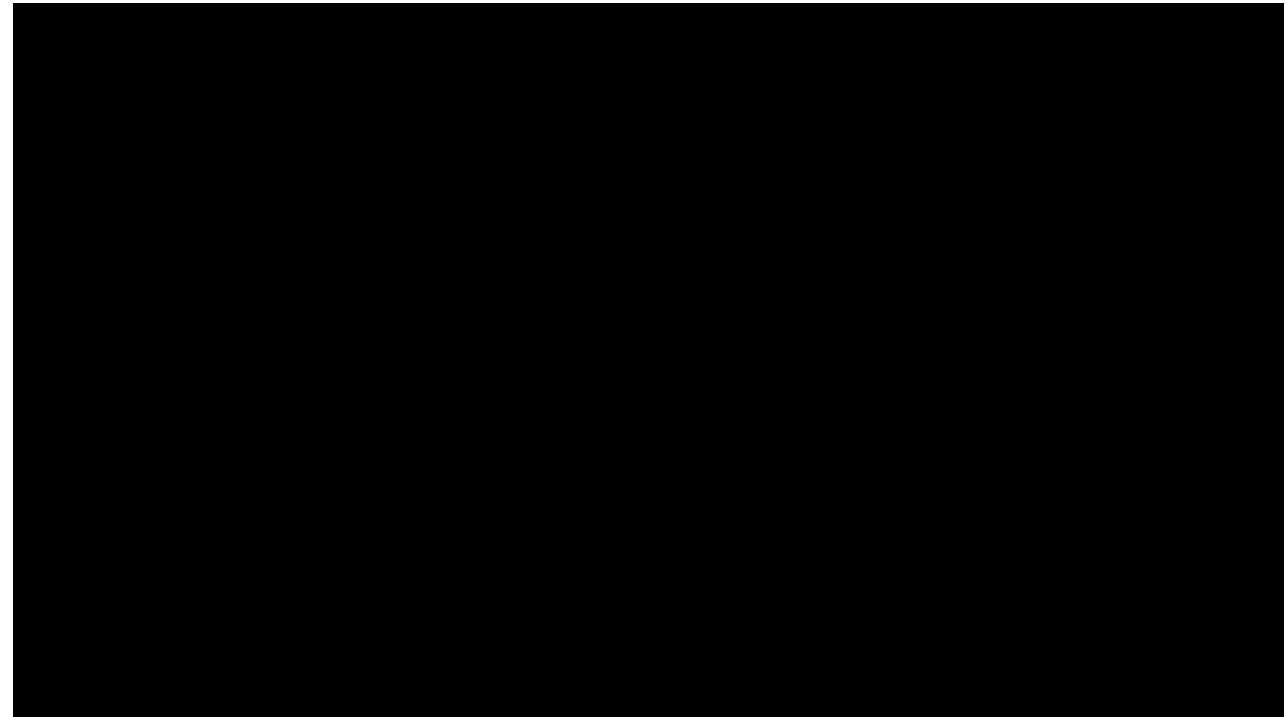
Our reactive moderation policies and processes complement our proactive moderation efforts. Across all of our product surfaces, individuals and entities in the EU can report accounts and content they believe violate our Terms, which (among other things) prohibit the dissemination of harmful and illegal content. We make it easy for users and non-users to submit a confidential report directly to our Safety & Moderation team, who are trained to evaluate the report, seeking legal input where needed.

Once we receive a report, we acknowledge receipt of the report and we review it. If the account or content is violative, we then take appropriate enforcement action and notify the relevant account holder, and, if the user's account was locked as part of Snap's enforcement, we provide them an opportunity to appeal our decision. If we find no violation, we take no action on the account or content. In either scenario, we notify the reporting party of the outcome, and provide the relevant user an opportunity to appeal our decision.

Reactive Moderation Process

On our [Support Site](#), we provide guidance on how individuals and entities in the EU can report Snapchat accounts and content directly within the app or through web forms that are easily accessible on our website.



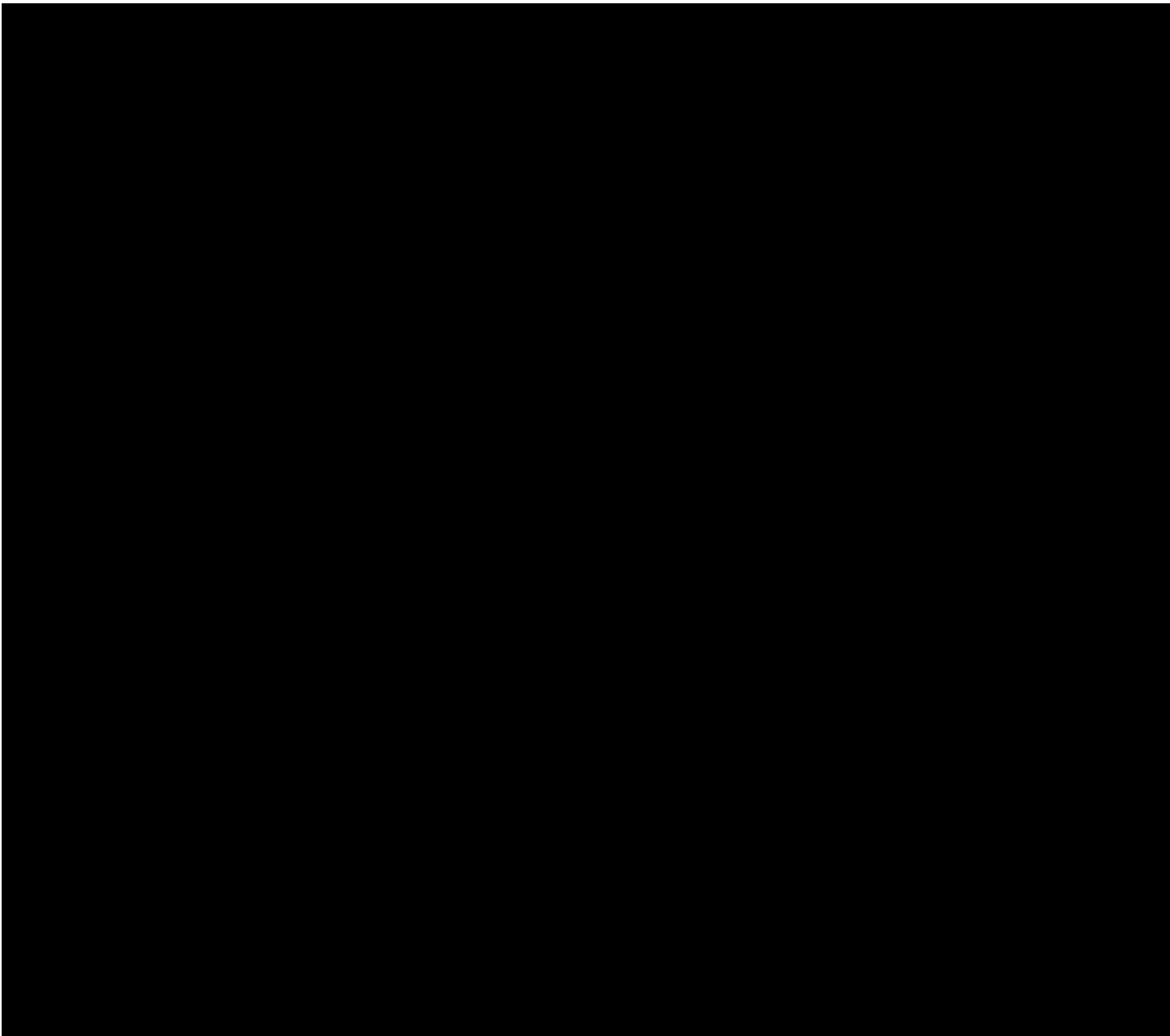


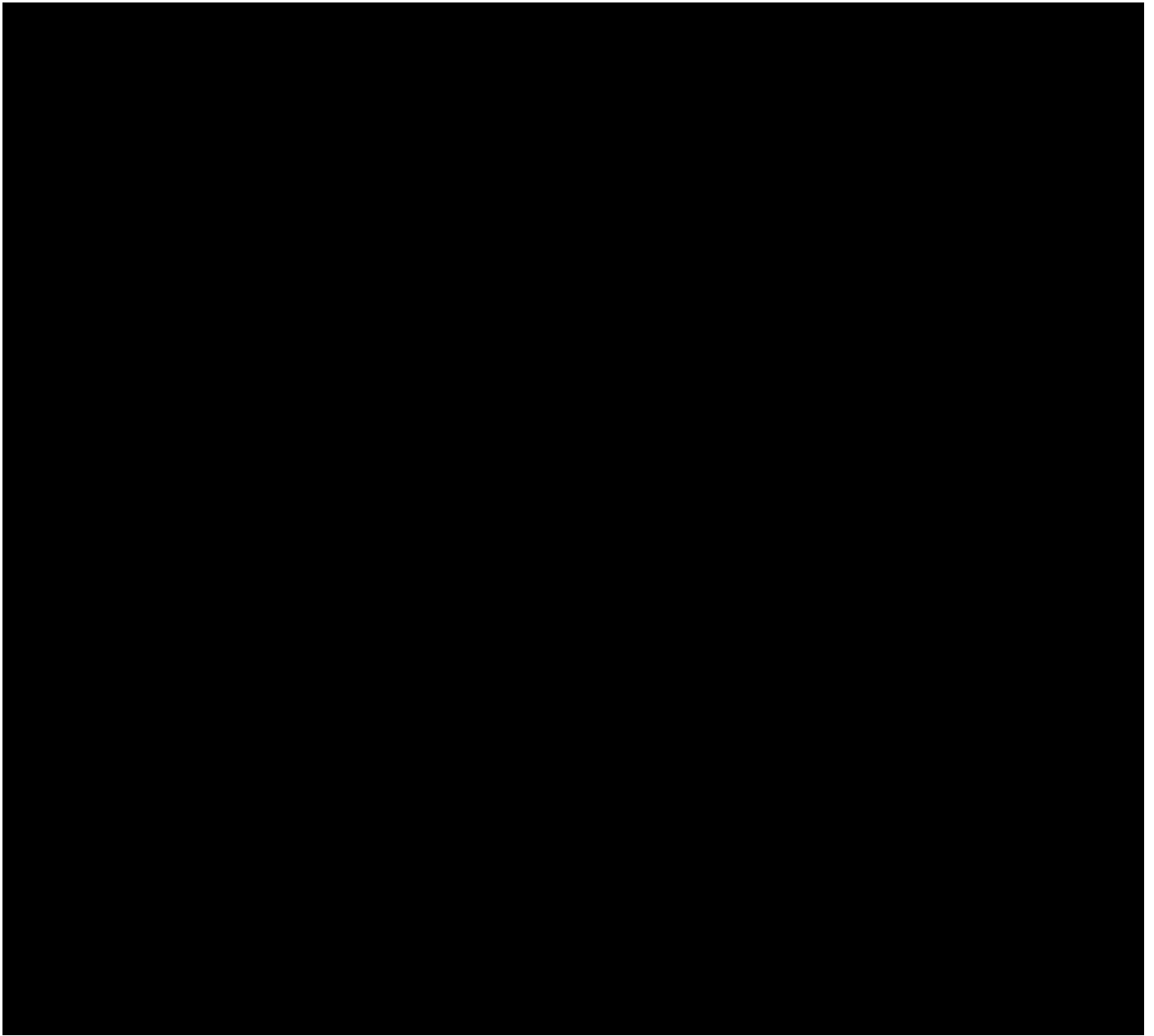
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]







[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

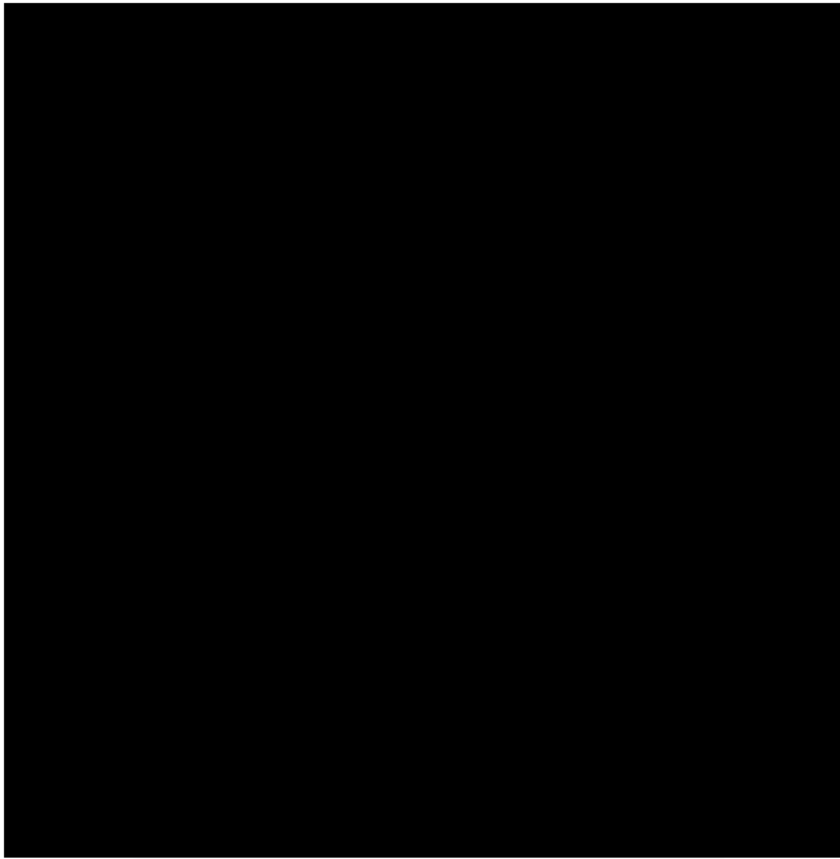
[Redacted text]

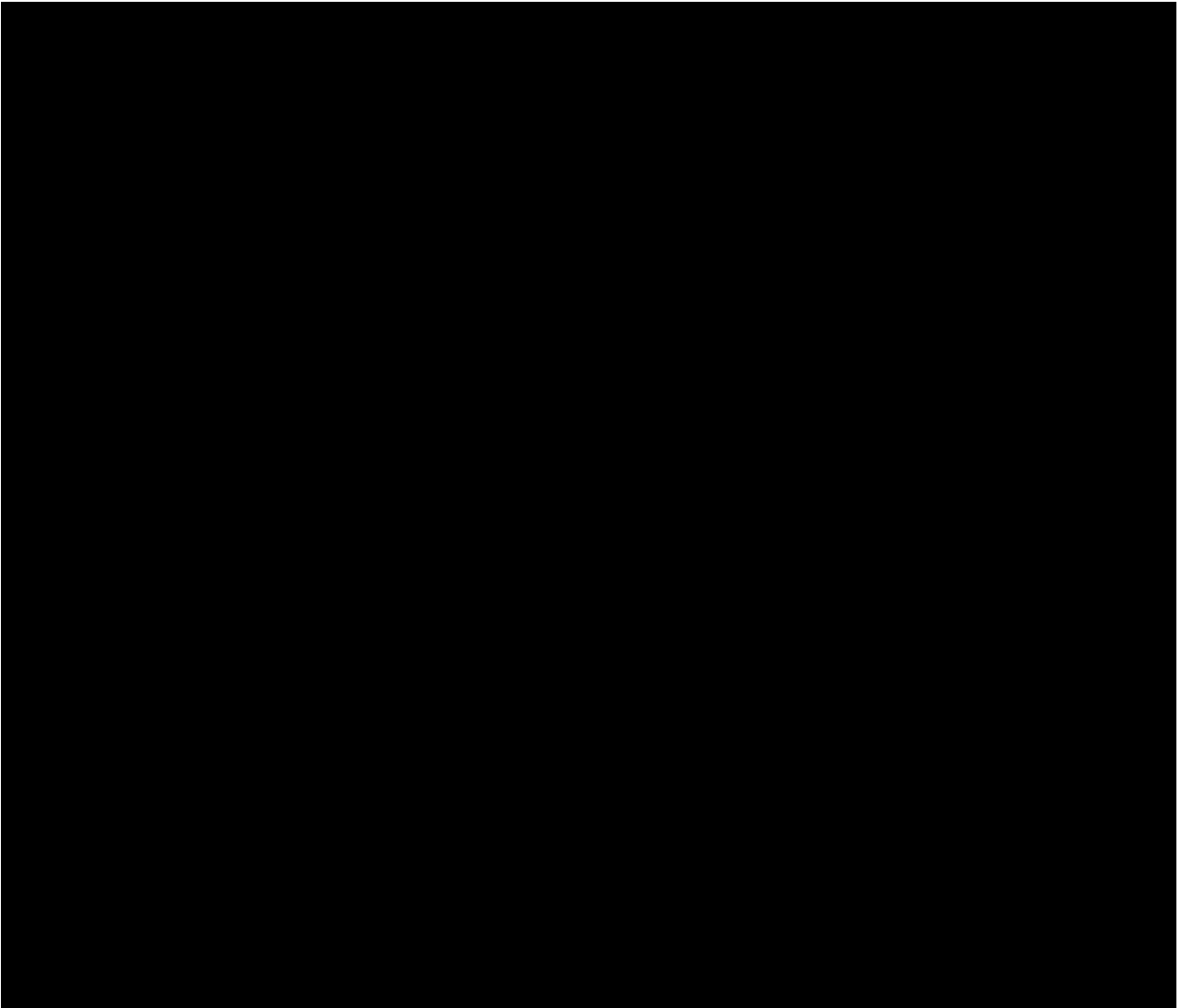
[Redacted text]

[Redacted text]

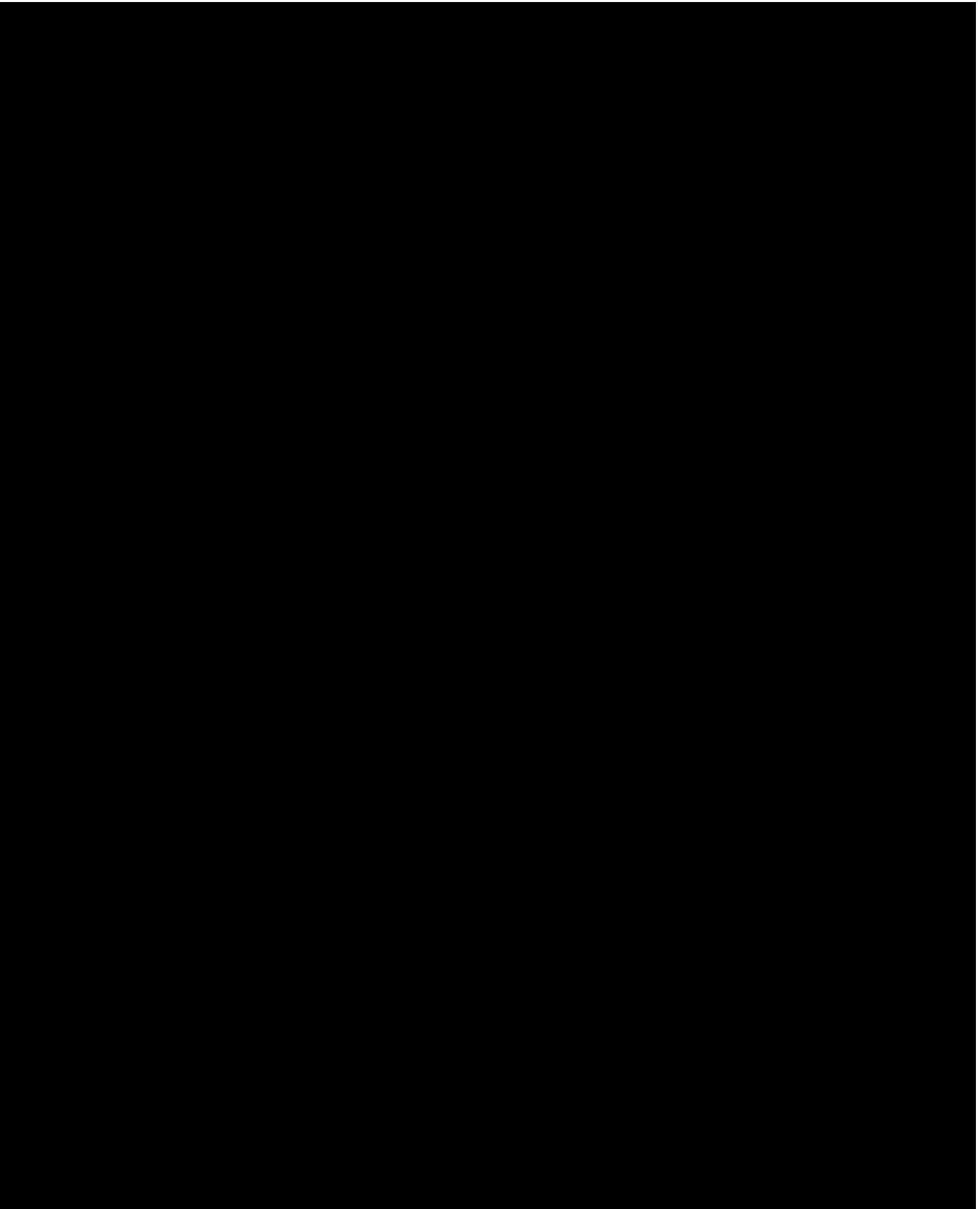
[Redacted text]

[Redacted text]











[REDACTED]

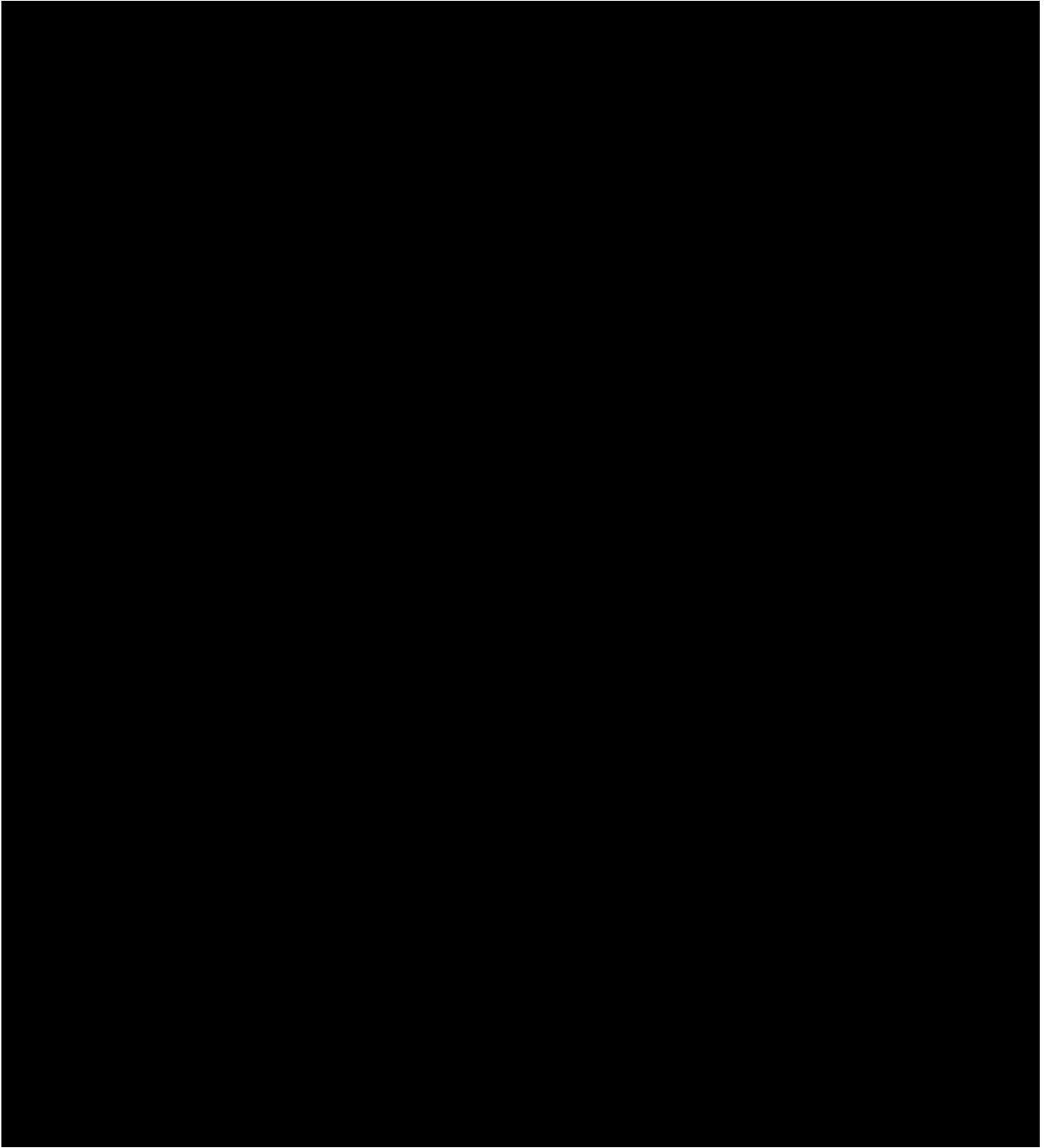
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

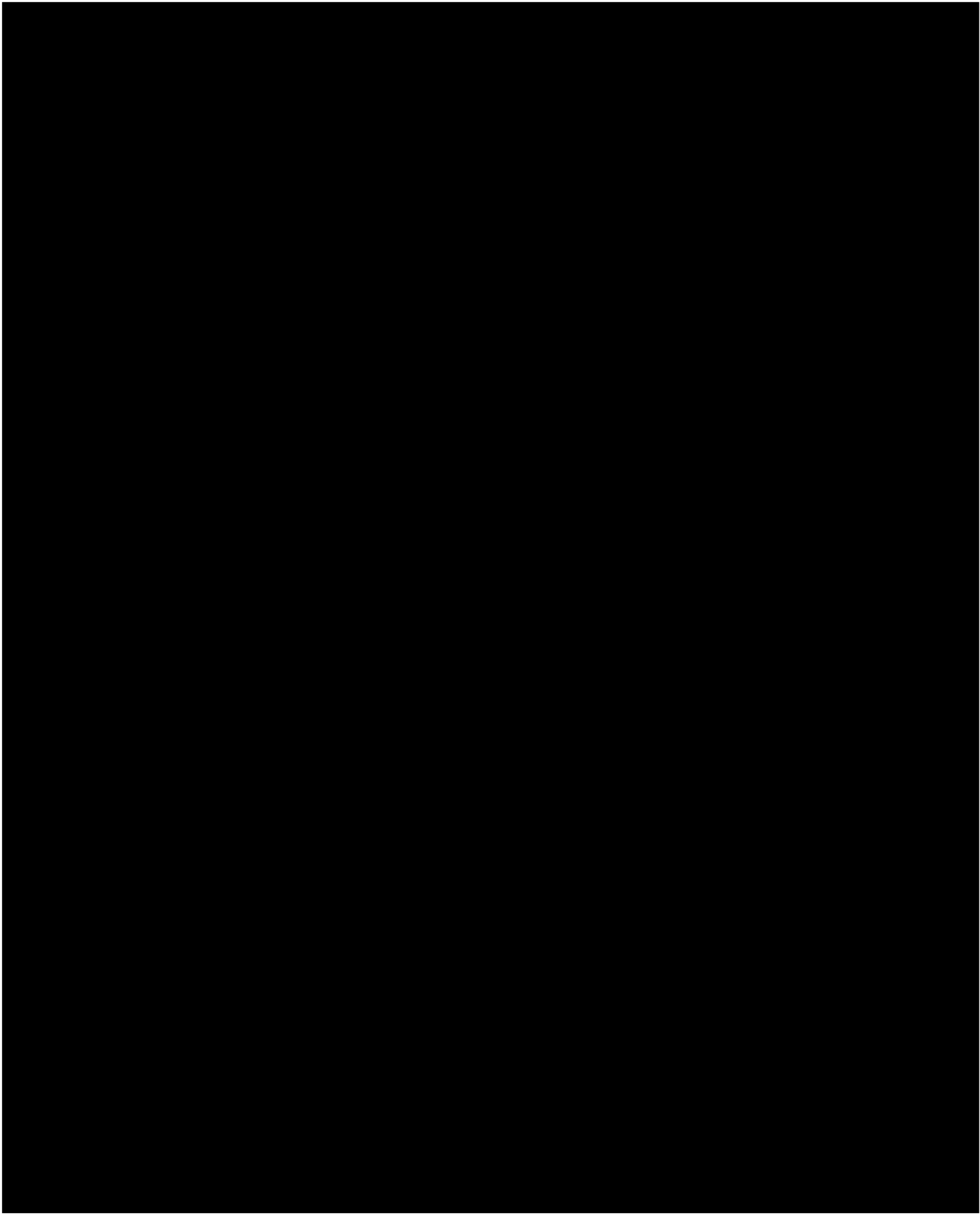
[REDACTED]
[REDACTED]
[REDACTED]

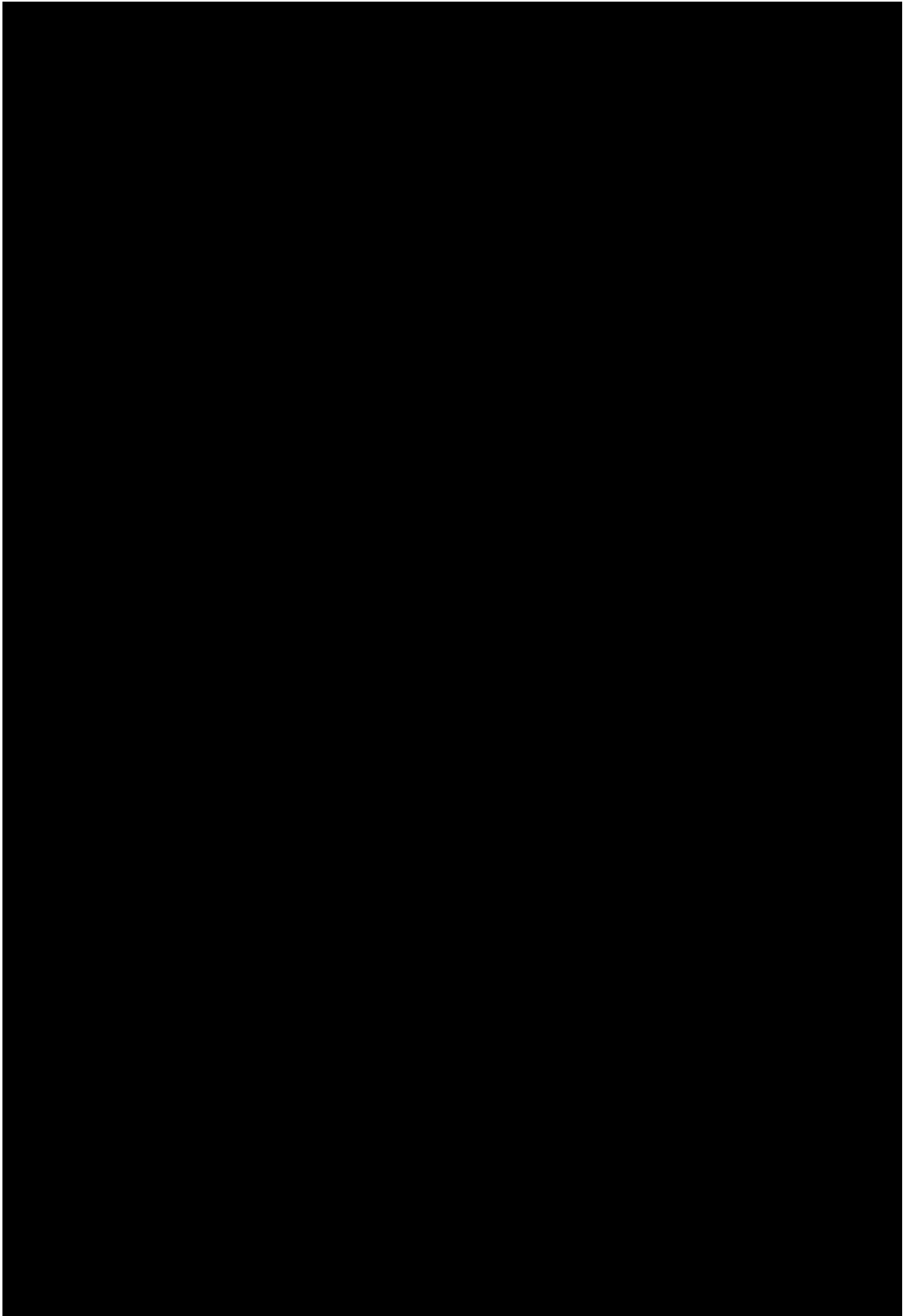
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]









5.4.6 Conclusion

Safety is a priority across Snapchat, and we use a combination of in-app reporting, automation tools, and human review to combat harms on the platform. All content must adhere to our [Terms](#), including our [Community Guidelines](#) and [Terms of Service](#), and some content must also adhere to our [Content Guidelines for Recommendation Eligibility](#). We strive to be transparent and consistent in our practices and enforcement, while striking the right balance between privacy and safety.

As explained in Section 4, we have concluded that Snap’s measures to moderate illegal or violating content, in combination with the other mitigations explained in this Section 5, are reasonable, proportionate and effective mitigation measures for the risks presented by Snapchat’s in-scope services.

5.5 Enforcement

5.5.1 Introduction

We strive to continuously update and improve our enforcement mechanisms to protect Snapchatters and our broader communities. As explained in the [Terms Section](#) of this Report, Snap has carefully developed its Terms with a view to mitigating the systemic risks it has identified for the EU (see [Section 4](#)). Integral to our risk mitigation efforts are Snap’s policies and processes to enforce these Terms. Below, we explain how we enforce our Terms in a transparent, consistent and equitable manner, balancing our commitment to safety with respect for the privacy interests of our community.

It is important to note that increases in reporting, enforcement and proactive law enforcement referrals over time do not necessarily indicate that Snapchat has become less safe. On the contrary, these upward trends correlate with a continued drop in Policy Violating Prevalence (PVP) on Snapchat overall (see our prevalence testing data in Section 6.4). In other words, as we get better at detecting and enforcing against more violations, the frequency of violations on Snapchat overall continues to decrease.

We are committed to continuously improving the safety of our communities on Snapchat and beyond. We use prevalence testing to identify and adapt to changing abuse trends on Snapchat so we can be better equipped to detect and address any gaps in enforcement.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



[REDACTED]

- [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

This strike system helps ensure that Snap applies its policies consistently, and in a way that provides warning and education to users who violate our Terms when appropriate. The primary goal of our policies is to ensure that everyone can enjoy using Snapchat in ways that reflect our values and mission; we have developed this enforcement framework to help support that goal at scale.

Transparency

Snap publishes within its Terms information regarding its policies and enforcement mechanisms related to misuse of the platform, including the following:

- Snap’s [Community Guidelines](#), including the [Severe Harms Explainer](#) and other associated Harm Category Explainers linked to from the Community Guidelines.
- [Snapchat Moderation, Enforcement, and Appeals](#) Explainer, which explains how our content moderation practices, enforcement, and appeals processes work. This includes a short guide to our strike system. We do not share further details about strike thresholds publicly because strike thresholds and related timeframes are subject to evaluation and occasionally change as part of our review processes and we do not want to give users the opportunity to “game” Snap’s enforcement system.
- [Content Guidelines for Recommendation Eligibility](#) which goes above and beyond the Community Guidelines, and limit other content from eligibility for recommendation that is appropriate for 13+ and is not considered illegal or harmful, but may still be unsuitable or unwanted in certain circumstances.



5.5.3 Notification of Criminal Offenses (Art. 18)

Proactive referrals to law enforcement and governmental agencies

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED] In addition, in instances where we become aware of potential child sexual exploitation content on our platform, our Trust & Safety team reviews the relevant information and, if appropriate, reports the content and account to NCMEC, as required by U.S. law. NCMEC then reviews those reports and coordinates with both U.S. and EU law enforcement agencies, as appropriate.

Law enforcement takedown and information requests (Articles 9 and 10)

EU Law Enforcement (“LE”) may submit orders for takedown of content or accounts via email to lawenforcement@snapchat.com. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

- | [REDACTED]
- | [REDACTED]
 - | [REDACTED]
 - | [REDACTED]
 - | [REDACTED]
 - | [REDACTED]
 - | [REDACTED]
 - | [REDACTED]
 - | [REDACTED]
 - | [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Law enforcement orders to provide information (Article 10)

Similar to the process for takedown requests, EU LE can submit specialized Article 10 requests via email to lawenforcement@snapchat.com.

[REDACTED]



[REDACTED]

- [REDACTED]
[REDACTED]
 - [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

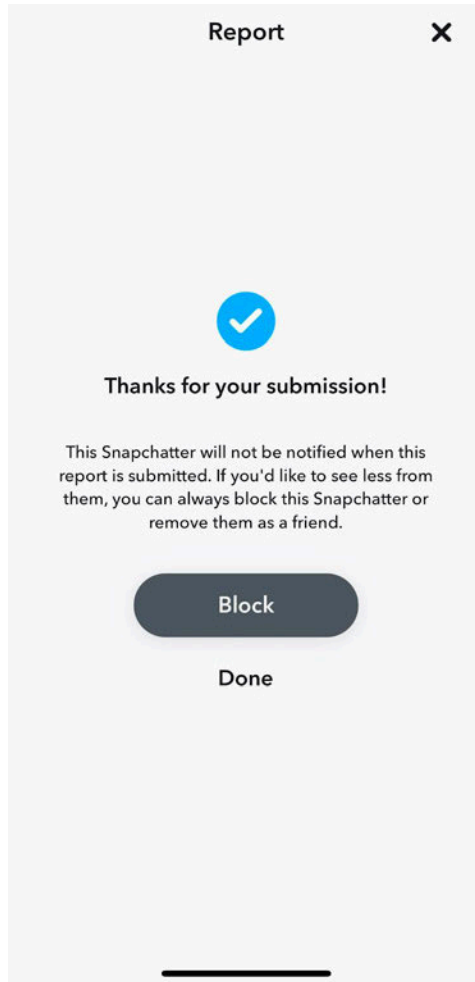
[REDACTED]
[REDACTED]
[REDACTED]

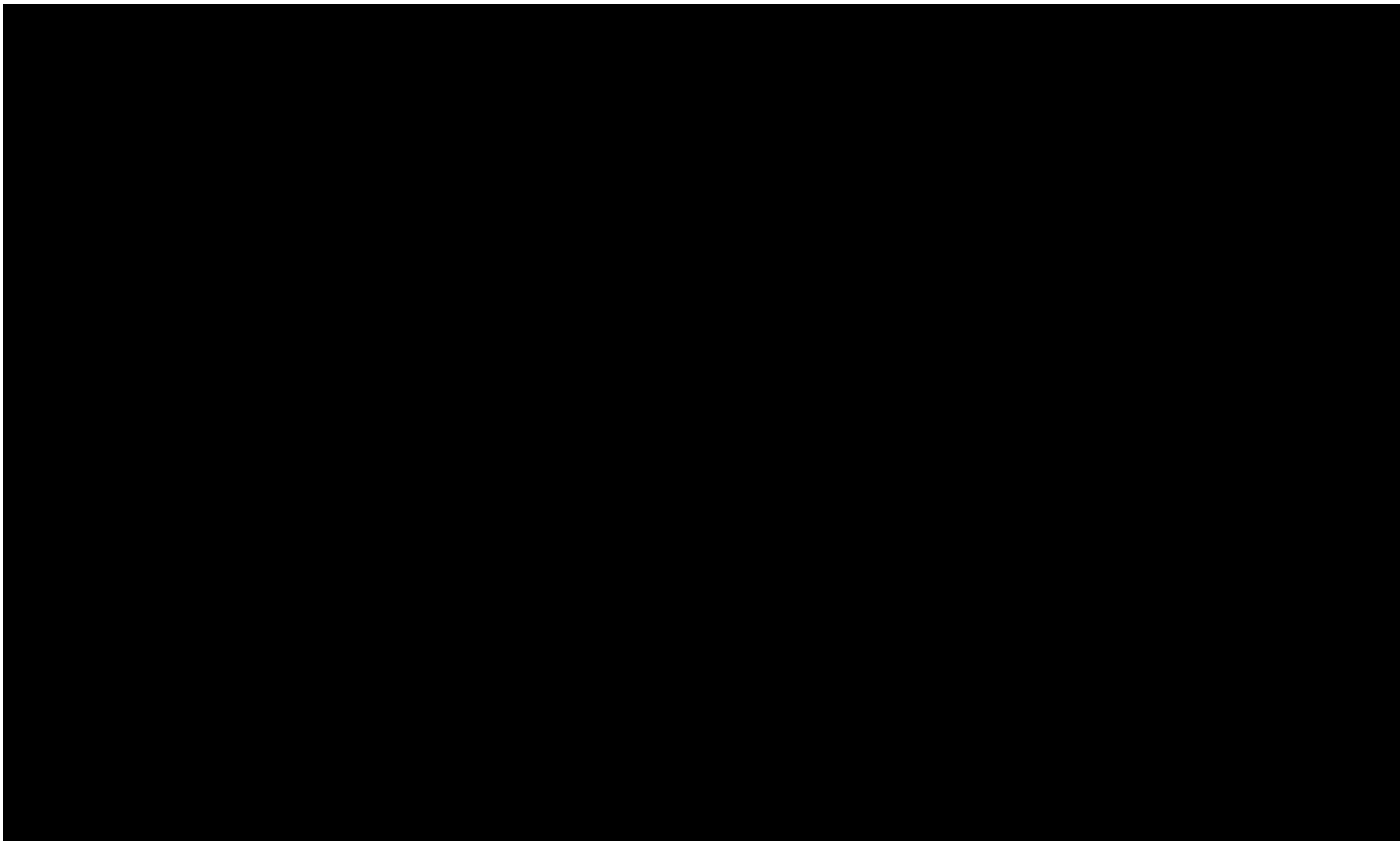
5.5.4 Notice and Appeals System

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

Notice to Reporter

After a reporter reports a piece of content or an account in-app for possibly violating our Terms or otherwise causing harm, we confirm receipt of their report and assure them that we are investigating it.





We endeavor to resolve all reports as quickly as possible while ensuring that we do a thorough review and achieve the correct result. Some reports can be resolved much more quickly than others, which may be more nuanced and require escalation to and/or input from other teams.

Account-Level Notice and Appeals

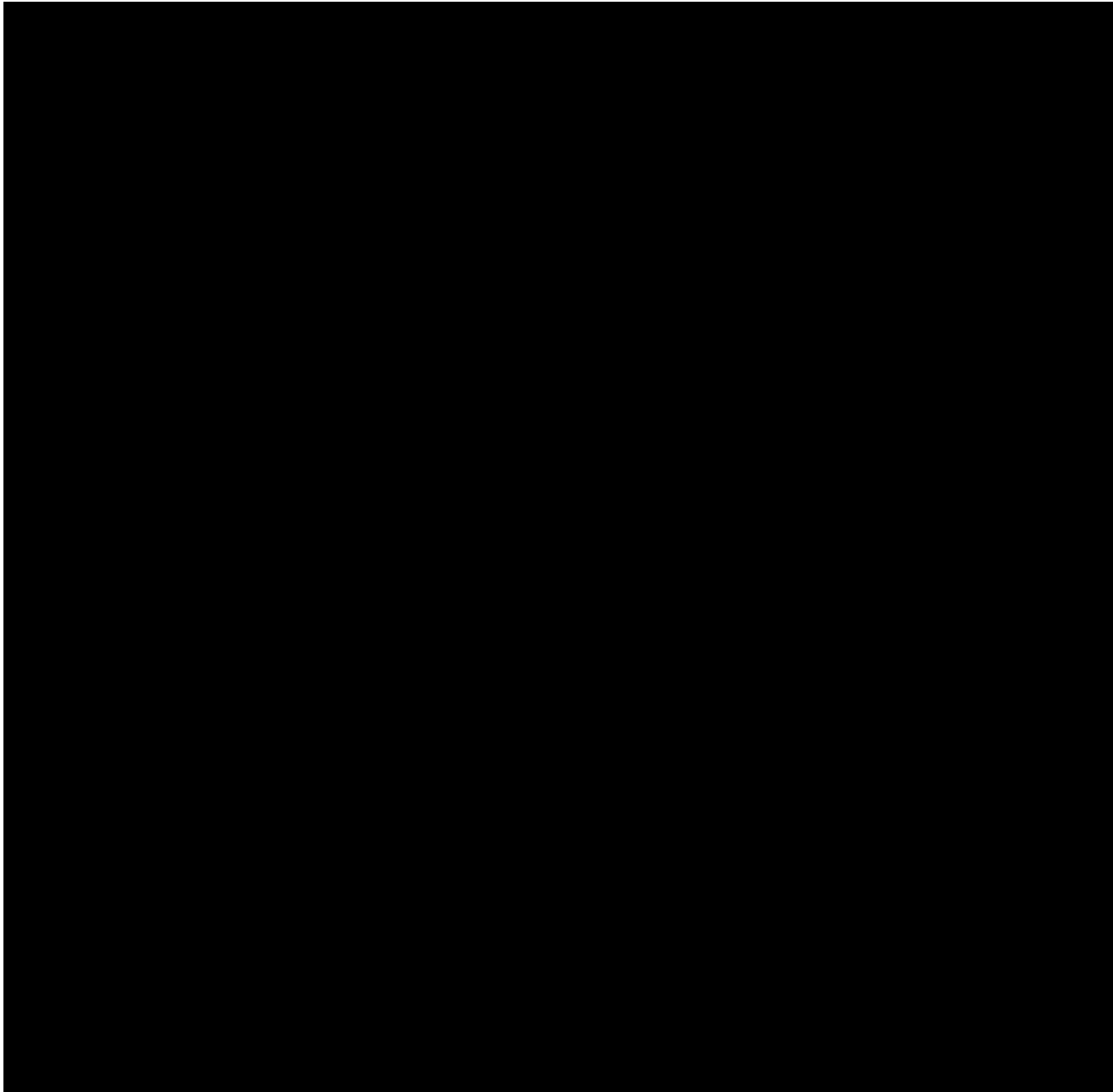




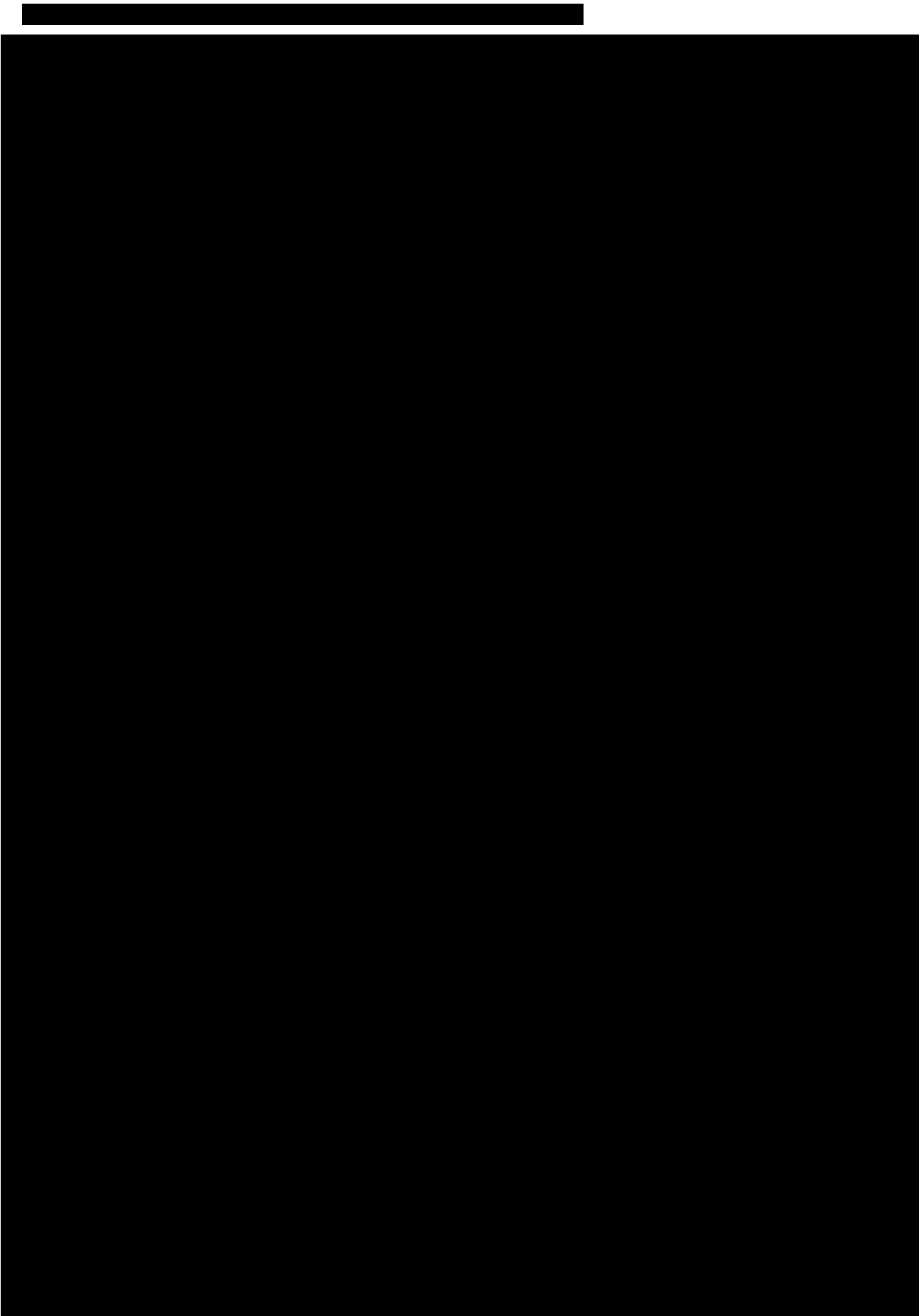




Upon submission of the appeal, an overview of the appeal process is provided. We allow the user to download their data in the interim.





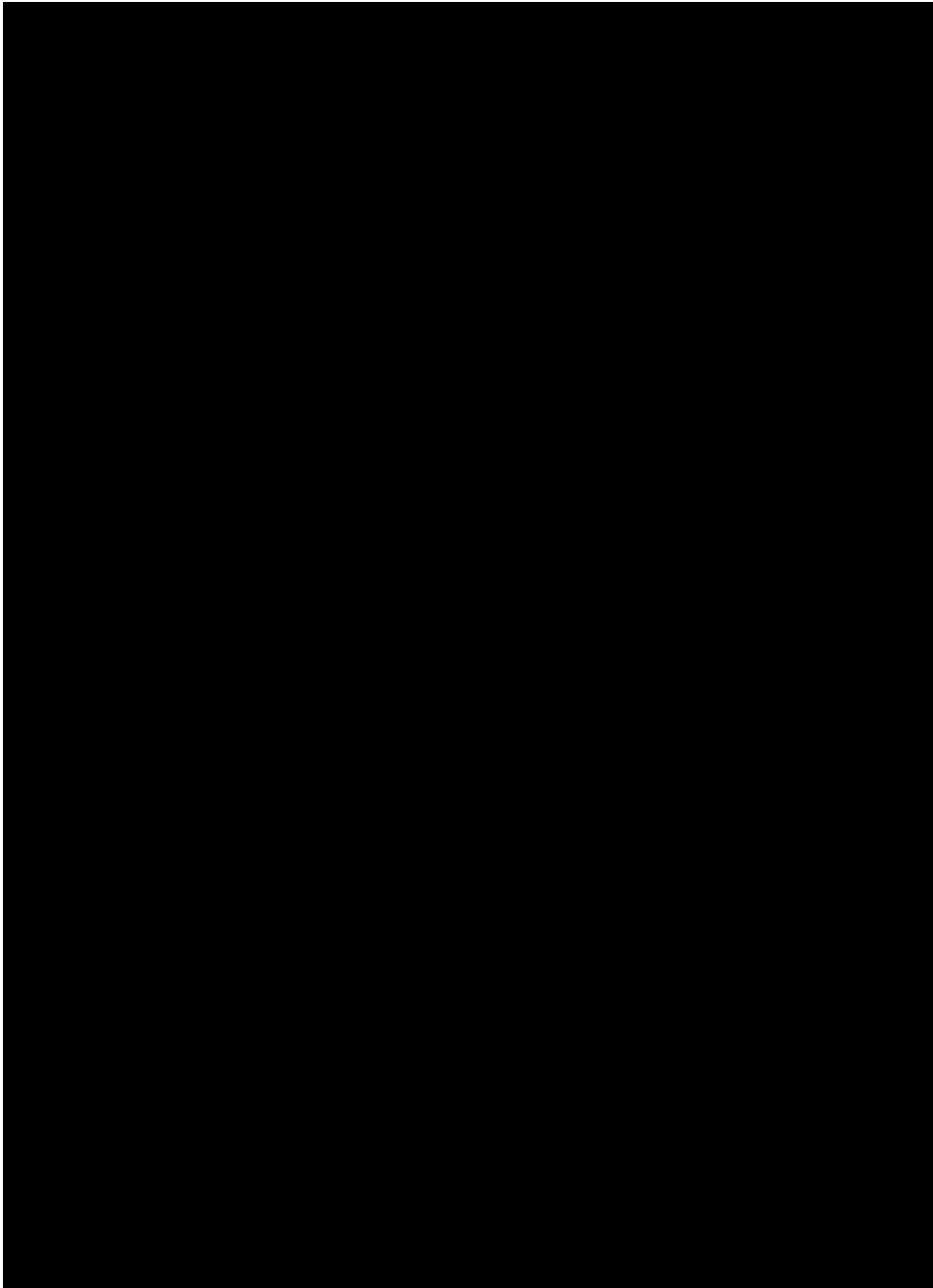




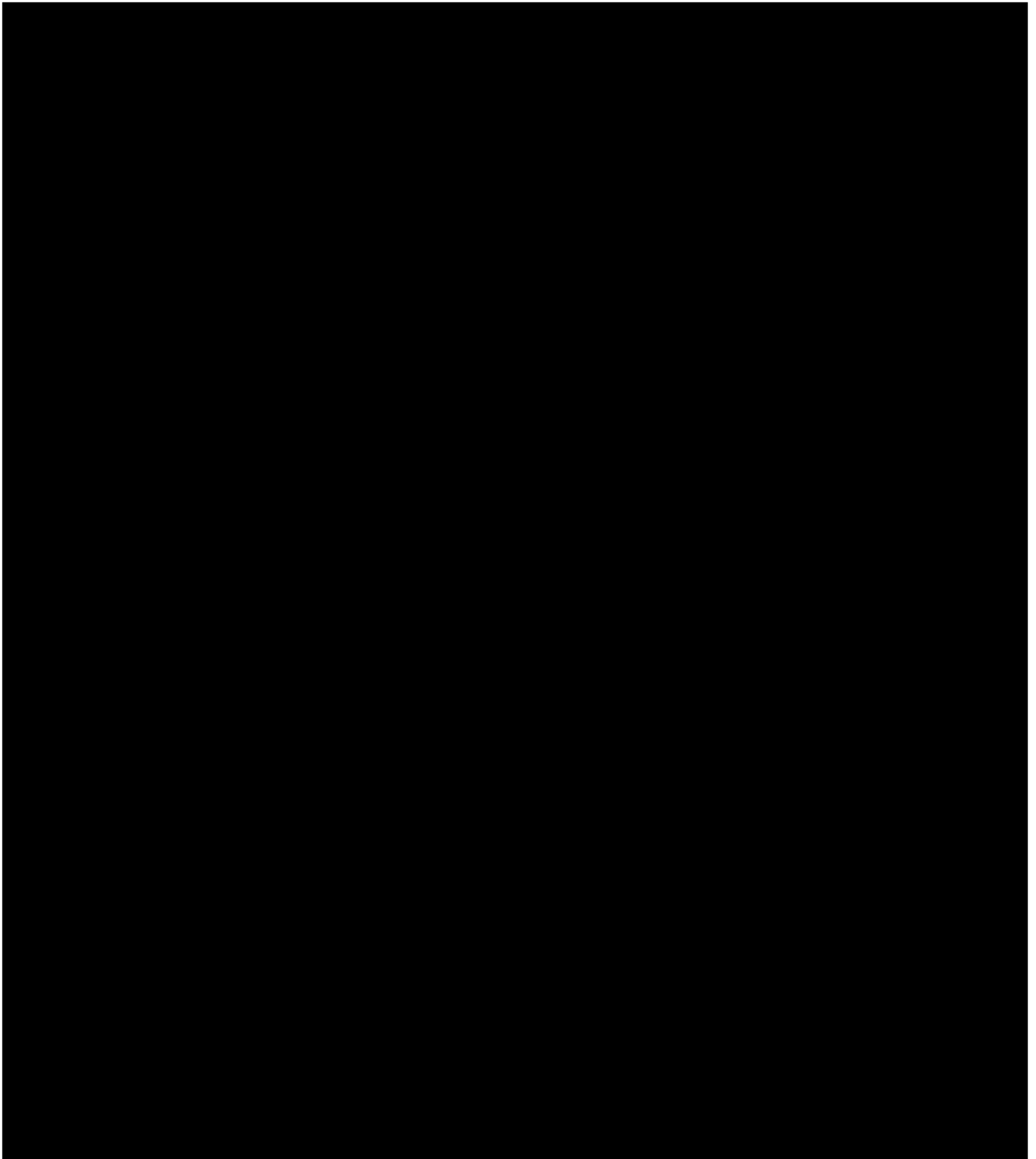
Content-Level Notice and Appeals

[Redacted]

[Redacted]









[Redacted text]

[Redacted text]

[Redacted text]

[Large redacted block]

5.5.5 Effectiveness of Enforcement

[Redacted text]

[Redacted text]



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

We regularly meet with NGOs and other stakeholders to discuss our measures and generally receive positive feedback. For example, we meet with Refuge, the UK's largest non-profit domestic abuse organisation for women, on a monthly basis. They have praised us for our responsiveness when individual cases have been escalated to us and they have positively singled out Snap as a receptive partner which offers direct channels of communication and cooperation.

Preventing, detecting, and eradicating illegal and other violating content on our platform is a top priority for Snap, and we continually evolve our capabilities. We continue to analyse enforced content and accounts to further develop our proactive measures to detect and root out bad actors. Our top ten risks are now in the 'low likelihood' category and we have observed steady declines in prevalence each year. See [Section 6.4](#) further details.

5.5.6 Protections against Misuse (Art. 23)

Suspending the Processing of Notices and Complaints

Snap's Trust & Safety team has procedures in place to suspend the processing of notices and complaints from individuals who frequently submit notices or complaints that are manifestly unfounded, for a period of up to one year. In egregious situations, we reserve the right to disable a user's account in relation to the above. These measures will continue to be reviewed and iterated.

5.5.7 Conclusion

Increases in reporting, enforcement and proactive law enforcement referrals over time do not mean that Snapchat has become less safe. On the contrary, these upward trends correlate with a continued drop in Policy Violating Prevalence (PVP) on Snapchat overall since our 2024 Report. In other words, as we get better at detecting and enforcing against an increased number of violations, the frequency of violations found on Snapchat continues to decrease overall.

We are committed to continuously improving the safety of our communities on Snapchat and beyond, and use prevalence testing to identify and adapt to changing abuse trends on Snapchat, so we are best equipped to detect and address any gaps in enforcement.

As explained in Section 4, we have concluded that Snap's measures to enforce its Terms, in combination with the other mitigations explained in this Section 5, are reasonable, proportionate and effective mitigation measures for the risks identified for Snapchat's in-scope services.



5.6 Algorithmic Systems

5.6.1 Introduction

Users are able to find new content on Snapchat primarily through our algorithmic personalization/recommendation service. While, in general, algorithmic content recommendation systems, like the one Snapchat uses, present a number of risks, we've designed our systems to mitigate these risks. This includes the use of appropriate terms, robust automated and human moderation, sufficient transparency with our users about the use of these systems, the ability to turn off personalisation, and the other mitigations outlined above.

This section describes specific mitigation measures that Snap has put in place with regards to Snapchat's algorithmic systems for the in-scope services on Snapchat to address the risks identified in its risk assessment pursuant to Article 34(1), DSA.

In line with Article 34(1), our risk assessment is proportionate to the risks identified taking into account their severity and probability, and the design of our recommender systems and other relevant algorithmic systems. While Snapchat uses several algorithmic systems across all of Snapchat's in-scope services: Spotlight, Discover, Map, Lenses, Public Profiles, Advertising, the risks identified in our risk assessment focused on the following algorithmic systems:

1. Content Recommendation Systems in Spotlight and Discover
2. Advertising Systems
3. Content Moderation Systems

The specific mitigations put in place for our Content Moderation Systems and Advertising Systems are covered in Sections [5.4](#) and [5.7](#) respectively. This [Section 5.6](#) is therefore primarily focused on the specific mitigations put in place for our Content Recommendation Systems in Spotlight and Discover, Map and Lenses (which we refer to in this Section as the "content recommender systems").

5.6.2 Content Recommendation Systems

Snap provides a free personalized content experience that is intended to entertain and delight users in the same app they use to communicate with their friends and family. Users find new content on Discover and Spotlight primarily through our algorithmic personalization/recommendation service.

Through our algorithm, users can also view location based content from the Snap community on Map. While Snap uses algorithms to rank what content to show users on Snap Map based on content location and 'recency', content shown on Snap Map is not personalized. Users can also find Lenses that are specific to their location but also more relevant and valuable as they are personalized to the users' interests.



[REDACTED]

Algorithmic content recommendation systems, like the one Snapchat uses, present a number of risks. For example, they may give rise to, amplify and/or result in the rapid and wide dissemination of illegal content and/or other harms identified in [Section 4](#), if not adapted and tested appropriately. We have designed our systems and processes to mitigate these risks. This includes the use of appropriate descriptive terminology, robust automated and human moderation, sufficient transparency with our users about the functionality of these systems, the ability to opt out of personalization, and the other mitigations as described in the testing and adaptation section of this document.

5.6.3 How do our Content Recommender Systems work?

To help users discover content they will be interested in, Snap's content recommender systems seek to understand the types of content viewers are interested in and not interested in. [REDACTED]

- [REDACTED]
- [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]



[illegible]

[illegible]



Benefits

Snap's recommender systems allow users to more easily discover interesting, entertaining, and relevant content. With over a million submissions a day of content, discovery methods like sorting by popularity, alphanumeric, timestamp, or curation are not practical.

Our recommender systems help viewers discover new interests they otherwise would have never found, and help creators who otherwise would not have been able to find an audience, allow users to learn, develop, play and have fun online. Users can explore different experiences, learn about topics of interest, and see what is happening around the world. Recommender systems are dynamic and responsive in that they can respond to viewers feedback.

We know users consider personalized recommender systems to provide significant benefit because:

- Viewers tell us (through their actions) that they prefer recommendations over other approaches and access to entertaining content is one of users' most frequent requests; and
- When we have tested removing personalization on Snapchat, we see a significant fall in user engagement (view time).

We also note that one of the reasons that traditional media services (i.e. linear television, newspapers, and magazines) are perceived to be in decline is because they are less entertaining to a diverse audience than the personalized alternatives provided by online platforms, such as Snapchat's in-scope services.

5.6.4 Adaptation of Snap Algorithmic Systems to Mitigate Systemic Risk

Enabling user choice in content prioritization

Snapchatters living in the European Union have the option to disable personalized public content recommendations. In Discover and Spotlight, users can disable personalized content by either tapping on '...' then 'Why am I seeing this content?' which will take you to Settings or by going directly to Settings and 'European Union Controls'. When users disable personalization, public content will be recommended to them with basic data only, such as the language set on the phone, age and country. Users will still see content, but it will be more random and less relevant to their interests. Users can also modify their content prioritization preferences across the platform by modifying notifications received by type or topic, toggling on or off Lifestyle Categories, and adjusting ad preferences within Settings.



5.6.5 Oversight and Administration

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]

Algorithmic System Review

[REDACTED] Snap conducts a comprehensive review of its Algorithmic Systems. The purpose of this review is to centrally catalog algorithmic systems that are significant to the functioning of Snapchat products as well as to safeguarding user safety and fundamental rights. This process is used to confirm understanding and documentation of significant algorithmic systems and review alignment of algorithmic systems with Snap’s standards of care for them.



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

- [REDACTED]
- [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED] common machine learning infrastructure for development, training, and testing models.



[REDACTED]

[REDACTED]

[REDACTED]

5.6.7 Adaption and Testing

In line with Article 35(1)(d), we explain in this part of the Report the extent to which we have adapted and tested our algorithmic recommendation systems to help address the risks identified in [Section 4](#) of this Report.

Summary

Snap has extensively adapted its algorithmic recommendation systems to ensure our content experience is beneficial to users, and that the risks of algorithmic personalization are mitigated.

[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]



114



[REDACTED]

	[REDACTED] [REDACTED] [REDACTED]	
[REDACTED] [REDACTED]	[REDACTED] [REDACTED] [REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]
[REDACTED] [REDACTED] [REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED]
[REDACTED] [REDACTED] [REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]

Considering each risk and its mitigation(s) in more detail:

Illegal or violating content

As explained in this [Terms Section](#) of this Report, all content on Snap must comply with our Terms which requires all public content on Snapchat to be suitable for users as young as 13, including our Community Guidelines. Additionally, content personalized by our algorithmic recommendation system must also comply with our more restrictive Content Guidelines for Recommendation Eligibility.

As explained in the [Moderation Section](#) of this Report, we have adapted our recommender systems and its processes to enforce our content policies with robust automated and human moderation, [REDACTED]

- [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
- [REDACTED]
[REDACTED]
[REDACTED]



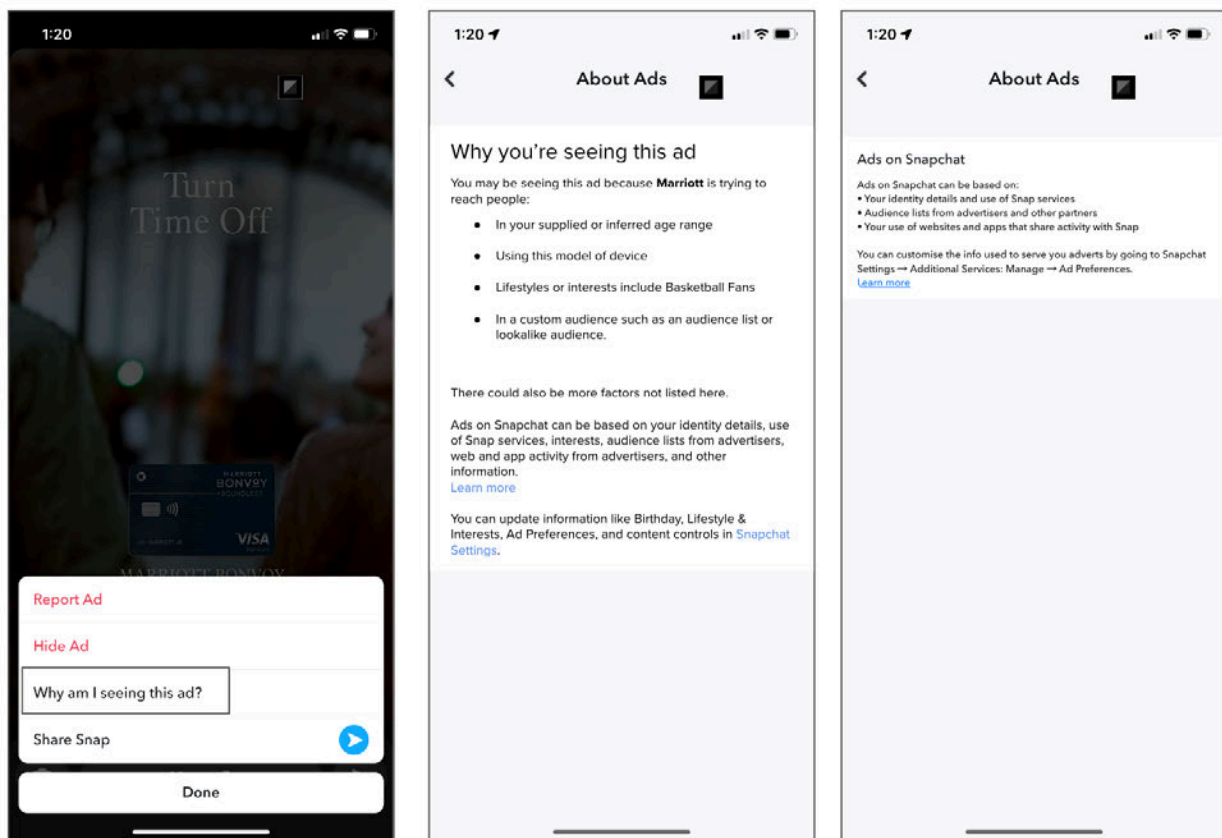


Lack of user understanding

Our recommender systems are complex and the process, the signals used in ranking and how significant each signal is to the recommender system can be challenging for users to understand.

To help users and answer frequently asked questions, and as part of our DSA compliance, we have:

1. Adapted our content to include links to articles available explaining how we personalize content in Spotlight, Discover and Ads [here](#). This includes a description of the main parameters used for our recommender systems, as well as the weighting applied to each signal.



2. Users may also reach out to our Support team if they have concerns or questions about how our algorithms work. We test this is appropriate by reviewing privacy and other consumer queries raised to our community support teams and our Data Protection Officer.

Intrusive personalized recommendations

We believe content is more relevant and entertaining when it's personalized to a user's interests, and not to someone else's. However, there is a risk that some users may experience personalized recommendations based on their inferred interest to be intrusive.



In Discover and Spotlight users can disable personalized content by either tapping on ‘...’ then ‘Why am I seeing this content?’ which will take the user to Settings or the user can navigate directly to Settings and ‘European Union Controls’. When the user disables personalization, the Discover and Spotlight experiences will be less personalized, and rely on essentials to determine what content to show the user, such as the language the user has set on their phone, their age, and country. Users will still see content, but it will be more random and less relevant to the user’s interests (as required under Article 38 DSA). If the user wishes to enable personalization again, users can do so either by tapping on the favorite icon () in Discover and Spotlight and then tapping ‘Enable’ in the ‘Show More Personalized Content?’ screen or by going to Settings in ‘European Union Controls’.

Users have the option to disable personalised Lenses through Personalisation Controls in Settings. If they choose to disable personalised Lenses, Snap will no longer recommend Lenses based on their profile. They will still see Lenses, based on basic data only such as their country and age, and Lenses will be more random and less relevant to their interests.

Discrimination

Algorithms that process special categories of personal data (as defined in GDPR) on a large scale are considered high risk and require explicit user consent. We have adapted Snapchat’s recommender systems so they do not track or identify special categories of personal data, including for the purpose of recommending content and ads.

Rapid and Widespread illegal or false content & crisis exposure

There is a risk of rapid and widespread illegal or false content on Spotlight, Discover and Map, as well as exposure to crisis situations and unexpected events like riots. We combat this risk by prohibiting illegal or false content in our [Terms of Service](#) and [Community Guidelines](#) and allowing users to report violations. More importantly, Spotlight relies on a combination of automated and human [moderation](#) on all submitted content before any video receives broad distribution. On Discover, Creators²⁰⁹ have their Stories distributed in Discover. Those that are approved have their Stories and the ‘tile’ art moderated. We also monitor reporting and hide rates on both Discover and Spotlight. In Map, Snap applies human and/or auto-moderation to decide what content is eligible to appear on Snap Map.

Filter bubbles

Our recommender system algorithms are designed to serve users with content that they will find engaging based on factors that include which categories of content they have previously watched. There is a risk therefore that, without safeguards, the algorithm will tag users who view content that may not be harmful on its own as being interested in that content and that repeated and frequent exposure to that content could be harmful. For example, while one piece of content

²⁰⁹ Creators include users whose content is eligible for distribution in Discover, such as Snap Stars and users who reach a certain number of followers.



related to dieting may not be harmful, if a user sees many or frequent videos about dieting, the user may feel inappropriately pressured to diet or may get a skewed perspective on how people manage their relationship with food.

We address this risk in a few ways. Firstly, we take significant steps to prevent and remove content that may become harmful when viewed frequently on Spotlight or Discover, including as explained above and in the [Terms](#), [Moderation](#) and [Enforcement](#) sections of this Report. Secondly, our content categories do not include harmful content categories and so in the unlikely event that a user does view harmful content, this will not be used by our recommender system algorithm to recommend similar content. Thirdly, in our Discover, Spotlight, Map and Lenses content recommendation systems, we have diversification rules in place to ensure that a particular category of content will only be capped or demoted within the recommender system for a given user. In other words, if a user is interested in makeup videos, we'll try to diversify the content by only showing a few makeup videos across a given period.

We evaluate our recommendations to users in terms of the number of categories of content we are introducing to them, while at the same time ensuring we do not overwhelm them with any particular type of content. This helps reduce the risk of filter bubbles, since users will be served diverse content even if our models show they have a strong interest in certain types of content.

Erroneously excluding content

There is a risk that our efforts to ensure appropriate content on Snapchat results in some content that is appropriate being mistakenly identified and incorrectly moderated. This may create for example, a risk to users rights to freedom of expression.

To combat this 'over-moderation', we evaluate and work to improve our automoderation in terms of precision and recall, and currently have high auto-approval precision for Discover, Spotlight, Map and Lenses. In addition, as explained in the [Enforcement](#) section of this Report, we have additional moderation transparency messages (statements of reason) and a more comprehensive appeals flow for moderated creators and content as part of our efforts to comply with the DSA.

Viewers could be watching our content but not enjoying content

There is a risk that the recommendation systems and models we build end up optimizing only for short-term metrics like engagement (i.e. time spent) in the Snapchat app, rather than in support of Snap's mission of "empowering people to express themselves, live in the moment, learn about the world, and have fun together". Our long-term objective when recommending content to users therefore goes beyond time spent and is focused on whether our users are enjoying themselves and are entertained and satisfied with their experience.

Snap evaluates the effectiveness at achieving this objective in multiple ways, in particular ensuring that we evaluate our algorithmic performance using a wide range of factors and not



[REDACTED]

solely relying on [REDACTED]

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

In addition, Spotlight has been designed not to distribute sensitive (i.e. shocking) content to 13-17-year-old users' Snapchat accounts, which includes non-glorifying discussion of self-harm and suicide content (such discussion is not prohibited on Snapchat but may still be sensitive). For users under 18, we will remove all content labeled as sensitive. For users over 18, we will limit its distribution [REDACTED]

To make sure users are enjoying themselves and are entertained and satisfied with their experience, we enable them to "Reset Suggested Content" in their settings. This allows them to "refresh" their public content feeds. When users enable Reset Suggested Content, most indicators used for personalisation of your public content feeds will reset but not all. Users may still see content or creators that they have followed and content tailored to their demographic, and past blocks and dislikes will still be taken into account.

We evaluate our algorithms across the above dimensions because we believe they are the drivers to the ultimate outcomes we are attempting to deliver for users: that they be (1) satisfied with our experience - which we survey regularly (i.e. quarterly) across all tabs in our app and (2) continue to use it (i.e. user retention).



We take into consideration the user's age when showing Lenses to our community. This way we can identify Lenses that are appropriate for users of a certain age and to the best of our ability in compliance with local laws.

5.6.8 Change Management

From a high level, Change Management over algorithmic systems at Snap is governed by the previously described Privacy and Safety by Design Review process. Material updates to algorithmic systems and material changes to model pipelines, input data, and third party user data are documented and reviewed.

5.6.9 Monitoring and Quality Assurance

Performance Monitoring

Snap monitors deployed algorithmic systems for anomalies and issues and establishes alerts to notify Engineering teams when potential issues arise. These alerts look for relevant spikes or anomalies in statistics.

Quality Assurance

Snap monitors algorithmic systems related to content moderation for quality and precision on a continuous basis. Monitoring may include:

- User Reports
- User Hides
- Content removal and user appeals
- Policy Violative Prevalence (PVP)
- Content Rejection
- User Reports
- User Hides
- Content removal and user appeals
- Policy Violative Prevalence (PVP)
- Content Rejection

Snap uses dashboards to visualize content moderation statistics and set alerts for spikes in content moderation activity. Snap Engineers may also investigate spikes in user reported content or automatically detected violative content to identify correlation between model deployment to feed back into broader Engineering teams.

5.6.10 Conclusion

Users find new content on Snapchat primarily through our algorithmic personalization/recommendation service. While algorithmic content recommendation systems, like the one Snapchat uses, present a number of risks, we've designed our systems to mitigate



these risks. This includes the use of appropriate terms, robust automated and human moderation, sufficient transparency with our users about the functionality of these systems, the ability to opt out of personalization, and the other mitigations outlined above.

As explained in Section 4, we have concluded that our adaptation and testing of Snapchat's algorithmic systems described above, in combination with the other mitigations explained in this Section 5, are reasonable, proportionate and effective mitigation measures for the risks identified.

5.7 Advertising Systems

5.7.1 Introduction

Snap relies on online advertising to support its business. Snap recognises that without mitigations its advertising systems also have a significant risk of giving rise to the concerns referenced in Article 34 of the Digital Services Act. Snapchat Ads Manager includes our digital ad products and tools created for advertisers who would like to easily create and manage ads that target relevant audiences on Snapchat. We process user information about Snapchatters to serve them with ads within Snapchat that we think they might be interested in. However, advertising systems in general might give rise to, amplify and/or result in the rapid and wide dissemination of illegal content and/or other harms identified in [Section 4](#), if not adapted and tested appropriately.

5.7.2 How do our Advertising Systems Work?

An overview of Snap's ads services can be found [here](#) and [here](#). In essence, Snapchat collects data about our users as they register, log in and use Snapchat. As is described in our [Privacy Policy](#), this data is comprised of:

- Information the user provides us
- Information we collect as the user interacts with Snapchat
- Information we collect from third parties

Snapchat Ads Manager and its various tools allow advertisers to leverage this data for targeted advertising. Advertisers can use our [Audience Insights tools](#) to see the estimated aggregated demographics, including age, as well as locations, interests and device overviews of their targeted audience. User-level data is not directly available to advertisers through these dashboards.

Some of Snap's advertising tools allow advertisers to benefit from Snap's use of data about their customers such as customer personal data provided by our advertisers and data collected from third-party services along with our users' personal data, to provide and improve ad targeting and measurement:



- Snap [Custom List Audiences](#) - An advertiser and/or their agent can use this service to upload customer list data to Snap via Ads Manager. See the [Custom List Audiences](#) section of our Business Help Center. Customer list data provided by advertisers is used to create an 'audience' of Snapchatters matching the information in the customer list data. This allows advertisers to target ads to that audience, or similar audiences, on Snapchat. See the [Custom Audiences Overview](#) in our Business Help Center.
- [Snap Pixel](#) and [Conversion API](#) - An advertiser and/or their agent can also use this service to help target their ads on Snapchat:
 - For Pixel, advertisers install a piece of JavaScript within their web pages which sends data to Snap when those pages are accessed by website visitors. See the [Install Snap Pixel](#) section of our Business Help Center.
 - For Conversion API, advertisers install Snap API code on their servers that facilitates passing web, app and offline events directly to Snap via Server-to-Server integration. See the [Conversions API](#) section of our Business Help Center.
- [Advanced](#) and [Estimated](#) Conversion are examples of the additional services that we offer to advertisers to target and measure the performance of their advertising using advanced privacy enhancing techniques.

Snap acts as a data processor of data relating to EU data subjects received from advertisers via the Custom List Audiences, Pixel and Conversion API services. It processes the information in accordance with advertiser instructions subject to its data processing agreement (which follows requirements set out in Article 28 of the General Data Protection Regulation (GDPR)).

Our ad ranking algorithm determines which ads are displayed to a Snapchatter who is in the selected audience for those ads. The ad ranking algorithm uses various signals, including prior ad interactions and social signals, to determine which ads that user is more likely to interact with and then combines this with the results of advertiser ad action for that Snapchatter, to select an ad to display. Snap analyzes prior ad interactions to target advertisements. For example, we may determine that a user is likely to swipe up on certain types of ads or download certain types of games when they see an ad on Snapchat. We may then use this information to show that user similar ads. This is explained on our [Snap and Ads Privacy and Transparency](#) page.

Snapchatter interactions with the ad (i.e. impression data) is then logged to (a) attribute impressions to conversion events (such as a purchase on an advertiser website or download of an advertiser app) to demonstrate the performance of the ad and (b) to further train the ad ranking algorithm.



5.7.3 Benefits

Snapchat is used by millions of people in the European Union. They use Snapchat because it fosters fast and authentic communication with those who matter most to them. It is why our community continues to grow.

We consider it is in the best interest of all our users, including 13-17s, for them to have access to the best, most entertaining version of Snapchat possible, allowing them to exercise their digital rights (such as access to information, association with others, have a voice and to play and have fun) regardless of their financial background and ability to pay. We receive feedback everyday from our users; calling for new features, functionality and improvements. We are only able to do this by raising revenue from other sources. In common with many others in the industry, this has meant turning to advertising.

Our ability to raise revenue by selling targeted advertising opportunities to advertisers means that:

- Snapchat is maintained and improved for the benefit of Snap and all recipients regardless of their ability to pay. If Snapchat was only available for a fee, it would only be accessible to those who could afford to pay the fee, restricting access to Snapchat and raising risks to fundamental EU rights to information and to access to services, particularly for Teens.
- Snapchatters benefit from being able to exercise digital rights and association with others online through Snapchat regardless of their financial background. This includes developing their voice, having fun and access to entertainment and play. Balanced use of their personal data also benefits Snapchatters by avoiding seeing advertisements that are not relevant to them (which is one complaint we have received in the past). Although Snapchatters are given options to manually hide advertisements, through the use of personal data, Snapchatters benefit from targeted advertising by seeing more relevant, age and interest appropriate adverts.²¹⁰ The greater the revenue Snap is able to generate the more resources Snap can dedicate to supporting access to the service and teens' development.
- Advertisers benefit from being able to promote their brand and products to a Snapchat audience most likely to be interested in them. This allows advertisers to focus their advertising and avoid spending on the display of advertisements to audiences that are not likely to be interested. Snapchat Ads Manager also allows advertisers to better measure the success of their digital marketing campaigns so their quality can be continuously improved. Advertisers are also conscious about safety on Snapchat. With this in mind, in March 2024, we announced a partnership with a leading global media measurement and optimization platform,²¹¹ to:

²¹⁰ N. Fourberg e.a., on 'Online advertising: the impact of targeted advertising on advertisers, market access and consumer choice', 2021, [url](#).

²¹¹ <https://forbusiness.snapchat.com/blog/snap-partners-with-integral-ad-science-brand-safety>



- Conduct a measurement sample study on the advertiser suitability of our public content, specifically Spotlight and Creator Stories. In the study results, IAS found that both Spotlight and Creator content on Snapchat is 99% brand safe.²¹²
- Jointly develop a new brand safety reporting solution that would give advertisers transparency into the percentage of safe and suitable content their ads are appearing against. The new solution launched in June 2024.²¹³

However, notwithstanding the benefits advertising systems bring to our users, to Snap and our advertisers, we recognise that our targeted advertising will only operate in the best interests of all our users provided that the processing of individuals' personal data (including by way of profiling) to facilitate the sale of ads that fund Snapchat does not result in our users being subject to 'economic exploitation'. Privacy and Safety are central to Snapchat's values. When we first introduced advertising to Snapchat, we ensured those advertising systems appropriately balanced the legitimate benefits explained above with individuals' fundamental rights and freedoms, in line with Snap's strong privacy and safety principles. We have continued to uphold these values throughout Snapchat's life, adapting and testing our advertising systems to mitigate risks they may give rise to as identified in [Section 4](#) of this Report.

5.7.4 Adaptation and Testing

In line with Article 35.1(e) DSA, Snap has adapted Snapchat's advertising systems and adopted targeted measures aimed at mitigating the risks presented by its advertising systems, including by limiting or adjusting the presentation of advertisements on Snapchat, to help address the risks identified in [Section 4](#).

[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]

²¹² "Brand Safety" is based on the GARM standard, only considering content classified at the "Floor" risk-levels .

"Creator content" is image and video user-generated content posted to Public Stories; IAS sampled a wide variety of creators in the U.S. to ensure a representative sample. Spotlight content is user-generated video content that appears in the Spotlight tab on Snapchat; IAS audited content from US, CA, GB, UK, NZ, AU markets. IAS audited both Spotlight and Creator content from Oct 30, 2023 - Jan 2, 2024.

²¹³ <https://forbusiness.snapchat.com/blog/snap-ias-solution-ga>

439



[REDACTED]

processes.		
Without transparency, all adverts have a higher risk of violating our terms or the law.	We have adapted our systems to include an accessible, explorable ads library.	We conduct pre-launch testing.

We're considering each risk and its mitigation(s) in more detail below:

Invasion of Privacy – Reasonable and Proportionate Targeting

We recognise that, as a platform, we have a responsibility to raise revenue in an appropriate manner, and we take this responsibility very seriously. We want to ensure advertisers are not targeting specific individuals on our platform and that users do not feel like their privacy is being compromised by our advertising. We also want to prevent advertisers from manipulating small audiences with microtargeted campaigns, particularly for political ads.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

- Most of the ads on Snapchat, including all political ads, require a specific minimum audience [REDACTED] to be targeted. [REDACTED]
[REDACTED] This prevents adverts from being micro targeted.
- Snap generally has a short retention period for user content. Unlike some of our peers, we do not store content for excessive periods solely for monetisation purposes.
- Advertisers can only use our data for ads targeting indirectly via the targeting tools available on Snapchat. Amongst other things, this allows advertisers to target audiences based on a limited number [REDACTED] of high level interest-based lifestyle categories (SLCs) audiences (none of which are available for targeting 13-17 year olds in EU, UK, Norway and Switzerland), which we have inferred a Snapchatter may be interested in. They are based on high level, non-sensitive categories inferences, such as Business News Watchers, Sports Fans, and Fashion & Style Gurus, that users can see and control in the app, as detailed in [this support page](#). The interest categories are intentionally short-lived (13 months), sufficient to allow a year-on-year comparison. All users can manage their advertising interest categories in settings and view them via our Download My Data tool (DMD). [REDACTED]
[REDACTED]
[REDACTED] None of these SLCs are aimed at 13-17s specifically and the user-level targeting data is not directly available to advertisers.



We feel confident that our approach to advertising is reasonable and proportionate, as we have a low incidence of issues in relation to age targeting. Our approach to targeting minimums is based on mathematical analysis by our privacy engineering teams.

Special category data – No sensitive data use

Special categories of data concern information revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation. This information may have a greater risk of causing harm if used to target ads. A famous example of this concern relates to supermarket Target, which profiled purchase information to determine if a woman was pregnant, and revealed the pregnancy to a teenage girl's father by mailing vouchers for baby accessories.²¹⁴

Our targeting parameters such as Lifestyle Categories do not include special / sensitive categories. In addition we require in our Terms relating to advertising that advertisers do not send us this data from their sites. We do not allow advertisers to target audiences based on sensitive categories. When we discover advertisers sending us this data we remove it.

Our legal team has reviewed and confirmed our Lifestyle Categories do not include special / sensitive categories. Any changes must be reviewed by our legal team as part of the mandatory product review which forms part of our privacy and safety by design processes.

Discrimination – Special Targeting Models

Certain advertisers are in regulated sectors where there are rules to prevent discrimination, such as the housing, credit or employment sectors (particularly in the United States) and our teenage users.

In order to ensure we are not using discriminatory targeting models particularly when there is significant legal impact to the consumers, we offer special targeting models that do not include gender or age, which we require for advertisers who are advertising in the housing, credit or employment (HCE) spaces, so that discriminatory factors will not go into who sees these ads. We do not allow advertisers to build audiences for their ads based on their own data about our teenage users regardless of those user's own ad settings (i.e. activity data from the advertisers own online properties and the advertiser's own customer lists).

²¹⁴ Drive Research: How Target Used Data Analytics to Predict Pregnancies ([url](#)).



We use pre-launch testing and mandatory legal and privacy engineering review of any significant changes to these models.

Harmful or illegal content – Advertising policies

Snap has [Advertising Policies](#) in place that direct the types of ads and targeting that are acceptable on Snap. Ads must comply with applicable laws and regulations in each geographic area where they will run. Ads are prohibited from collecting sensitive categories of data. This policy also helps ensure advertisements submitted are in line with Snap's Community Guidelines.

As explained in [Section 4](#) above, our advertisers could use our advertising systems to disseminate information that is illegal or could otherwise harm users, impact their fundamental EU rights or negatively impact public security or health.

As explained in the [Terms Section](#) of this Report, we ensure advertisers are clear about their obligations, we have robust [Advertising Policies](#) to prevent inappropriate and illegal advertising on our platform. The systems used by advertisers to create and submit advertising (such as our Snap Ads Manager), have been adapted to require agreement to these Terms and provide easy access to guidance on what is required.

We test advertisers' compliance with these Terms using our Advertising Review process before advertising can be published. See below for more information.

Policy-violating or illegal content – Advertising Review

Notwithstanding that advertisers agree to our Terms, they may still deliberately or mistakenly seek to publish advertisements that violate our advertising policies or the law.

As explained in the [Moderation Section](#) of this Report, in particular the part relating to [advertising moderation](#), we use a combination of automated and human review to prevent ads that violate our policies or the law from appearing on Snapchat. We reject hundreds of thousands of adverts globally each month. We have a global team that supports ad moderation across 15+ languages and is composed of both full time employees (FTEs) and contractors. Ad Review team members are responsible for reviewing ad submissions to ensure ads abide by Snap's creative policies and technical requirements. Ad Review team members use [Snap's Advertising Policies](#) to assess compliance. Ads must comply with Snap's Community Guidelines and Advertising Policies in order to be approved. Grey area ads are discussed with Snap's Legal and Policy teams. Depending on the seniority, members of the Ad Review team also collaborate with the Sales team to create a consistent review experience for our Snapchat partners.

Fraudulent advertising accounts for the majority of these rejections and our advertising review teams are particularly vigilant for this form of violating advertising. This also includes ensuring inappropriate ads are not targeted at Teens. Our review takes account of the targeted audience i.e. if the ad is for alcohol and the selected demographic for the ad includes Teens, then it will be



rejected. We use inferred age, as well as declared age, to help ensure Teen users see ads that are appropriate for their age. Inferred age is regularly checked to ensure it is up-to-date.

We monitor ad reporting and enforcement data to ensure our review process is catching a reasonable and proportionate level of violating adverts.

We aim to ensure that all advertising review is maintained within a 24 hour SLA window from when the advertisement is created by the advertiser. More information on Snap's ad review process, including the timeliness of review, are located on [Snapchat's For Business website](#).

Bypassing Moderation Controls – Advertising Reporting

Although we have an advertising review process in place to prevent the publication of advertisements with information that violates the law or our policies, it is possible that some of these advertisements may be missed or incorrectly reviewed and be published.

As explained in the [Enforcement Section](#) of this Report, our advertising systems have been adapted with an easy mechanism for advertisements to be reported by Snapchatters from within the Snapchat app as being inappropriate along with the reason for the violation. Based on the number of reports, we will take down the ad or send it to human review for additional moderation.

All ads that are reported are reviewed by our human moderation team. Upon reporting the ad, Snapchatters are able to select a reason and write in comments. Both the reporting reason and the comment are provided in the moderation task, as well as the number of reports. We closely monitor sentiments of the ads on our platform and when ads are taken down, we inform the advertiser. We also monitor the aggregate number of reports for advertisements on a regular basis.

We monitor ad reporting and enforcement data.

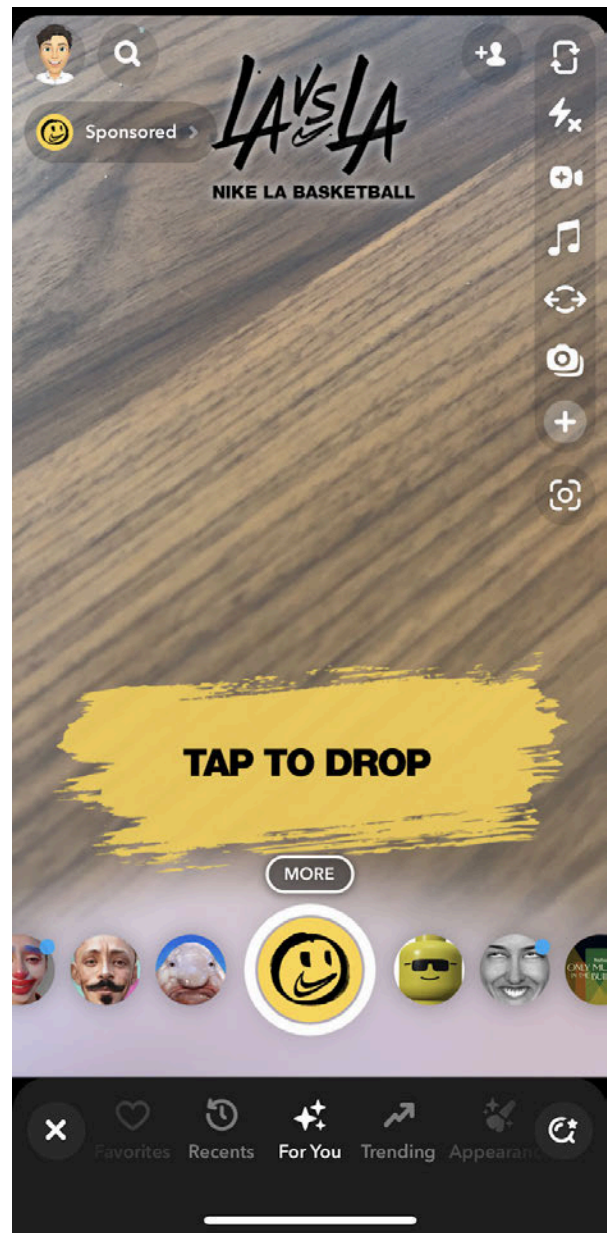
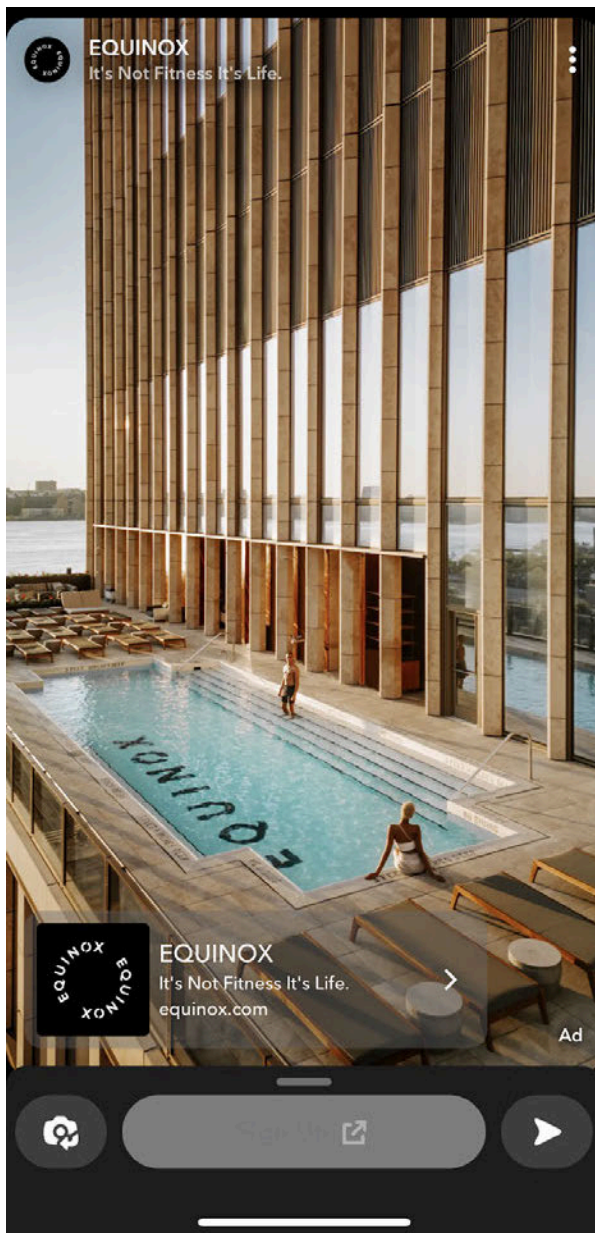
Unclear Commercial Intent – Ad Markers

If users are not aware when content is an ad or sponsored or other commercial content, there is a risk that this may lead to confusion, deception and exploitation.

We automatically place an “Ad” marker on all paid ads that run on Snapchat. Sponsored Lenses say “Sponsored”. Our commercial content policy requires all organic content posted by influencers to be marked appropriately. We now offer a “Paid Partnership” tag tool that influencers and users may use when they post commercial content to help them comply with this policy and their legal obligations.

Ad marker example

Sponsored Lens example

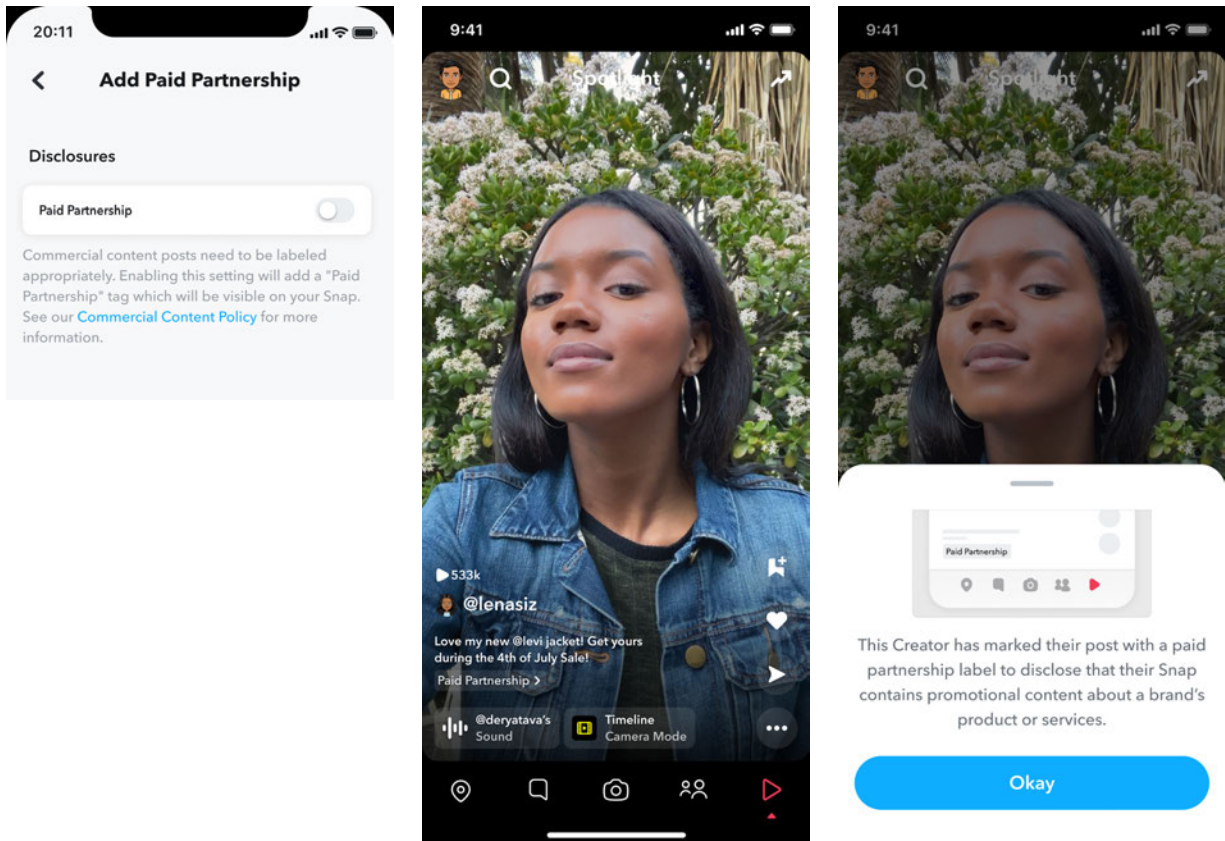


See below for examples of the “Paid Partnership” tag tool that influencers and users may use when they post commercial content to help them comply with this policy and their legal obligations.

Add Paid Partnership

Paid Partnership label

Paid Partnership Explainer



Political ads – Transparency Safeguards

Political advertising presents a higher risk of misinformation and Negative Effects on Democratic and Electoral Processes. Snap intends to discontinue political advertising in the European Union to comply with the European Union's Political Ads Regulation which comes into force in October 2025. Snap ensures that ads shown are in line with Snapchat's [Advertising Policies](#), which contain a subsection on political and advocacy advertising policies. Snap allows political, election-related, advocacy and issue ads, but sets additional requirements and places additional transparency safeguards in relation to publishing these types of ads:

- Political advertising must comply with all applicable laws and regulations, including all national election laws, copyright law, defamation law;
- All political advertising must include a “paid for by” message in the ad that is followed by the name of the paying person or entity. Snap may also require a “paid for by” disclosure on ad content that links to political content, ad content for political merchandise, or in other cases in Snap’s sole discretion;
- Like all ads on Snapchat, political ads must comply with Snap's Terms of Service, Community Guidelines, and our Advertising Policies.
- We encourage political advertisers to be positive. But we don't categorically ban “attack” ads; expressing disagreement with or campaigning against a candidate or party is



generally permissible if it meets our other guidelines. That said, political ads must not include attacks relating to a candidate's personal life.

- Snap will review political ads on a case-by-case basis. To get started, political advertisers are required to fill out our [political advertiser form](#). Snap reserves the right to reject, in our sole discretion, or request modifications to ads that we believe violate the standards listed above or that are otherwise inappropriate. Our discretion will never be exercised with the intent to favor or disfavor any candidate, political view, or political party.
- Snap may publicly display and otherwise disclose information relating to political advertising, including ad content, targeting details, delivery, spend, and other campaign information.
- Snap has for some time provided transparency for political ads with its [political ad library](#).

As explained in the moderation section of this Report, higher risk adverts (including political adverts) are subject to human review.

Personal Data Use for Targeting – User Choice

Opting out of Personalized Advertisements

If you live in the European Union, you have the option to disable personalized public content recommendations in Settings under 'European Union Controls.' Snapchat prevents personalized content from being served to EU users who have opted out of personalized content. Our Ad Server automatically prevents targeted ad types (such as pre-defined audiences and custom audiences) from being approved for ads reaching Snapchatters who have opted out. When you disable personalization, public content will be recommended to you with basic data only, such as the language you have set on your phone, your age, and your country. You will still see content, but it will be more random and less relevant to your interests.

Do Not Track

Snapchat respects the device iOS App Tracking Transparency setting regardless of user Ad Preference choices.

Ad Preferences

For EU users who have enabled personalized content, additional ad preferences are available to modify in Snap.

Users have choices about how Snap and its third-party ad partners use information about you for advertising purposes both on and off Snapchat. Snapchat settings allow users to select whether to receive Audience-Based ads, Activity-Based ads, Ads from Third-Party Ad Networks, or none of these ad options. The settings on this page are automatically disabled if you are under 18 in the EEA or UK.



Snapchat infers your interests to better suggest content to you and personalize your experience. Lifestyle categories are used to target ads (excluding users under 18 in the EU and UK) and personalize other content. Snapchat displays these inferences to Snap users and allows them the ability to toggle off specific Lifestyle Categories as desired. Snap's Ad Policies including age restrictions for Alcohol and Gambling ads will control the ads you see regardless of these settings.

<

Ad Preferences

You have choices about how Snap and its third-party ad partners use information about you for advertising purposes both on and off Snapchat. The settings on this page are automatically disabled if you are under 18 in the EEA or UK. These settings do not override or impact the iOS App Tracking Transparency setting for your device. Snapchat respects the device iOS App Tracking Transparency setting regardless of the choices here.

Audience-Based

These are ads targeted based on audience lists we receive from advertisers and other partners. For example, advertisers that already have information about you, like an email address, may want to reach you on Snapchat or other services on which we serve ads. If this setting is enabled, we may use audience information from advertisers and other partners to customize ads for you.
[Learn more](#)

Activity-Based

Snap may target the ads we show you based on information about your activities off Snapchat or other services on which we serve ads. If this setting is enabled, we may use your activity outside the service in which you see the ad to target ads for you. For example, if you search for a movie on a website that shares data with Snap, you may see ads for other movies. If you opt-out, we will not use your Snapchat data to target you with Snapchat ads on third-party advertising platforms that act as a data controller.
[Learn more](#)

Third-Party Ad Networks

Snap may use third-party ad networks to serve ads on Snapchat. To do this, Snap sends the networks a limited amount of data, for example, IP Address, Mobile Ad ID, and whether you saw the ad or interacted with it, so the network can provide advertising services, like ad targeting, measurement and optimization to us and its advertisers. When this setting is enabled, we may allow third-party ad networks to serve you ads on Snapchat.
[Learn more](#)

<

Lifestyle & Interests

Snapchat infers your interests to better suggest content to you and personalize your experience. Lifestyle Categories are used to target ads (excluding users under 18 in the EU and UK) and personalize other content. Snap's Ad Policies including age restrictions for Alcohol and Gambling ads will control the ads you see regardless of these settings.

Action & Thriller Genre Fans

Adventure Seekers

Advocates & Activists

American Football Fans

Arizona Cardinals Fans

Arts & Culture Mavens

Atlanta Falcons Fans

Automotive Enthusiasts

Automotive Shoppers

Bachata Music Fans

Baltimore Ravens Fans

Baseball Fans

Basketball Fans

Transparency and Control

Some users may have specific vulnerabilities or other reasons to be concerned about any use of their personal data for targeting ads. If users do not understand how advertising works, they may not be able to confirm whether they should be concerned or exercise any choices they may have.

As explained in the [Transparency Section](#) of the Report, our privacy center provides extensive information regarding our processing of personal information. This includes a [dedicated page](#) explaining how we use personal data for advertising purposes. We offer choices for users to control the data that's used to determine the ads they see. In the European Union, we have introduced controls to turn off most personalized ads except those based on real time location, language, age and device type, and this is always turned off for teen users in the European Union

447

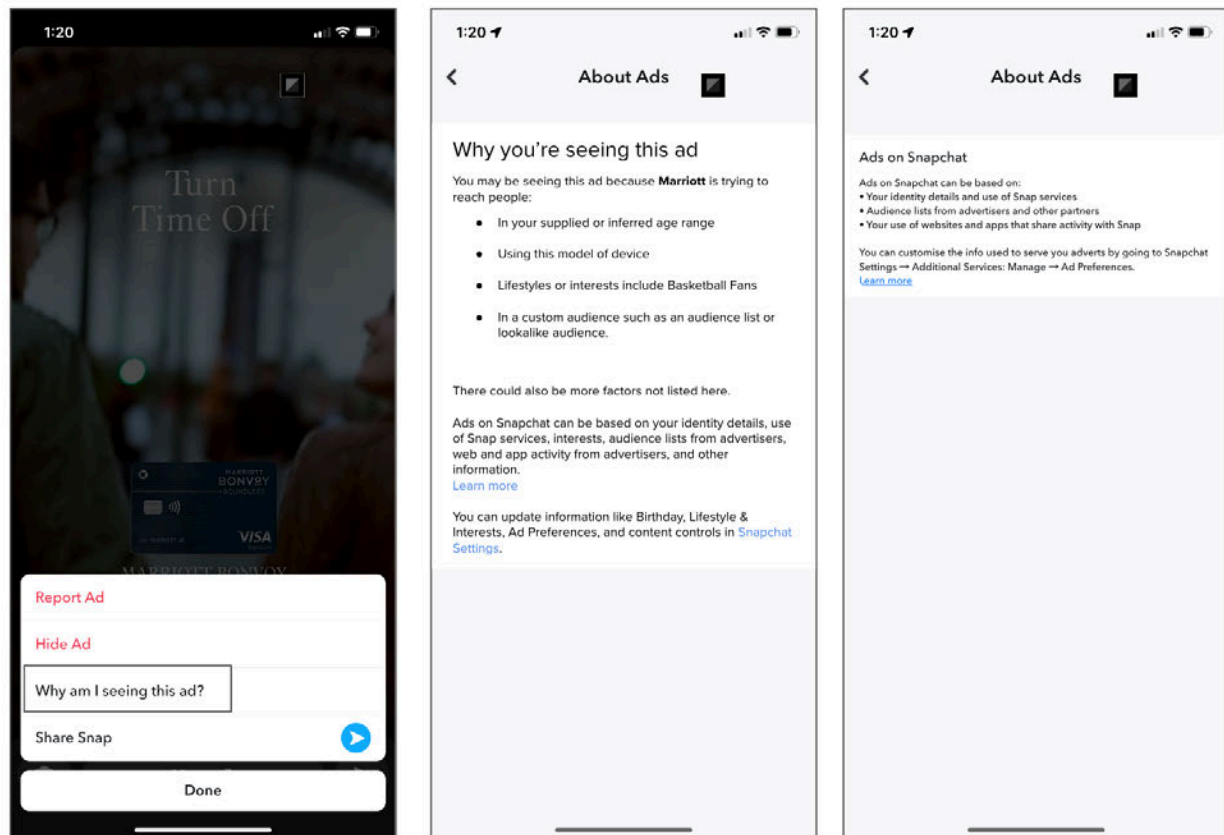


and UK, Norway and Switzerland. All users can restrict our use of third party data and being included in advertiser supplied audience matches for ads targeting.

We use pre-launch testing and our ad review process to help ensure these controls work as designed.

Why am I seeing this ad?

Snapchat also allows Snapchatters in the EU to click “Why am I seeing this ad” to see additional information on the targeting parameters for the advertisement in question.



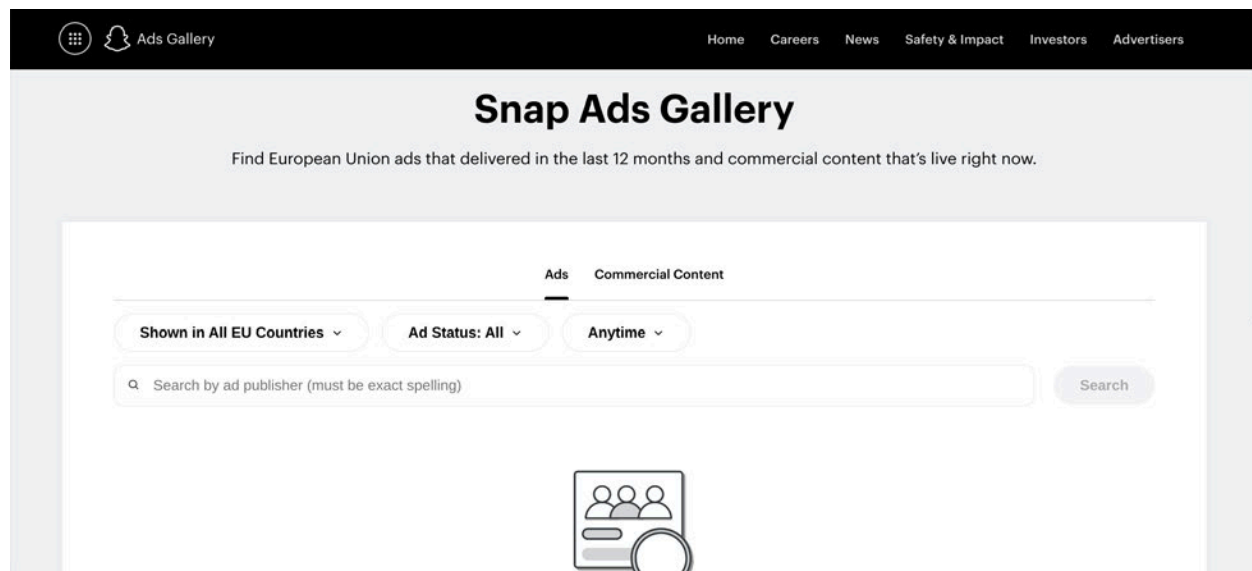
Lack of visibility – Ads Gallery

There is a higher risk that advertising will violate our terms or the law, in particular content misleading information, if the Snapchatter community and wider society does not have visibility into the history of ads over the past year that have run on Snapchat and some details about the targeting and reach of those ads.

Snap has an [ads library](#) (as required under Article 39 DSA) which provides increased transparency for ads - not just political - that are currently running, and historically have run in the past year, directed to EU users on Snapchat. This ads library is available to anyone, can be searched / filtered / sorted based on pre-defined parameters (e.g. country targeted, advertiser name, etc) and includes an API interface as well. This allows anyone to check who has paid for an



advert and, if different, on whose behalf is the advertisement being published. In the Commercial Content section of the Ads Gallery, we also include links to all live organic content that has been marked with the “Paid Partnership” tag. The Ads Gallery maintains advertisements for one year after the advertisement airs.



Users can search the Ads Gallery by ad publisher and can filter results by multiple criteria, including by country, ad status, and publication time.

Information included for each ad is shown in the screenshots below. When a user clicks on the “See Details” link they are taken to the Ad Details modal on the right. Per DSA guidelines, data includes:

Main Ad Modal

- Ad Publisher - the natural or legal person who paid for the advertisement
- Brand Advertised - the natural or legal person on whose behalf the advertisement is presented
- Ad Start Date and Ad End Date - the period during which the advertisement was presented
- Ad Creative - the content of the advertisement, including the name of the product, service or brand and the subject matter of the advertisement
- Total Impressions - the total number of recipients the service reached

Ad Details Modal

- Impressions by Member State - aggregate numbers for the recipients reached by country (if 0 recipients were reached, the ad will not appear).
- Targeted devices and demographics - whether the advertisement was intended to be



presented specifically to one or more particular groups of recipients, specifically, devices and demographics; These options do not support exclusion targeting.

Ads Gallery - Ads Data

Ad Publisher: Nike, Inc. Active

Brand Advertised: Nike

Ad Start Date: Mar 20, 2023

Ad End Date: N/A

Total Impressions: 5,368

[See Details](#)

Ad Details

Nike Active

Organization Charged
Nike, Inc.

Ad Start Date
Mar 20, 2023, 1:00 AM GMT-7

Ad End Date
N/A

Total Impressions
16,199

Austria	Germany
5,512	10,687

Total impressions include lifetime or the last 12 months of impressions for a given country, not unique impressions.

Demographics
18 - N/A

Devices
Android

Per the Advertising Review section above, we use a combination of automated and human review to prevent ads that violate our policies or the law from appearing on Snapchat. Ads that were delivered and subsequently taken down are marked as Rejected in the Ads Gallery.



Ad Details

Sloggi and Amazon Fashion EU Rejected

This Ad was rejected

Your ad contains content that could be considered inappropriate, including nudity, sexual content, body objectification, sexually provocative behavior or sexually suggestive imagery. This may include obscuring heads or faces, or focus on particular parts of the body unrelated to the advertised product or service. Ads for lingerie and other intimate clothing must not be overly sexual or depict nudity.

Organization Charged

Jellyfish UK - Amazon Fashion

Ad Start Date

Sep 11, 2023, 1:00 AM GMT-7

Ad End Date

N/A

Total Impressions

2,412,597

France	Italy	Spain
891,154	870,066	651,377

Total impressions include lifetime or the last 12 months of impressions for a given country, not unique impressions.

The Snap Ads Gallery is maintained by the Ads API and Ads Manager teams. The ads library underwent pre-launch testing to ensure it met design specifications and will continue to develop based on further testing.

Freedom of Expression

The purpose of Snap Ads is to amplify advertisers’ commercial messages, and as a result the content is rarely political or rather expressing views. We have specific procedures for political ads. As a result, the risk of a negative impact on freedom of expression from Snap’s other mitigations listed above is low.

5.7.5 Conclusion

Targeted advertising on Snapchat is necessary to ensure we can continue to provide a free service to all users regardless of their ability to pay. We have taken extensive steps to ensure our approach to targeted advertising appropriately balances the interests of Snapchatters, Snap and advertisers.

. We have also put in significant measures to prevent fraudulent and other advertising that may be harmful or against the law. We reject thousands of them each month to keep Snapchat’s community safe.

As explained in Section 4, we have concluded that our adaptation of Snapchat’s advertising systems described above, in combination with the other mitigations explained in this Section 5, are reasonable, proportionate and effective mitigation measures for the risks identified for Snapchat’s in-scope services.



5.8 Protection of Minors

5.8.1 Introduction

Snap’s utmost priority is the safety and wellbeing of our users, especially young people, aiming to ensure they have positive experiences on our service. Privacy, safety, and security are foundational tenets of the company and sit at the core of our value proposition to our users. Since Snapchat’s inception, we have embraced a [privacy and safety by design](#) approach and recognise that our platform architecture and product choices play a major role in the protection of our users, including minors. We put significant thought and consideration to ensure our values are reflected in the architecture of our platform, and in the design and implementation of our products, policies, and enforcement actions.

We take the protection of minors aged 13-17 (“Teens”) seriously on Snapchat. Our key tenets include acting in the best interests of Teens, offering strict default settings for all users, and respecting Teens’ freedom to express themselves safely, while recognizing their right to information about the world. We aim to achieve these tenets by positioning parents and guardians to help guide Teens in their responsible use of our platform, attaching a heightened safety interest to Teens using our products, and establishing processes to ensure we develop products in a way that upholds these tenets. We have implemented these tenets through the use of Family Center, focusing on age-appropriate content, reporting and blocking mechanisms, and putting in place appropriate protections and limitations on private messaging, friending, public content, and advertising.

5.8.2 Administration and Oversight



[Redacted]

[Redacted]	[Redacted]
[Redacted]	[Redacted] [Redacted]
[Redacted]	[Redacted] [Redacted]

5.8.3 Overview and Approach

[Redacted]

- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]

- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]



[Redacted]

- [Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]

- [Redacted]
[Redacted]
- [Redacted]
[Redacted]
- [Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]

[Redacted]
[Redacted]

- [Redacted]
[Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]

- [Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]

[Redacted]
[Redacted]



- [REDACTED]
- [REDACTED]

Age Appropriate Design Code

Snap has adopted the Age Appropriate Design Code (AADC) guidance for developers to incorporate appropriate and proportionate measures to safeguard the privacy, safety, and security of users aged 13-to-17 within platform products and features. This was first established by the Information Commission's Office (ICO) in the UK, but has been used as a global best practice standard in other countries.

This design code includes 15 core standards:

1. [Best interests of the child](#): The best interests of the child should be a primary consideration when we design and develop products.
2. [Data protection impact assessments](#): Product must be covered when appropriate by minor data protection impact assessment.
3. [Age appropriate application](#): Requires a risk-based approach to recognising the age of individual users to ensure we effectively apply the AADC standards.
4. [Transparency](#): The privacy information we provide to Snapchatters, such as our privacy center, support pages and in-app notices, must be concise, prominent and in clear language suited to the age of the child.
5. [Detrimental use of data](#): We should not use children's personal data in ways that have been shown to be detrimental to their wellbeing, or that go against industry codes of practice, other regulatory provisions or government advice.
6. [Policies and community standards](#): Uphold our own published terms, policies and community standards (including but not limited to privacy policies, age restriction, behavior rules and content policies).
7. [Default settings](#): Settings must be 'high privacy' by default (unless we can demonstrate a compelling reason for a different default setting, taking account of the best interests of the child).
8. [Data minimisation](#): Collect and retain only the minimum amount of personal data we need to provide the elements of our service in which a child is actively and knowingly engaged. Give children separate choices over which elements they wish to activate.



9. [Data sharing](#): Do not disclose children's data unless we can demonstrate a compelling reason to do so, taking account of the best interests of the child.
10. [Geolocation](#): Switch geolocation options off by default.
11. [Parental controls](#): If providing parental controls, give the child age appropriate information about this. If our online service allows a parent or carer to monitor their child's online activity or track their location, provide an obvious sign to the child when they are being monitored.
12. [Profiling](#): Switch options which use profiling 'off' by default unless there are appropriate measures in place to protect the child from any harmful effects.
13. [Nudge techniques](#): Be mindful of and avoid using nudge techniques to lead or encourage children to provide unnecessary personal data or weaken or turn off their privacy protections.
14. [Connected toys and devices](#): If providing a connected toy or device ensure we include effective tools to enable conformance to the Code.
15. [Online tools](#): Provide prominent and accessible tools to help children exercise their data protection rights and report concerns.

We have been actively supporting the efforts of the Commission and other stakeholders' to establish an EU-wide Age Appropriate Design Code and/or guidance on the application of Article 28, and to consider whether further mitigation measures may be reasonable, proportionate and effective for online platforms, 'gateways' and other online services. The Commission has since published its guidance on Article 28 DSA, and we remain committed to working closely with the Commission and others on its implementation. We have assessed the extent to which we meet the recommendations in the guidance in [Section 4.4.3](#) (Negative Effects on Minors) and have also referenced the guidance in [Section 5.12](#) (Codes).

Privacy, Safety, and Security of Minors on Snapchat

Privacy, safety and security are key priorities of the company and at the core of our value proposition to our users. Snap has dedicated extensive resources to incorporate protections aimed at safeguarding the rights of Teens on the platform, greatly reducing the likelihood of rights infringement. At the highest level Snap follows Age Appropriate Design Codes (or similar) established by the United Kingdom, France, California, etc, as well as our own key tenets described earlier in this playbook. Snap's approach to privacy and safety by design means that we generally design for our youngest users first and work upwards. This means our first layer of protection for minors includes the mitigations that are designed for Teens but apply to all to advance safety across our community.

Snap has put in place a range of mitigation measures to help protect the privacy, safety, and security of minors on Snapchat. This includes general platform safeguards such as our Teen friendly terms and support pages, our moderation and enforcement processes, our parental tools—Family Center, in-app reporting, and Teen specific content moderation and restrictions,



In addition to these safeguards for Teens, we consistently enforce our policies disallowing users under the age of 13 from creating or maintaining an account. Persons who are younger than 13 will be blocked from creating an account at the time of registration; accounts that are discovered to be operated by persons under the age of 13 are removed from the platform at the time that Snapchat discovers such violations.

Advertisements for Minors

Snapchat automatically disables advertisements based on profiling for users aged 13-17 in the EU. Additionally, safeguards have been put in place to help Teens understand and recognize Lenses and ensure that advertisers and advertisements on our platform comply with our requirements.

Snap has implemented additional safeguards and protections for minors related to advertisements, as described in detail in the product-specific subsections below.

Identifying Minors

As the creators of a central communications tool for young people, we take seriously our responsibility to protect teens on our platform. We know age verification is an industry-wide challenge everyone is trying to solve, and we are already working with industry peers, regulators, and third-party technology providers on possible approaches. We look forward to continuing these productive conversations to achieve methods that work for everyone.

Snap currently takes a risk-based approach to provide an age-appropriate experience across Snapchat, consistent with best practices such as the UK Age Appropriate Design Code (AADC). As explained in our response to the Commission's RFI on minors in December 2023:

Registration and access to Snapchat

In order to download the Snapchat app, users first need to create an account with either Apple or Google to access their app stores (Apple App Store and Google Play Store). Both the Apple App Store and Google Play Store have age restrictions, they require users to create an account before they can access the stores, and the age restriction for those accounts is 13+ and in some cases 14, 15 or even 16+ (see [Apple](#) and [Google](#) age restriction terms).

Both Apple and Google rely on declared age to determine if a user is 13+. If a user provides an age under 13 account creation is persistently blocked unless parental approval is provided. Both Apple and Google offer state of the art and easy to use parent tools (see [Apple](#) and [Google](#) family link terms). This means that in order to download an app - for example Snapchat - from the Apple or Google Play Store a user needs to declare to be 13+ or parental approval has been provided.

Although Snap has asked for access to Apple and Google's parent tools and age signal to ensure consistency, increase visibility of our guides and settings, reduce the burden on end users and



ensure a level playing field with Apple and Google's own apps, Snap does not currently have such access.

As a result, Snap independently asks the user to confirm their age as an additional age assurance measure, despite age already being provided by the same user as part of the Apple and Google account registration flow, as follows:

Declared age to limit access to Snapchat to its target 13+ audience:

- a. Our declared age process has been designed to meet industry standards.
- b. In our Terms of Service, Privacy Policy, and other documentation, we make clear that Snapchat is intended for users 13 years old or older. Users must affirmatively add their birthdate when registering for an account, and we deny users declaring they are under the age of 13 the ability to create accounts.
- c. If we determine, or are otherwise made aware through an in-app report from a user, parent, or law enforcement, that an account belongs to someone younger than 13, we take immediate action to prioritize and respond to the information. Our trained internal team will review and disable the account, including immediately deleting the data associated with the account.

In respect of a)

- We do not use inferred age techniques to prevent individuals under the age of 13 from registering or accessing the app. Reliable age inference is not feasible without data based on user activity once registered and engaging in the app. We do not have such activity level data for any new user at registration, nor do we have this data at any time for users under the age of 13 (since all Snapchatters are declared to be 13 or older).
- As explained below, we have stronger age assurance in place to protect minors from certain content and features of Snapchat targeted at more mature audiences which allow us to protect against potential U18 use of adult accounts despite not having absolute knowledge of U13 use. Our approach is stronger as it allows U18 protections to be applied if there are changes after registration, for example, in potential situations where a parent or other adult user may register an 18+ account using their own information but then provide account access and usage to an U18 bypassing any static age assurance applied on registration.
- Further, as shown in Section 1 ([Snapchat Community](#)) the vast majority of Snapchat users access the app to use our messaging services to communicate with friends - not too dissimilar from traditional SMS or other messaging services, which typically do not have any age gates at all. Such interpersonal communication services fall outside the scope of the DSA.



[REDACTED]

In respect of c) and our trained internal team:

- The team is trained to prioritize these tickets. When our privacy operations team receives a ticket they act upon this promptly, often within a couple of hours. [REDACTED]
[REDACTED]
[REDACTED]
- If our team is made aware of an account belonging to someone potentially under 13 through external sources (eg. through in-app reporting, Law Enforcement requests), a ticket is created and routed to our human Trust & Safety team who is also trained to prioritize these types of tickets. The team of moderators responsible for reviewing these reports consists of several dozen FTEs.²¹⁵ We do not track the specific response time for this type of ticket, these metrics are tracked across all reporting types. [REDACTED]
[REDACTED]
[REDACTED]
- We provide moderators with training sessions on policies, processes, tooling, current events and cultural norms to be effective at their work. Our moderators are trained through small group training classes and also review a multitude of scenarios while shadowing high-performing peer moderators. Through practice and instruction, they apply our policies and enforcement measures in a manner that protects our Snapchat community. This training is conducted over a multi-week period, in which the moderator is educated on Snap's policies, tools, and escalations procedures. After the training, moderators must pass a certification exam before being permitted to process content.
- In addition to the general moderation training (see below), these team members receive specific training and guidance from our (privacy) legal teams. [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

In respect of c) and the use of keywords to detect under age users:

- In general, Snap does not scan keywords during account creation or from account information to detect whether a user might be underage. As noted in Section 1 (Snapchat Community), more than 80% of time spent is on private surfaces. Snapchat is primarily intended and used as a communications tool by our users with their close friends and family, and fundamentally different from traditional social media where the majority of the content is public. We apply a privacy-first approach to user communications, such as Chats, and those are not subject to scanning for purposes of learning a user's age or

²¹⁵ For a more detailed breakdown of our human moderators please see Section 5.4 (Content Moderation).



profile. We assessed that doing so would be contrary to the fundamental privacy rights of individuals, as well as privacy laws, including the GDPR and ePrivacy Directive.

- However, since our 2023 Report and the RFI, we have begun testing an additional mechanism to detect under age registrations and access to Snapchat. This involves scanning five text fields within user public profiles for certain key phrases:

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

We assessed the scanning of public profiles to be less intrusive as this content is public, and on balance would not infringe users' fundamental privacy rights. Where clear statements are made that the user is under age, the corresponding account is further investigated and appropriate enforcement action taken (for example, to remove the account).

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Access to certain content / features

We rely on a combination of declared age and inferred age techniques for stronger age assurance to limit under 18 access to certain content and features targeted at more mature audiences.

2. A combination of declared- and inferred-age techniques for stronger age assurance to protect minors. We not only rely on a users' self-reported age, but have techniques to infer a users' actual age, which considers a combination of various influential signals [REDACTED]

[REDACTED]

to limit under-18 access to certain content and features targeted at more mature audiences, including:

- a. Discover (Publisher Partners): At the time of onboarding, Snap provides its [Commercial Content Policy](#) to Publisher partners. Publishers must adhere to Snap's Commercial Content Policy. This Policy requires Publishers to ensure their content is appropriate for a 13+ audience. If it's not, our policy requires publishers to age-gate their content. Publisher partners are also given more detailed guidelines on how to comply with the policy (including with respect to age-gating). In addition to this policy, as explained in Section 5.4 (Content Moderation) our global moderation team reviews Publisher Tiles for compliance to ensure content shown in "Stories/Discover" meets our [Content Guidelines for Recommendation Eligibility](#) (our standard for "appropriate content"). Our Partnerships team also performs periodic checks of Publishers and Content Creators to ensure that the content they are posting is compliant with these guidelines. Publisher



partners receive notification (either emailed or in-app, depending on the creator type) if their content is found in violation of our guidelines. Publisher partners who violate frequently and/or severely are further penalized, after editorial review. These penalties can involve a period of restricted visibility, a suspension during which all publishing is banned, or a permanent channel ban. In rare instances, we've ended relationships with entire organizations. In addition to penalties, Snap regularly communicates updates and clarifications to our guidelines in emailed newsletters to partners, and holds periodic training seminars for publishing partners.

- b. Spotlight and Discover (UGC):
 - i. Content that is prohibited by our [Community Guidelines](#) is prohibited everywhere on Snapchat and our [Content Guidelines for Recommendation Eligibility](#) specifies additional categories of content that, while permitted on Snapchat, will not be eligible for recommendation to a wider audience on Spotlight and Discover. These mitigation measures apply to all Snapchatters.
 - ii. In addition, Spotlight and Discover have been designed not to distribute sexually suggestive content to 13 - 17 Snapchat accounts and only recommends sexually suggestive content to a 18+ Snapchat account if that content has been created by a creator that the account has subscribed to, or if the user repeatedly favorites suggestive content. This uses our combined declared and inferred age techniques.
 - iii. Regarding the machine learning classifiers, we use in-house classifiers that were trained [REDACTED] to scan and identify sexually suggestive content using state-of-the-art computer vision models. When user-generated content on Discover and Spotlight is scanned by our machine learning classifiers, content that is scored above our threshold and considered "suggestive" is then removed from the content recommended to teens. Snap assessed (and continues to assess) the effectiveness of our machine learning classifiers via (i) quality testing and product/engineering review before deployment and (ii) ongoing review against in-house human labelling of publicly available content. We do not deploy new machine learning classifiers until they achieve at least 80% precision. Note that we also conduct routine quality checks of our human review where our precision is 95%+.
 - iv. Our systems are designed in a way that a Snapchatter (who is over 18 years old) who has subscribed to a content creator that has posted suggestive content will not see more than one sexually suggestive content video out of seven in their Spotlight feed. If a user hides a video labelled as sexually suggestive or a sexually suggestive creator ([here](#) is an example of a piece of Spotlight content that was marked as suggestive), we stop showing that type of content to that user. If a user hides a creator, we stop showing them that creator.
- c. Lenses: We age-gate certain Lenses (e.g., related to alcohol, gambling, NFTs, etc.). For example, a Vivino Lens will only be shown to 18+ users in the EU, mitigating the risk of such content being shown to users who are under the legal drinking age.
- d. Ads: We restrict ads based on the user's age. For example, ads for dating services must be targeted to users over 18 and must not be provocative, overtly sexual in nature, or reference transactional companionship. Similarly, ads for alcohol products must be



[REDACTED]

age-targeted to at least 18+, or the applicable minimum drinking age in the respective country where the ad is running.

Our age inference model is used as an integral part of our age assurance method to limit under 18 access to content and features targeted and suited to more mature audiences. The inferred age model on Snap uses a variety of influential signals [REDACTED]

[REDACTED], rather than only relying on the age that users provide when signing up to the platform. This helps, for example, prevent regulated ads from being served to those users who have declared themselves to be over the age restriction but we have modelled as likely to be under the appropriate age for such regulated content.

For example, if a user's self-declared age is 20 years old, yet the signals derived from the user's activity within the app and the ages of their friends strongly indicate that they are likely under 18 years old, we can internally "override" their supplied age and flag the user as a minor, and therefore filter regulated ad content (e.g. alcohol) from being displayed to them.

The inferred age model does not utilise any third party age assurance providers. To minimise disclosure of personal data, the model is processed within Snap. [REDACTED]

[REDACTED]

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

As flagged above, we do not use inferred-age techniques to prevent individuals under the age of 13 from registering or accessing the app. Reliable age inference is not feasible without data based on user activity once registered and engaging in the app. We do not have such activity-level data for any new user at registration, nor do we have this data at any time for users under the age of 13 (since all Snapchatters are declared to be 13 or older). The use case for our age inference model has always been geared toward mitigating the risks for users under 18, for example to ensure users under 18 do not see regulated ads, such as ads for alcohol. As Snap isn't capable of determining that a user is under 13 from our inferred age model, we have never applied it for the purpose of preventing users from registering an account. In line with our [COPPA](#) obligations, if Snap receives actual knowledge that a user is under 13 (for example, from a verified parent's request), Snap promptly deletes the account and associated data.



Oversight

We have a dedicated working group overseeing our age-assurance efforts. This cross-functional group consists of 10+ FTEs with representatives from the product, privacy legal, product legal, policy, and trust and safety teams. These team members bring to the table extensive experience and knowledge in the areas of operations, policy, global privacy laws and regulations, privacy-focused product decision making and online safety.

Different iterations of this group activate as needed to meet in working groups that are focused on exploring, discussing and assessing possible ideas and concepts and solutions related to age assurance through risk assessments, discussions with our product team, and collaboration with external legislators, regulators, peer platforms, vendors, experts, NGOs and other stakeholders.

For example:

- This group regularly meets to discuss Snap's age and parental assurance methods, and the legal and regulatory requirements in this space.
- This team has assessed current industry practice with regards to age and parental assurance, including in particular the mechanisms of: (i) device operating systems Google [Family Link](#), Apple [device parental controls](#) and [Family Sharing controls](#) and [Microsoft Family Safety](#) and (ii) other online services such as Whatsapp, Tiktok, Instagram and Youtube.
- The team regularly engages in stakeholder meetings, such as those organised by the [Centre for Information Policy Leadership \(CIPL\) Privacy](#) in Europe.
- This group also advocates for a holistic approach to age and parental assurance. An example of this would be our work with the UK Government and House of Lords during the passage of the Online Safety legislation in Parliament to successfully achieve a requirement for Ofcom (the UK's communication regulator responsible for regulating the UK's Online Safety Act) to consult holistically on child safety and age assurance, including considering the role that infrastructure such as app stores / device operating system accounts have in providing privacy friendly, secure, effective and efficient solutions.
- This team has also met with a number of industry leading third party age assurance providers to assess the technical, legal, financial and user impacts of those services being integrated into Snap. These efforts have proven successful and both Google and Apple have recently announced APIs to share age signals with app developers. We hope to build on this work to: (i) develop a comprehensive industry wide framework, with CIPL and the We Protect Alliance and (ii) require Google and Apple to share further first party age signals/software pursuant to their obligations under the Digital Markets Act.
- This team has reviewed research from external stakeholders, on the positive and negative impacts of age and parental assurance, as well as consulting with users of Snapchat (including our own families and friends).



- Representatives from this group presented to Snap's Safety Advisory Board²¹⁶ on Snap's approach to age assurance and Family Center, including potential options and challenges with age assurances. On age assurance in particular, this group of experts advised us that:
 - This was an industry wide issue that needed broad stakeholder discussion and need for service independent solutions (i.e. considering devices, app stores etc) that best met the needs of children and parents/responsible adults.
 - They also felt our 13+ age limit was inhibiting our ability to recognise and keep potential younger users safe and advised us to consider lowering the age so Snapchat's U18 experience was available to those that wish to use and less incentive for children and parents/responsible adults to try to access via 18+ accounts / bypass age assurance methods. They recognised the challenges that COPPA presents in this regard.
 - They felt it was better to have a safe platform for all than to rely on excessive age / parental assurance. This would limit the need for age / parental assurance to the smaller number of mature areas.

Ongoing evaluation

Snap continues to evaluate its approach, and consider possible concepts and approaches with industry peers and third-party age-assurance vendors, to ensure we keep pace with industry practice. We are also supporting legislators and NGOs in the UK, France, and elsewhere in the EU to enhance the role of app stores, online devices, and web browsers in providing appropriate interfaces for age assurance and parental controls to facilitate consistent, effective and efficient approaches for the online ecosystem.

Snap earned praise over years by the Children's Advertising Review Unit (CARU) in the United States for exceeding standard protections to keep underage users off Snapchat, and for providing numerous safeguards for our users once they are on our platform. Despite this, we continue to reassess our age assurance efforts, including engaging with industry partners, regulators, and third party age assurance vendors, to ensure we keep pace with developments in the space. We actively participate as thought leaders in industry roundtable forums, as well as with policy makers in the UK, US and the EU to understand the evolving landscape of age assurance, and its critical importance coupled with its inherent challenges. We've also held multiple exploratory and deeper dive sessions with leading vendors across the age estimation and identity verification marketplace in recent months, as we consider third party technologies including biometrics, ID scan, and financial transaction methods (among others), to enhance our current approach. As noted above, these efforts have been fruitful as both Google and Apple intend to introduce APIs to share age signals with app developers. As above, we hope to build on this work to require Google and Apple to share further first party age signals/software through industry collaboration and/or pursuant to their obligations under the Digital Markets Act.

²¹⁶ The Safety Advisory Board is explained in Section 6.7 (Snap Advisory Groups).



In terms of parameters, Snap consistently focuses on the potential impact to key areas when evaluating the effectiveness of age assurance measures. Such factors include the tradeoff between safety of minors and compromising user privacy/data security, the accuracy and reliability of age estimation technology (particularly for younger or ethnic minority populations), the fairness of methods that may disadvantage users without official government IDs or bank accounts, and the harm to industry competitiveness from the exorbitant cost of adopting third party technology at scale.

It is also worth noting that many organizations have concerns about the effects of introducing age assurance and support device OS account and app store based solutions for age assurance to support online platforms’ own mitigation measures, for example ICMEC and NCOSE in the US. Prominent European children’s NGOs have expressed similar support when meeting with Snapchat. To that end we’ve seen in the US that a growing number of states have passed or introduced legislation that requires app stores to perform age verification of all users and to share such age signals with app developers.

We remain focused on thoughtful enhancements to our risk-based, age and parental assurance approach that include balancing the need for safety, accuracy, fairness, and user privacy among other important factors and taking a holistic view of the online ecosystem used by children and their parent(s) (or other responsible adult(s)). Taking account the results of our risks assessment set out in Section 4 and current industry wide practices, we continue to conclude that our approach is proportionate, reasonable and effective.

European Commission Art 28 DSA Guidelines

Considering the mitigation measures outlined throughout this section and the Risk Assessment, and taking into account the recently issued Guidelines on the application of Article 28, we have carried out an analysis of these recommendations and assessed our organisation’s position in relation to them. This analysis seeks to map existing measures against the expectations set out in the guidelines and evaluate whether the platform’s current practices align with the requirement to ensure a high level of privacy, safety, and security for minors. The table provides an overview of the obligations, which serves as the benchmark for our assessment.

Please note that we are conscious that not all recommendations are currently implemented; in some cases, we consider our existing measures to be proportionate and effective, while in others further steps remain under consideration. In any case, we value the guidelines, which - though not mandatory - provide useful direction on best practices. At the same time, it is important to recognise that each platform and online service has its own nature, and therefore measures should always be applied in a manner that is proportionate and adapted to be effective in those specific circumstances.

Section	Summary of the Guidelines	Snap current measures
---------	---------------------------	-----------------------



General Principles	Platforms should adopt measures proportionate to their specific risks, always balancing safety with fundamental rights. They should respect Children's Rights under the Charter and UNCRC, embedding privacy, safety, and security by design from the outset. Finally, services must follow age-appropriate design, adapting features to minors' cognitive and emotional development.	As explained in Section 6.3., Snap embeds privacy and safety by design and follows Age Appropriate Design Codes from jurisdictions like the UK, France, and California, as well as its own framework. Guided by key tenets - acting in Teens' best interests, enabling safe expression, ensuring access to information, supporting parental guidance, and applying heightened safeguards - Snap designs products with the youngest users in mind first. These protections for Teens form the baseline safety layer applied across our platform, as explained in the mitigation section of this section.
Risk Review	Organizations should systematically assess how their products and features impact Children's Rights by identifying risks through the OECD's 5Cs typology (Content, Conduct, Contact, Consumer, and Cross-cutting risks), mapping them by likelihood and severity, and evaluating both positive and negative effects on minors. These assessments should be reviewed annually or after significant changes, such as the introduction of AI features or recommender systems, and should actively involve minors, guardians, child-rights experts, and independent stakeholders. To ensure accountability, the findings should be shared with regulators and published transparently.	Snap already conducts (and publishes, when required) systematic risk assessments, mapping them by likelihood and severity, as evidenced in this report. This approach is applied not only to meet the DSA requirements for VLOPs, but also to comply with similar regulatory frameworks in other jurisdictions, such as the UK's Online Safety Act.
Service Design	Platforms should focus on robust age assurance, with a preference for secure verification methods over self-declaration. Estimation techniques should be used cautiously, ensuring transparency, accuracy, and data minimisation. Access to high-risk content	In line with Article 28 requirements and the Commission's guidelines, Snap has implemented a series of service design measures aimed at protecting minors. Age Assurance: Snap applies a risk-based approach to age assurance, building on its



	such as pornography, gambling, and addictive features should be strictly limited to verified adults, with multiple assurance methods and appeals processes in place. Registration flows must remain simple and child-friendly, collecting only essential data and avoiding nudges that encourage underage users to bypass restrictions. To protect minors after registration, account settings should default to the highest levels of privacy, such as private accounts, disabled geolocation, and restrictions on unsolicited contact. Interfaces must avoid manipulative design practices (like infinite scroll and dark patterns) while offering tools for time management and friction nudges. AI features, including chatbots and filters, should remain optional, transparent, and easily disabled. Recommender systems must be regularly tested to prevent amplification of harmful content, prioritize explicit user choices over profiling, and allow minors to reset feeds or opt out of profiling altogether. Additionally, commercial practices should ban exploitative marketing tactics, mandate clear advertising labels, and safeguard children against hidden or manipulative in-app purchases. Finally, content moderation should be clear, responsive, and continuous, combining human and AI systems to detect harmful content across languages, prioritize minors' reports, and prevent harmful AI-generated interactions such as grooming prompts.	13+ entry threshold. Snap continues to evaluate its approach in collaboration with industry peers and age-assurance vendors. While current measures are deemed proportionate, Snap actively participates in policy discussions in the EU and other jurisdictions to strengthen consistency across the ecosystem (e.g., through device-level or app store-based solutions). Defaults and Account Settings: Snap ensures that Teens' accounts default to the strictest privacy settings - including limited visibility, restricted friending, disabled geolocation sharing by default, and controls on who can communicate with them. Features such as Family Center allow guardians to monitor interactions in a supportive, transparent way. These design choices align with the principle of acting in the best interests of minors while still preserving their right to self-expression. Interfaces, Tools, and AI: Snap has implemented product review processes with the aim to identify and prevent manipulative design patterns. Although generally out of scope of this Report, we note that with the introduction of AI tools, Snap has applied safeguards such as content moderation, user warnings, and ongoing red-teaming with HackerOne to identify vulnerabilities. Recommender Systems and Content Moderation: Snapchat's recommender systems (e.g., Spotlight, Discover) are designed to reduce exposure to harmful content. Moderation is carried out through a hybrid human-automated model, with 24/7 coverage, trusted flagger partnerships, and escalation systems. Moreover, on Discover, our policies are written with the understanding that people as young as 13
--	---	---



		may be viewing the content and should be age appropriate. Snap also invests in moderator training, support, and ongoing upskilling to ensure accurate and safe enforcement. Commercial Practices: Snap has adapted its advertising systems to prohibit exploitative or harmful practices. This includes bans on targeted ads to minors, the rejection of thousands of harmful ads monthly, and the use of strict review processes to block hidden or manipulative marketing practices.
Reporting Support & Guardian Tools	The guidelines recommend platforms to provide clear, accessible, and age-appropriate tools that allow minors to report harmful content, seek redress, and access support. Reporting mechanisms should be visible, simple, multilingual, and include feedback loops so children understand outcomes. Complaints must be free, timely, and confidential. In parallel, platforms should offer support features such as block, mute, “show me less” options, and links to external helplines. They should also provide guardian tools to help parents guide minors, but only as a complement to safety- and privacy-by-design protections, ensuring Children’s Rights to privacy, safety, and autonomy remain central.	As explained in Section 5.4 (Content Moderation) Snap provides in-app reporting tools that are simple and accessible for all users, including Teens, allowing them to flag harmful content, unwanted contact, or underage accounts. These reports are prioritized when submitted by minors and handled by a 24/7 Trust & Safety team, which combines automated detection with trained human moderators. Beyond reporting, Snap offers supportive features such as Heads Up, which surfaces expert educational resources when users search for sensitive topics, as well as easy-to-use block and mute functions and access to external hotlines. Guardians are supported through Family Center, which enables oversight of Teens’ connections and interactions, and a dedicated Parents’ Site with guidance and resources.
Governance	Organizations should establish child safety policies and dedicated teams with senior oversight, ensure staff are trained on Children’s Rights, and foster child	Snap has put in place a governance framework to oversee DSA compliance (which includes child safety), led by a Cross-Functional DSA Governance Team



	participation in design and safety processes. They must monitor compliance regularly, share data on risks and mitigations, and keep terms and conditions transparent and adapted for minors. Ongoing evaluation with input from children, guardians, and experts is required, while transparency obligations call for clear, accessible explanations of systems and tools using plain language and visuals.	and an Independent Compliance Function with direct access to senior management. As explained in Sections 6.6 and 6.7, Governance is further reinforced through external input from the Safety Advisory Board and the Council for Digital Well-Being, as well as systematic monitoring via prevalence testing and the Digital Well-Being Index.
--	---	--

Transparency to Minors

We continue to make efforts to provide users with information regarding our services in a way that is clear and comprehensible across age groups. We verify readability of our key terms and conditions and privacy notices using automated readability tools. The vast majority of these documents are shown to be understandable for our users. Our main terms and conditions is, necessarily, a formal legal document and contains longer provisions and more complex language which our automated readability tools indicate may be more difficult for our younger users to understand. To improve readability in particular for our younger users, we arranged the terms and conditions into sections that provide a sensible flow, including appropriate and succinct section headers, and added short summaries at the bottom of each section. We have tested the readability of these short summaries, and confirmed they are understandable for our users, including Teens. In addition, we have provided short explainers for our Community Guidelines to facilitate user understanding of this important document and know what they should and should not be doing on Snapchat.

Our Privacy Center was designed for our youngest Snapchatters and was intentionally developed to be easy to read and understood by all members of our community. We created our privacy and safety hub, with pages such as our Privacy by Product page, to give Snapchatters a high-level summary of our privacy and safety practices across each of our products and features. We also created a video to visualize our privacy practices, and use icons and other best practices as recommended by privacy and safety experts and the recognised Age Appropriate Design Codes.

As outlined in the Introduction, our EU Snapchat community consists of a diverse range of ages and genders. Snapchat services are not primarily directed at or used by minors. While Snapchat does have a young demographic, only a relatively small percentage of European Union users fall within the 13-17 age category. The largest age category of European Union users falls within the 18-24 age category.



5.8.4 Safeguards

In addition to our defaults for all users, we have added protections in place for Teens, to help mitigate risks in a number of ways.

App Store Level Safeguards

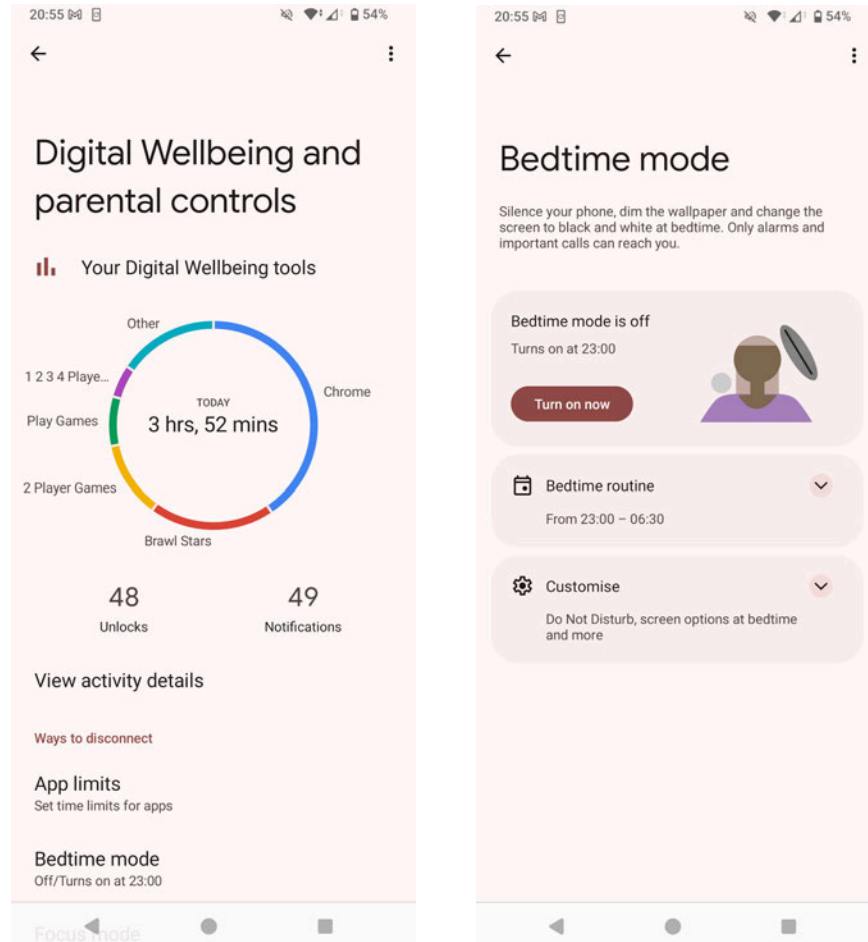
App stores contain and facilitate a vast range of apps presenting a wide array of different risks and representing the entire spectrum of online risks. They have a special place in the ecosystem and therefore a uniquely high risk. They should thus be considered as ‘high-risk situations’. In addition to the specific risks presented by the individual apps hosted, app stores can generate higher and exponential risks for a minor than the ones created by each individual service (e.g. a minor accessing harmful information on different services and combining harms).

Larger app stores already recognise that they present unique risks and require additional mitigations. In fact, larger app stores already apply age gates to prevent users from downloading an app if their app store account age is below the app’s minimum category specified by the app provider (and where applicable parents via the device operating system’s account level family controls - see below). Like most online platforms, app stores usually rely on the app store account’s declared age (which is often the same account used by the device operating system). We also have noted that app stores rely on predetermined age categories which do not necessarily capture nor fit the age categories defined by Snapchat and other developers at the app level. A classical example is the app store category 12+, which does not align with the age threshold of 13+ that is commonly specified by application services.

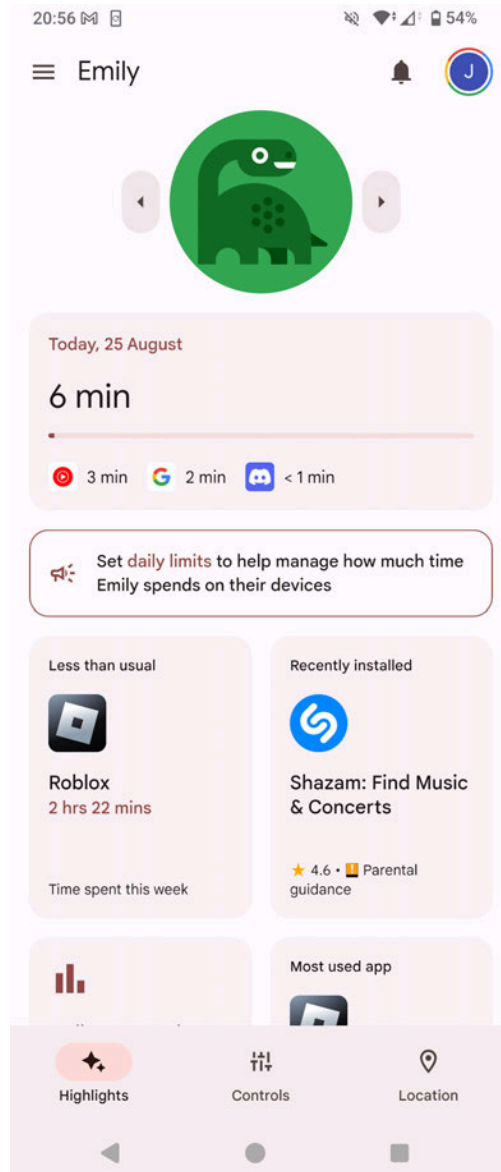
There are improvements which could be made by app stores providers for the benefit of the entire ecosystem including: (i) stronger age assurance at the app store / device (i) OS account level and sharing that signal with developers to support their minor protection measures; and (ii) allowing developers to set a more precise minimum age.

Device-Level Safeguards

Many additional controls are provided for teenagers, parents and other responsible adults. Many devices now come with wellbeing settings, such as bedtime mode that turn off device and app notifications and turn the screen the black and white to encourage sleep.



Additional controls are also provided via the device operating system's account level family controls (e.g. Google [Family Link](#), Apple [device parental controls](#) and [Family Sharing controls](#) and [Microsoft Family Safety](#)). For example, via these controls, parents and other responsible adults are able to view usage, set time limits, and disable access for each app which the teenagers have on their device.



We have also noted that the providers that operate these family controls (who are also gatekeepers pursuant to the EU Digital Markets Act) provide deeper levels of visibility and control for their own first party services. This level of interoperability and access would be very helpful for our own Family Centre (which is explained below) as it would increase the awareness and accessibility for parents and other responsible adults who may not have a Snapchat account. As explained below, we are actively encouraging further multi-stakeholder dialogue to have drive solutions that provide equal access and interoperability across the industry.

Platform-Level Safeguards

There are several protections that we put in place at a platform level to mitigate the risk of malicious users of Snapchat. In particular, we have inference models in place that look at platform



[REDACTED]

wide meta-data signals to identify suspicious accounts. We use this information at a product level to implement additional safeguards for Teen and adult users.

Product-Level Safeguards

In addition to our defaults for all users, we have added protections in place for Teens, to help mitigate risks in a number of ways.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



Public Content

Once users decide to share a Snap via My Story, by default only friends can view it. Snapchatters can choose to share to everyone, only to friends, or to a customized few. This emphasis on sharing with friends and giving users controls over who can view their content is in line with how Snap takes into account privacy and safety when designing its features.

[REDACTED]

[REDACTED]

[REDACTED]

Teen stories are deleted by default. Their My Story view setting is defaulted to friends only. Friends lists are private.

Viewing Public Content

Teen accounts are restricted from access to certain content that is generally considered suitable for 13+ but may contain certain shocking or sensitive content some may not find appropriate.

[REDACTED]

[REDACTED]



Spotlight

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED] We have developed machine learning classifiers which work to identify sexually suggestive content and filter it from the experience before human intervention. In addition, our Spotlight content is evaluated by human moderators upon reaching a threshold number of views, and before being even more widely distributed. These steps reduce the likelihood of Teens accessing illegal or violating content, or content that may negatively affect their rights, security and health.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

We also aim to prevent older users from seeing content from younger users and to protect Teens from being contacted by older users. We seek to achieve this by, for example, implementing the following measures:

- We limit the recommendation of content created by Teens to older users
- Adults cannot comment on Teen's Spotlight content on Snapchat.
- Users can also choose to disable comments on any post.
- Teens are protected on Spotlight by not having their usernames displayed.

Snap Map

As under 16 users cannot have Public Profiles, they will not have their Public Stories featured on Snap Map when tagging a place or venue to a Public Story (which would occur for 16+ accounts). Public posting options for 16-17 accounts are more [limited](#).

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



[REDACTED]

[REDACTED]

Additionally, Map filters out suggestive content from being recommended to users ages 13-17 and age-gates certain types of locations to prevent them from showing on maps for minors, including bars and tattoo parlors. Moreover, Snapchat displays a "low mutual friends" warning when a user is about to share their location with someone who has few or no mutual connections. This safety feature helps users make more informed decisions by flagging potentially unfamiliar individuals before location sharing occurs.

Advertisements

We restrict ads based on the user's age. For example, ads for dating services must be targeted to users over 18 and must not be provocative, overtly sexual in nature, or reference transactional companionship. Similarly, ads for alcohol products must be age-targeted to at least 18+, or the applicable minimum drinking age in the respective country where the ad is running.

Advertisers must comply with our Ad Terms, Advertising Policies, and applicable national advertising codes. We prohibit ads that address or intend to appeal specifically to children under the age of 13.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

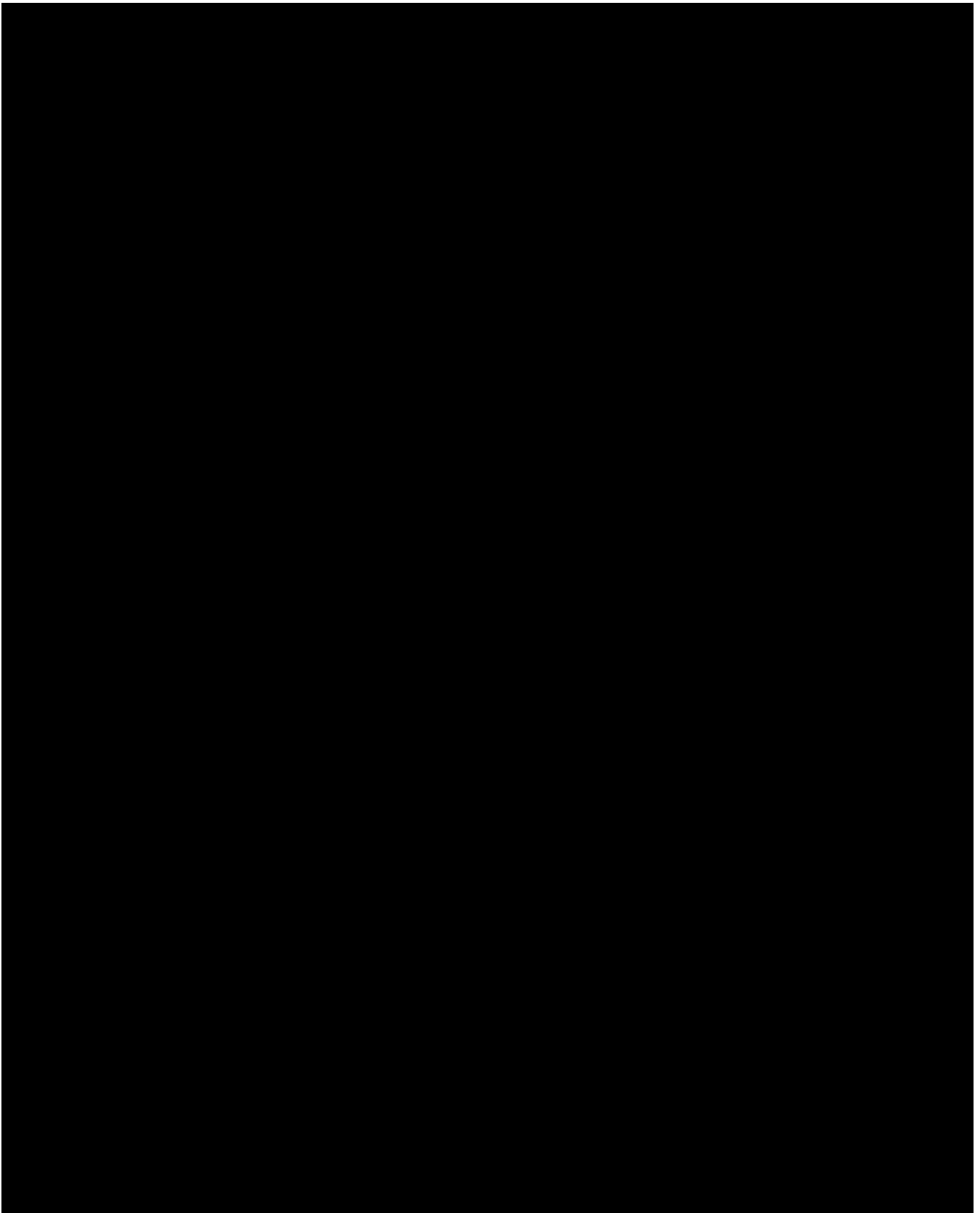
[REDACTED]

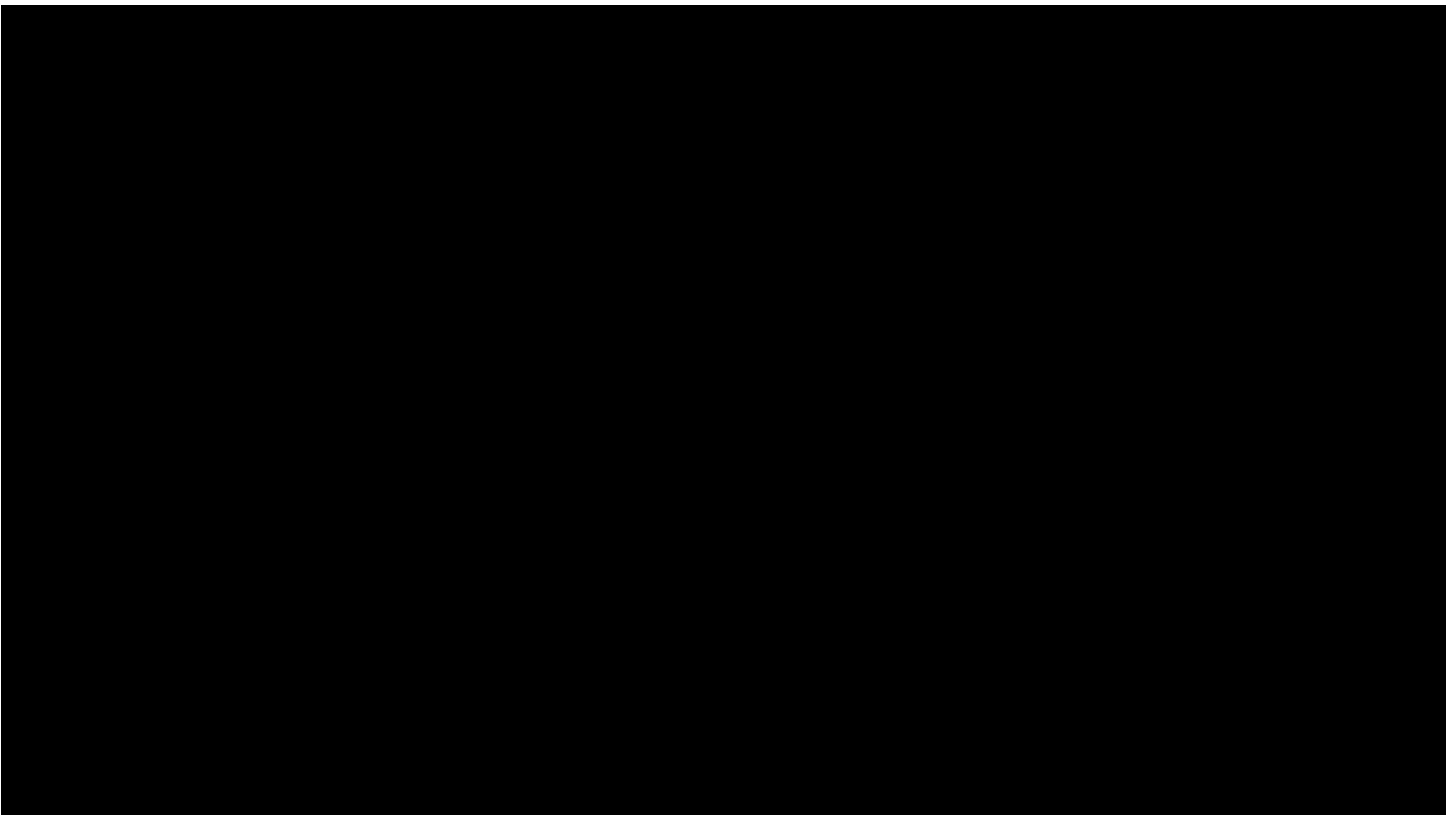
[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

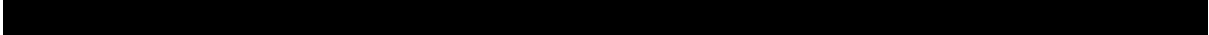




Our Ads API automatically prevents targeted ad types (such as pre-defined audiences and custom audiences) from being approved for ads directed at users ages 13-17. Ads API also prevents personalized optimization goals from being used for ads reaching Snaphatters aged 13-17 in the EU and UK (such as video views and story open).

Ads Restriction

We also restrict ads based on the user's age. 



Reporting and Blocking

- In-App Reporting: Teens have the ability to report abuse they may observe or experience within Snapchat. They can easily report Snaps, Chats, Stories, and Accounts by navigating to the clearly marked “Report” option in the menu on each of these feature screens or by pressing and holding on the content itself. Users follow our simple reporting flow and provide their reason for reporting and any additional comments that might be relevant. Moreover, through the "My Reports" section in the Account Settings, teens can view and track the status of the safety reports they've submitted. Reports are reviewed by our Trust & Safety teams that operate 24 hours a day, 7 days a week, and violating content and accounts are subject to enforcement. See the [Enforcement](#) section of this Report for more information.



- Blocking: All users have the [option to Block](#) another user. This prevents the friend from viewing friend content posted by, or sending Snaps and Chats to, the blocking Snapchatter. Since our 2023 Report, in an effort to prevent bullying and potential repeat harassment, we have introduced improvements to our blocking tools: Blocking a user will also now block new friend requests sent from other accounts created on the same device.²¹⁷
- Removing Friends: All users also have the [option to remove a friend](#) from their friends list. Once removed, the Snapchatter will no longer be able to view content accessible only by friends and, by default, should not be able to Chat or Snap.

Private Messaging

- [REDACTED]
- [REDACTED]
- [REDACTED]

Friending

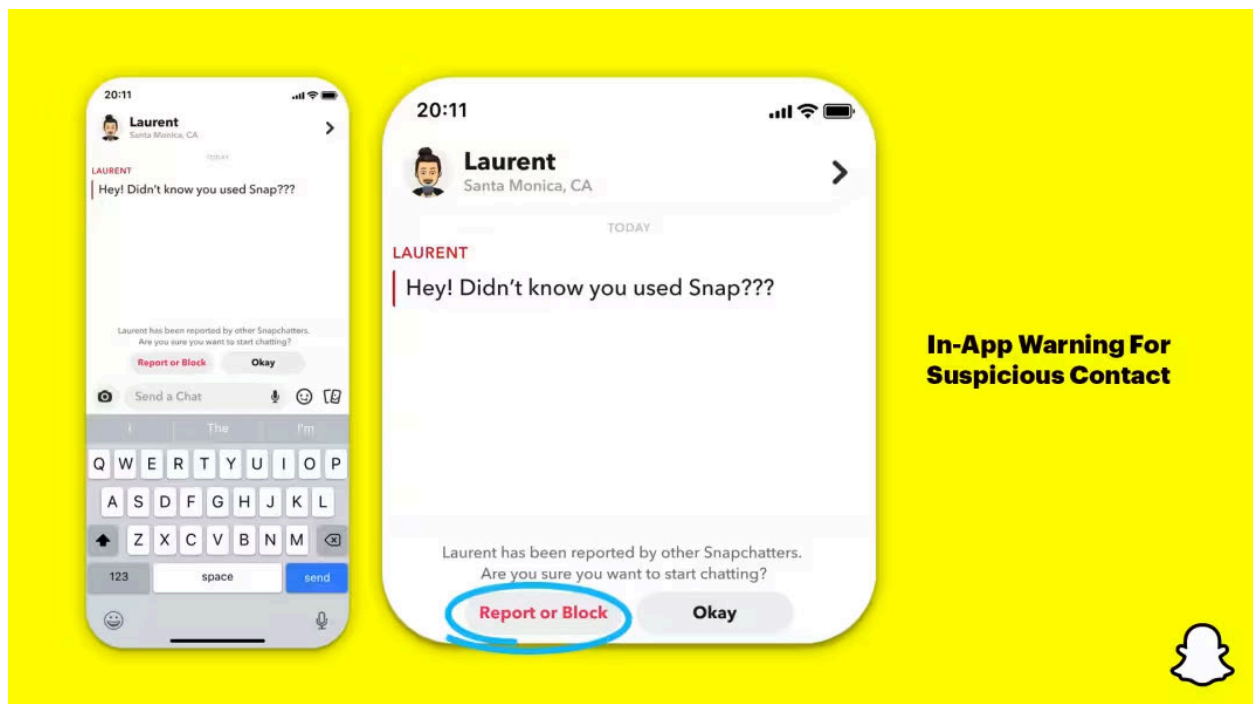
- [REDACTED] Similar protections apply to prevent Teens from searching for unknown adults. We prevent delivery of a friend request altogether when Teens send or receive a friend request from someone they don't have mutual friends with, and that person also has a history of accessing Snapchat in locations often associated with scamming activity.²¹⁸
- [REDACTED]
- By default: Users need to accept bi-directional friend requests or already have each other in their contact book to start communicating directly with each other. This design decision adds friction and prevents users from communicating with each other prior to accepting a friend request or being in one's contact book.

²¹⁷ <https://values.snap.com/news/new-features-to-help-protect-our-community>.

²¹⁸ <https://values.snap.com/news/new-features-to-help-protect-our-community>.



- No Public Friends Lists: Once users have accepted friend requests, the friend lists remain private. Snapchat does not disclose the friend lists of users to other users, nor do we expose the total number of friends that a user has. This protects the privacy of the user and their friends. On most other platforms friend lists are public by default or there is an option to share them publicly. These types of features create the ability for strangers to contact vulnerable groups (e.g. younger users).
- Friend Check-Up: Prompts Snapchatters to review their friend lists and remove those they are no longer in contact with, keeping their network up-to-date and focused on close friends.
- In-App Warning: We provide pop-up warnings: (1) when a teen receives a message from someone they don't already share mutual friends with or have in their contacts²¹⁹ and (2) if they receive a chat from someone who has been blocked or reported by others, or (3) is from a region where the teen's network isn't typically located.²²⁰



Family Center / Parent Tools

Our in-app parental supervision tool, Family Center, gives parents, caregivers, and other trusted adults visibility into their teens' friends list and who they have messaged with in the last seven days, as well as the ability to: (i) restrict their teen's access to Spotlight and Discover content tagged as 'sensitive' by our moderation team, (ii) disable their teen's ability to engage with the My AI chatbot; and (iii) quickly request their teen's location (which the teen must approve before location is shared). Parents can also control whether the teens have access to sensitive content

²¹⁹ <https://values.snap.com/news/new-safeguards-for-snapchatters-2023>.

²²⁰ <https://values.snap.com/news/new-features-to-help-protect-our-community>.



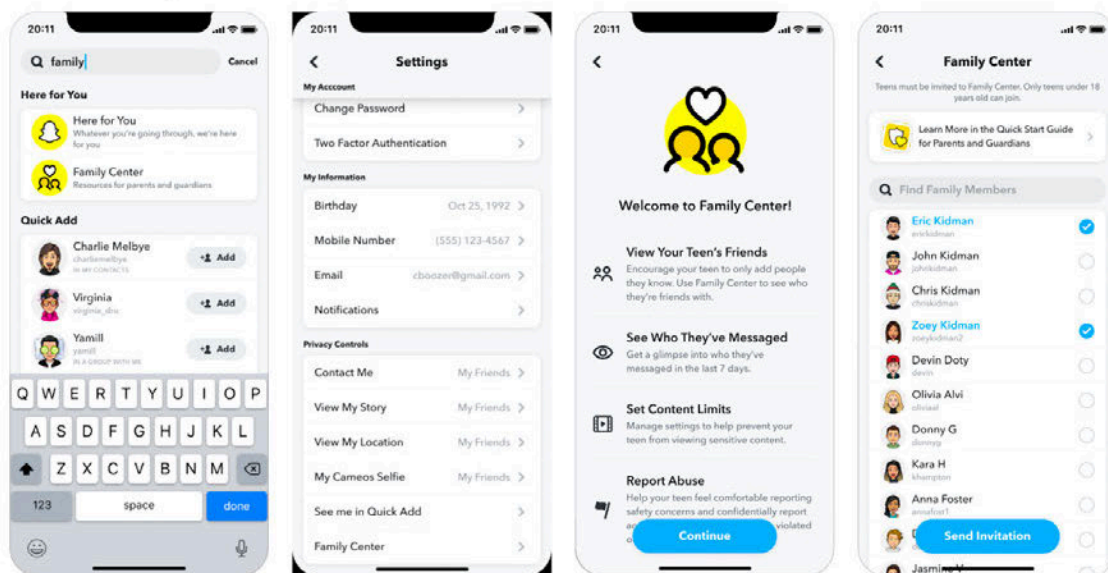
on their Discover and Spotlight feeds. Parents are also able to easily report accounts that may be in violation of our Community Guidelines and have access to helpful resources directly in the app.

Our goal in designing Family Center was to empower both caregivers and teens, balancing parents' desire for more insight with teens' desire for autonomy and privacy - notably ensuring that teens' messages remain private. We continue to put care and time into establishing this balance in a thoughtful way, engaging in user research and surveys, competitive research, focus groups and interviews with both teens and parents, feedback sessions with dozens of online safety experts and academics, including members of our current Safety Advisory Board, and extensive cross-functional internal reviews, including by our Product Legal and Privacy Engineering teams.

In their annual report,²²¹ Jugendschutz.net, the joint competence center of the German Federal and State governments for the protection of children and young people on the Internet, highlighted Family Center as a positive example in the area of parental tools and support on social media platforms. The report concluded that Family Center can help teens and parents talk about negative experiences, contacts, or time spent on the platform. At the same time, it noted the opportunities for teen control, such as having to agree to parental oversight.



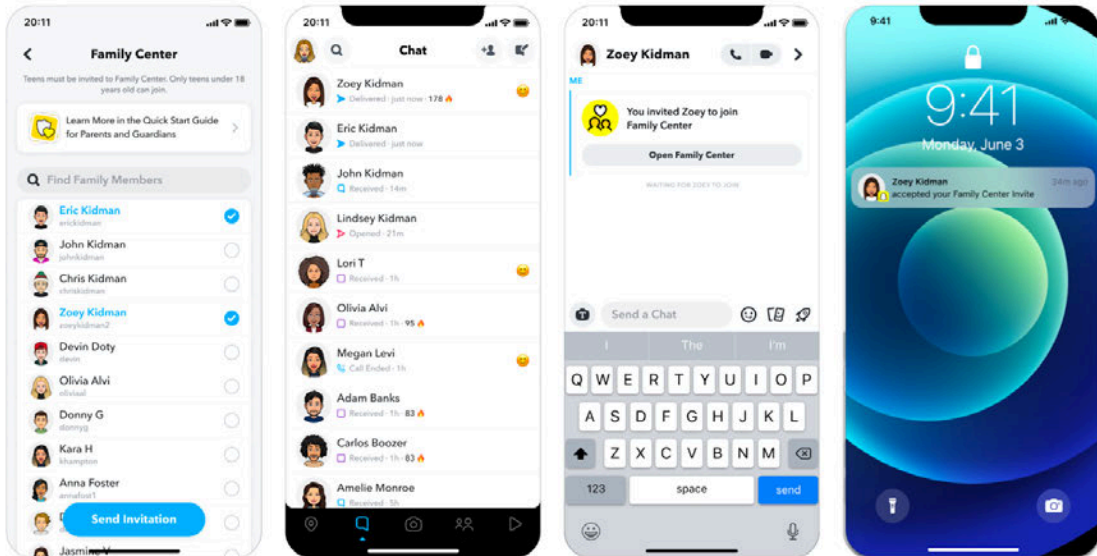
Finding Family Center



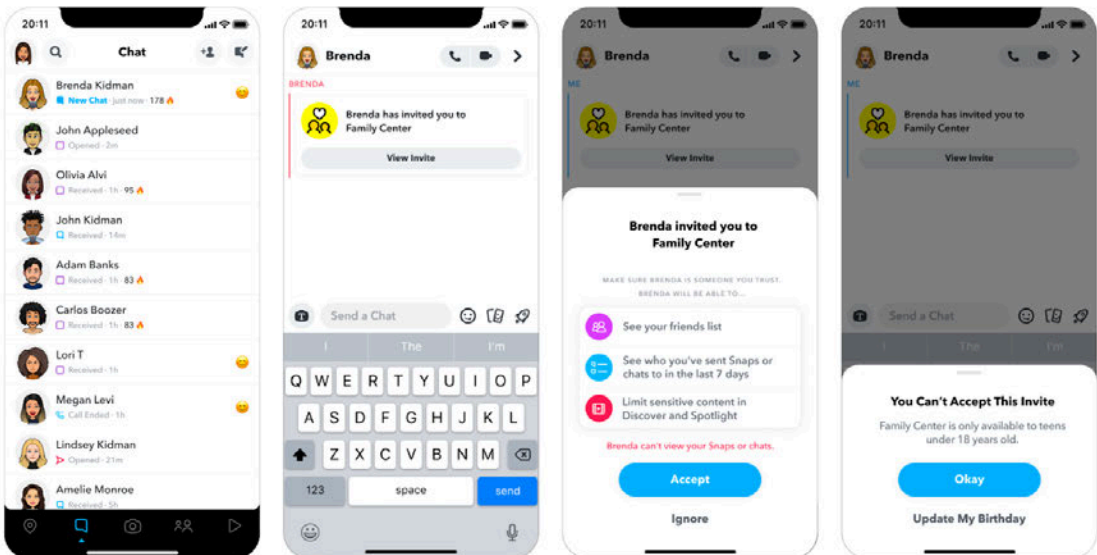
²²¹ Jugendschutz, 'Jugendschutz im Internet - 2022 Bericht', April 2023, [url](#).



Inviting Teens to Family Center

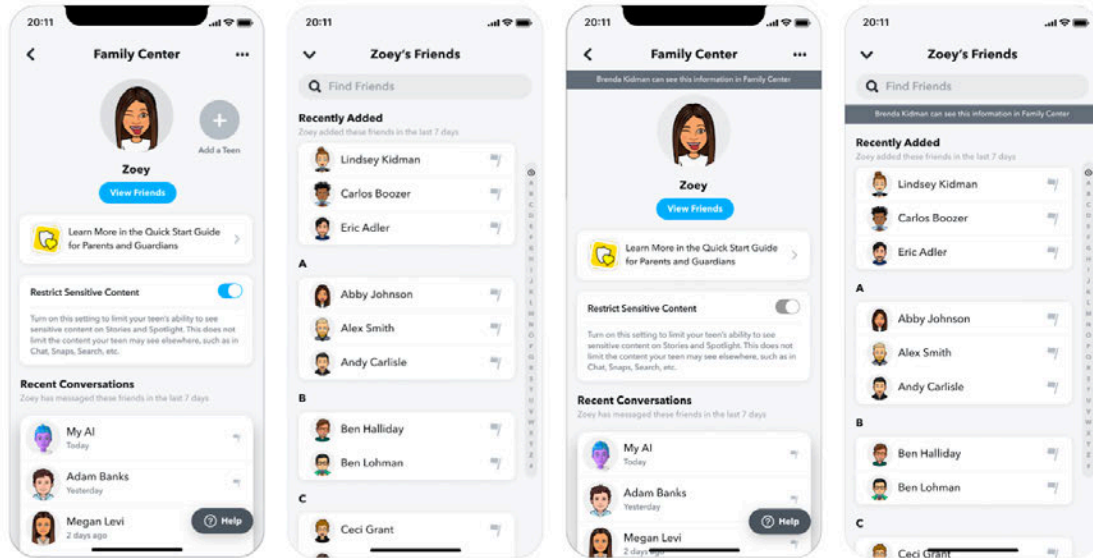


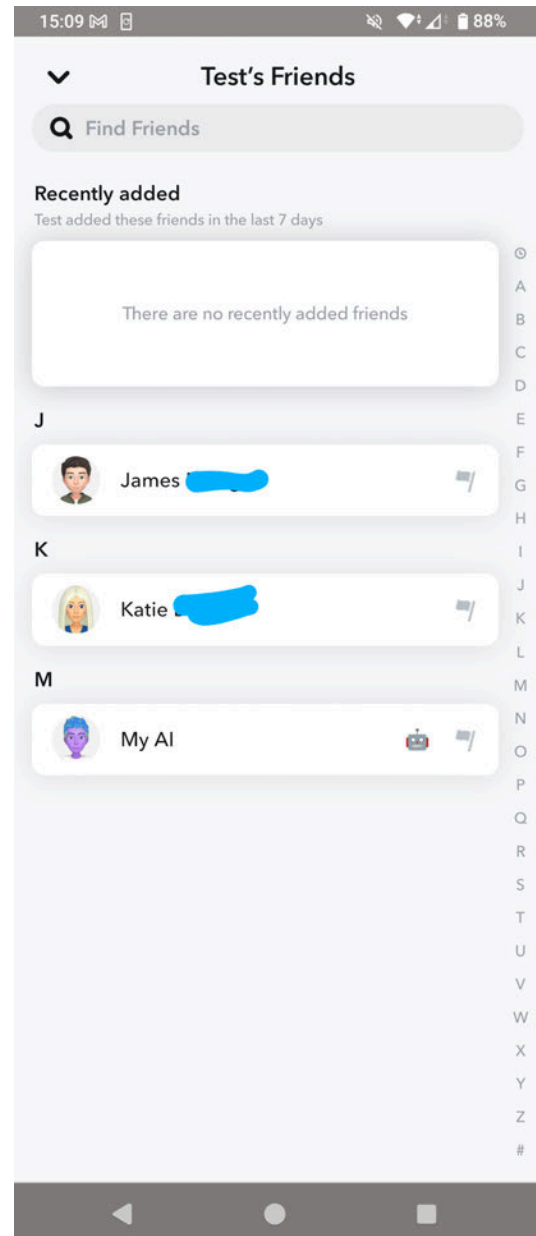
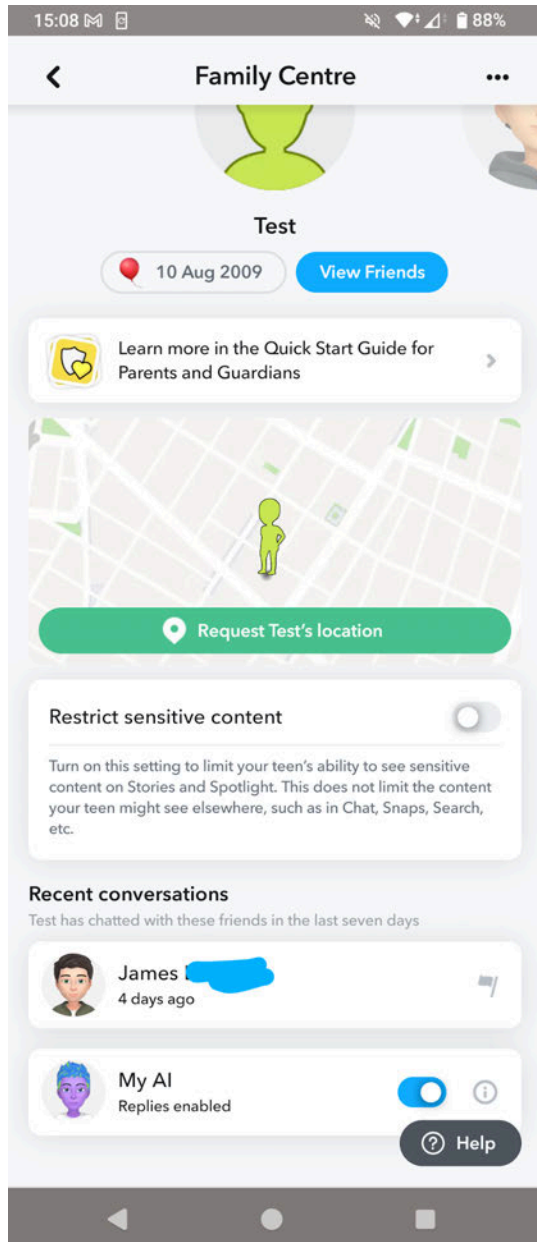
Joining Family Center





Using Family Center





5.8.5 Conclusion

We take the protection of minors seriously. We've designed Snapchat to protect their privacy, safety, and security. Our key tenets include acting in the best interests of Teens, offering strict default settings for all users, and respecting Teens' freedom to express themselves safely, while recognizing their right to information about the world. We aim to achieve these tenets by positioning parents and guardians to help guide teens in their responsible use of our platform, attaching a heightened safety interest to Teens using our products, and establishing processes to ensure we develop products in a way that upholds these tenets. We have implemented these tenets through the use of [Family Center](#), focusing on age-appropriate content, reporting and



blocking mechanisms, and putting in place appropriate protections and limitations on private messaging, friending, public content, and advertising.

As explained in Section 4, we have concluded that the targeted measures we've taken to protect the rights of the child, including age verification and parental control tools, in combination with the other mitigations explained in this Section 5, are reasonable, proportionate, and effective mitigation measures for the risks identified for Snapchat's in-scope services. We have actively worked with the Commission and others to establish guidance on a high level of privacy, safety and security for Art 28 of the DSA, and are working through that guidance to consider whether further mitigation measures may be reasonable, proportionate, and effective for online platforms, 'gateways,' and other online services.

5.9 Content Authenticity

5.9.1 Introduction

Snap is aware that there is intense interest and concern surrounding the ways in which advancements in generative AI technologies are impacting online platforms. Content authenticity is a very challenging topic without a silver bullet.

Snap is very conscious of the issues and has implemented a number of measures. Snap recognises the potential for AI generated or transformed content to be distributed through the inscope services of Snapchat, such as Spotlight and Discover, as explicitly called out in our [Community Guidelines](#). Just like any other content distributed through these channels, this content may constitute illegal content or information that otherwise violates Snap terms and could contribute to the systematic risks outlined in Section 34 of the DSA. Snap continues to carefully monitor developments and industry practice, including regarding whether and how best to use prominent markings and other measures to distinguish content that falsely appears to be authentic or truthful.

5.9.2 Risk Assessment Results

Snap gave due consideration to the risks and harms that could arise from dissemination of user content in its risk assessment results section of this Report. In particular:

- Section 4.1.10 (Harmful False Information) - In this section, Snap recognised that “fake news,” (online) “disinformation” and “deep fakes” had gained a lot of attention in the media and academic and political debate over the last years. We recognised that such content presented a risk of significant harm. This applied to all content formats, whether or not generated using AI tools. However, when considering evidence relating to Snapchat specifically, we found very low prevalence rates of this type of harmful content.



We concluded that ‘Harmful False Information’ fell within the lowest likelihood and risk prioritisation category relative to other harms being monitored on Snapchat.

- Section 4.3.1 (Negative Effects on Democratic and Electoral Processes) - In this section, Snap recognised that online platforms may have a negative effect on the electoral processes and the exercise of political rights by amplifying digital disinformation or deceptive content relating to political matters or processes. Again, this applied whether or not generated using AI tools. However, when considering evidence relating to Snapchat specifically, we found only limited occurrence of content harmful to democracy. Independent reports of electoral interference on Snapchat are vanishingly rare. In connection with a major, high-profile election in 2022, we onboarded Snap to the Election Integrity Partnership (EIP),²²² a partnership among leading research centers and civil society organizations who monitor online harms to democratic processes; as participants in the EIP threat escalation program, our teams received only one single incident report from the researchers monitoring risks on Snapchat. We concluded that ‘Negative Effects on Democratic and Electoral Processes’ fell within our lowest likelihood and risk prioritisation category relative to other harms being monitored on Snapchat.
- Section 4.3.2 (Negative Effect on Civic Discourse) - In this section, Snap recognised that digital content platforms could contribute to Negative Effects on Civic Discourse. For example, we noted:
 - The potential for personalized content and algorithmic biases lock users into echo chambers, reinforcing existing beliefs and potentially leading to polarized communities, which hinders open dialogue.
 - The risk of amplified dis- and misinformation negatively impacting public opinion on important civic issues.
 - The possibility of amplification of extreme or sensational content to retain user attention leading to heightened polarization and a hostile online environment.

However, when considering the evidence relating to Snap specifically, we again found a very low prevalence of content related to harming Civic Discourse relative to other categories being monitored. We concluded that ‘Negative Effect on Civic Discourse’ fell within the lowest likelihood and risk prioritisation category.

5.9.3 Mitigations

Although there is not a high prevalence of Harmful False Information, Fraud and Spam or impersonation, we take harmful information of this nature on Snapchat very seriously and Snap has implemented a number of mitigation measures.

Guidelines, policies, and practices

Snap maintains robust policies – applicable to both the dissemination and the creation of generative AI content – that function to mitigate risk and advance safety.

²²² Election Integrity Partnership (2020), [url](#).



Creation

Snap has developed several internal policies relating to generative AI. In particular,

- (1) Content and Product policies: We have developed a suite of policies that disallow the generation of harmful content (including deceptive political content). Our policy and moderation teams work in partnership with engineering and data science colleagues to ensure that our AI products are responsibly trained on these policy parameters.
- (2) Acceptable Use: We have similarly developed Acceptable Use Policies that prohibit the use of our AI tools to attempt to generate violative content at the prompt-level.

With regards to the Content and Product policies, we have drafted and implemented internal Generative AI Policies to govern the internal development of generative AI features, such as MyAI. Our Product and engineering teams refer to this policy as they train models or adapt models from third parties. Our Safety team applies this policy when testing new features.



- [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

These aligned very closely with the rules for content dissemination, which are explained below. In addition, our generative AI tools feature a broad range of mitigation measures, depending on the tool, and include for example: specific transparency statements, abusive language detection and query related measures, age appropriate and/or canned responses, reporting mechanisms, off-by default settings, data minimisation, data sharing, testing and parental controls (see Section 5.8 on the Protection of Minors). Although out of scope of the DSA and this Report, risks and mitigations relating to our generative AI tools are assessed via our privacy and safety by design product reviews (see Section 6.3).

Dissemination

In the context of dissemination of content on Snapchat's online platform, in scope of the DSA, we understand well that online platforms may have a negative effect on the electoral processes and the exercise of political rights by amplifying digital disinformation or deceptive content relating to political matters or processes.

Our [Community Guidelines](#) and [Terms of Service](#) set out the rules on what content is allowed on Snapchat. They are focused on preventing harm to Snapchatters and the broader community from content and behaviour, whether or not caused by generative AI or any other form of IT tools (such as Photoshop). These rules apply to all content formats across our platform, including content that is AI-generated. While the rules are agnostic to content format or creative tools, the Community Guidelines specifically note: "We implement safeguards designed to help keep generative AI content in line with our Community Guidelines, and we expect Snapchatters to use AI responsibly. We reserve the right to take appropriate enforcement action against accounts that use AI to violate our Community Guidelines, up to and including the possible termination of an account."

Our rules and internal enforcement guidance include clear provisions related to content risks, for example for Civic Discourse and electoral processes. In particular, our Community Guidelines prohibit spreading false information that causes harm or is malicious, such as denying the existence of tragic events, unsubstantiated medical claims, undermining the integrity of civic processes, or manipulating content for false or misleading purposes (whether through generative AI or through deceptive editing).

Our Community Guidelines rules on false information refer to a more detailed [Explainer](#) that prohibits content that undermines the integrity of civic processes, or deep fake content or other media that is manipulated for false or misleading purposes. The Community Guidelines further explain that these prohibitions extend to the following types of harmful content:



- Procedural interference: misinformation related to actual election or civic procedures, such as misrepresenting important dates and times or eligibility requirements for participation.
- Participation interference: content that includes intimidation to personal safety or spreads rumours to deter participation in the electoral or civic process.
- Fraudulent or unlawful participation: content that encourages people to misrepresent themselves to participate in the civic process or to illegally cast or destroy ballots.
- Delegitimization of civic processes: content aiming to delegitimize democratic institutions on the basis of false or misleading claims about election results, for example.

Sharing such content will violate Snap's Community Guidelines irrespective of whether it is AI-generated or user-generated, or whether it is generated on Snapchat or on another platform.

Snap has a suite of internal policies and guidelines to help our content review and trust and safety teams apply the Community Guidelines to user generated content disseminated via our online platforms (such as Spotlight and Discover). They provide more granular information for our content review teams. For example, we explain that obvious jokes, memes, satire and non-libelous comments about prominent social figures are OK; whereas false political narratives meant to undermine elections, or harmful / defamatory deepfakes, are NOT OK.

In addition, our platform does not widely distribute an unvetted feed of algorithmically curated political information; we disallow all political content²²³ from Spotlight (our broadcast platform for User Generated Content) unless it's from trusted news partners and creators, and pre-moderate that surface to ensure that other such political content is not distributed. This safeguard ensures that Snap is not algorithmically promoting political statements from unvetted sources, and generally reflects Spotlight's function as an entertainment platform. (Consistent with our commitments to fundamental rights of expression and access to information, Snapchat provides other, non-algorithmically amplified spaces for users to express their views and political observations, such as Chat and My Story; users can also seek access to political information from known publishers and creators whom Snap has on-boarded for distribution on the Stories tab).

With regards to advertising, as explained in our previous communications and meetings with the Commission in response to its RFI on Gen AI and its consultation on the Guidelines for electoral processes, we do not require ads to label when advertisement includes generative AI content nor do require advertisers to disclose to us the tools they used to edit or create their ad creative. Instead, our approach is to subject all of our ads to a review process, and political ads are also subject to fact checking. Deceptive ads are rejected, irrespective of whether they use AI, photoshop, or other digital editing tools. Ads that are not deceptive, and otherwise comply with

²²³ For these purposes, "political content" means content related to political campaigns and elections, government activities, and/or viewpoints on issues of ongoing debate or controversy. This includes content about candidates or parties for public office, ballot measures or referendums, and political action committees, as well as personal perspectives on candidate positions, government agencies/departments or the government as a whole.



our Ad Policies, are approved to run (and if they are a political ad, they must include a “paid for by” disclaimer and are catalogued in Snap’s political ads library).

User Guidance on Generative AI features

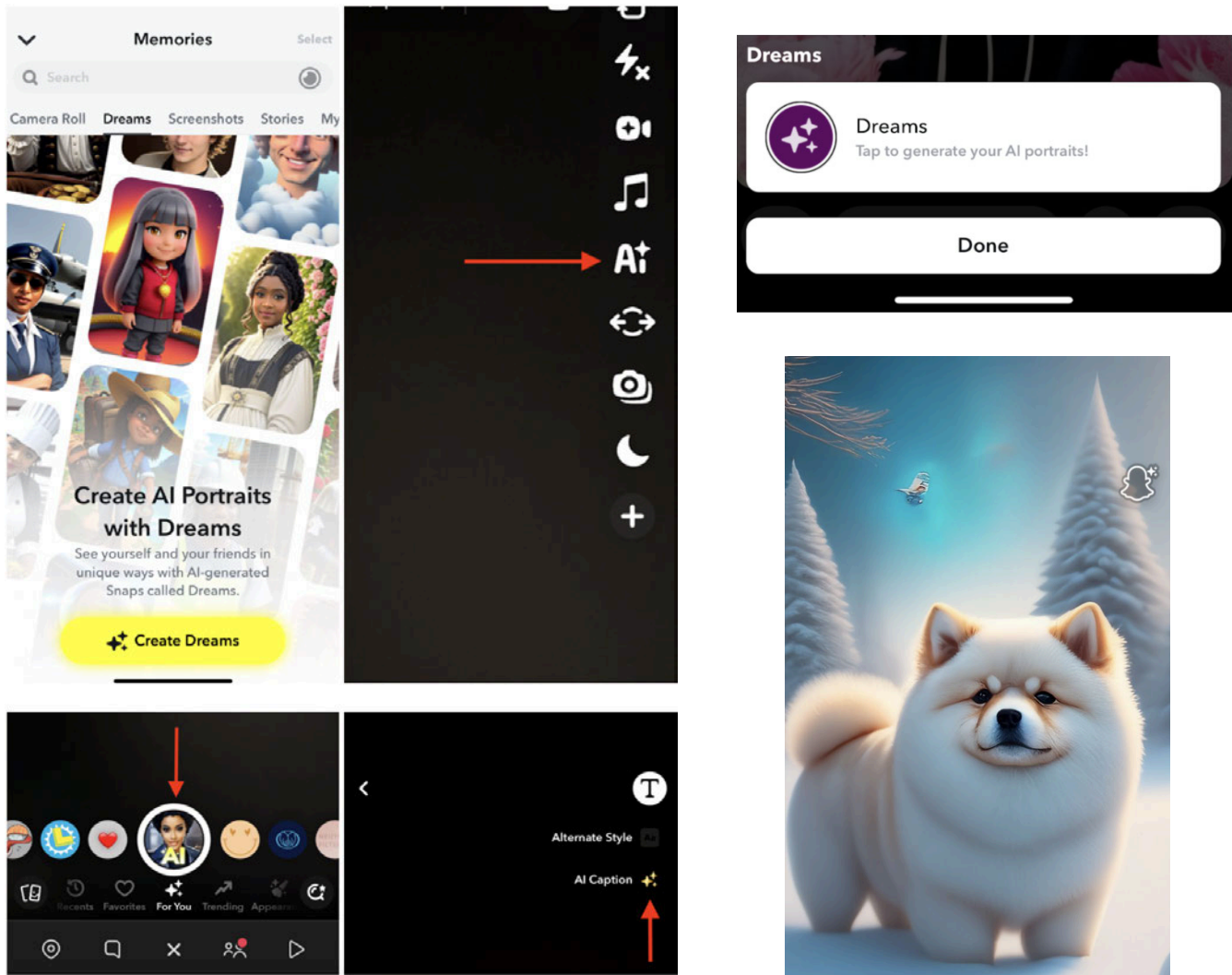
On-platform features for creating generative AI content are not part of Snap’s inscope services and are out of scope of this Report (save for certain commonplace ad creation tools). Nevertheless, outside of its DSA obligations, Snap has released a [generative AI support site](#) that explains what generative AI is and provides additional transparency around Snap’s practices with regard to generated images which are detailed below. To that end, we provide indicators across most Snapchat experiences that involve AI.

We use a range of methods to notify users that they are interacting with or viewing AI-generated content. These include:

- The sparkle icon () displayed alongside features or outputs powered by AI.
- Context Cards or disclaimers that explain when a feature is AI-driven.
- On-screen tool tips or labels that appear within relevant features like Dreams and AI Snaps
- In some cases, visible watermarks, such as a Snap Ghost with sparkles, are added to exported or saved images created using generative AI tools.



Here are examples of what these AI indicators look like in Snapchat:



These signals help users understand that the content was not created by a human and may not depict real-world events or scenarios.

We are mindful of the importance of clarity and consistency. For this reason, Snap continues to expand and refine our transparency efforts. For example:

- When users export AI-generated images to their Camera Roll, these images will often include a watermark indicating they were created using Snap's AI-powered tools.
- When AI-generated content is shared with others on Snapchat, we may include Context Cards to ensure recipients understand that the content was created with AI.
- We are exploring new ways to further enhance visibility and user understanding of AI involvement across the product.



At the same time, we recognize that watermarks and labels alone are not foolproof - they can be missed, altered, or removed. That's why we rely on a multi-layered approach that includes product disclosures, interface cues, content moderation, and user reporting tools to help mitigate risks of confusion or misuse.

Snap remains focused on helping users engage safely and responsibly with generative AI. We believe that transparency, when done effectively and consistently, builds trust and helps ensure that AI-powered features are understood and used appropriately.

Enforcement

Snap enforces these Community Guidelines fairly and consistently, using internal policies and guidelines, and applies outcomes that are commensurate with the severity of risk.

Accounts that we determine are used to perpetrate serious, high-severity harms will immediately be disabled. For other violations of our Community Guidelines, Snap generally applies a three-part enforcement process:

- Step one: the violating content is removed.
- Step two: the Snapchatter receives a notification, indicating that they have violated our Community Guidelines, that their content has been removed, and that repeated violations will result in additional enforcement actions, including their account being disabled.
- Step three: our team records a strike against the Snapchatter's account.

A strike creates a record of violations by a particular Snapchatter. Every strike is accompanied by a notice to the Snapchatter; if a Snapchatter accrues too many strikes over a defined period of time, their account will be disabled.

This strike system ensures that Snap applies its policies consistently, and in a way that provides warning and education to users who violate our Community Guidelines. The primary goal of our policies is to ensure that everyone can enjoy using Snapchat in ways that reflect our values and mission; we have developed this enforcement framework to help support that goal at scale.

Partnerships

Snap continues to be actively engaged in the roll out of the EU AI Act, including with regards to the drawing of the related codes of practice for providers of general-purpose AI models and those regarding the detection and labelling of artificially generated or manipulated content and under the AI Pact.

More broadly, tackling risks stemming from generative AI requires (among others) broad industry-wide technical solutions which have not been clearly identified so far. This is why Snap is actively engaging with its peers and industry experts in different fora to share best practices and advance the technical debate. These partnerships, industry collaborations and efforts include:



- OpenAI Integration: My AI is powered by OpenAI's ChatGPT, and Snap closely partners with OpenAI in relation to providing the My AI service, including sharing feedback on moderation of content.
- Tech Coalition / Working Groups on Generative AI: Snap is a member of the Tech Coalition's Working Group on Generative AI Content, and a member of the GenAI Briefing Subgroup. The Working Group on Generative AI Content meets regularly to facilitate dialogue and information- and idea-sharing around mitigating content-level generative AI risks. The GenAI Briefing Subgroup meets periodically to plan expert briefings for Tech Coalition members on topics related to Generative AI risks; such briefings have included representatives from government, law enforcement, civil society, and the research community.
- Tech Accord to Combat Deceptive Use of AI in 2024 Elections: Snap was an initial signatory to the Tech Accord to Combat Deceptive Use of AI in 2024 Elections. This compact seeks to set expectations for how signatories will manage the risks arising from deceptive AI election content created through their publicly accessible, large-scale platforms or open foundational models, or distributed on their large-scale social or publishing platforms in line with their own policies and practices as relevant to the commitments in the accord. The Accord was announced at the Munich Security Conference in February 2024.
- ITI AI Futures Initiative: Through its membership in the Information Technology Industry Council (ITI), Snap has participated alongside other private sector actors in the AI Futures Initiative. Led by technical and policy experts spanning the tech ecosystem, the Initiative is a forum through which participants are developing action-oriented recommendations for AI policy and working to address emerging questions around AI. Deliverables to date have included the issuance of Global AI Policy Recommendations to help guide governments around the world as to develop responsible regulatory approaches to AI-related issues.
- HackerOne - Red-Teaming Collaboration: Snap partnered with HackerOne on red teaming exercises to test the strict safeguards Snap has in place around AI. Together with HackerOne, we made significant developments in the methodology for AI safety red teaming that has led to a more effective approach to surfacing previously unknown problems. We refer to the HackerOne blog for more details: <https://www.hackerone.com/ai/safety-vs-security>
- As an active member of the EU Internet Forum, Snap will support the upcoming dedicated working group on generative AI matters.
- We are also members of the Centre for Information Policy Leadership (CIPL) and the Future of Privacy Forum (FPF) which work with industry stakeholders (like Snap), NGOs and government agencies in each region to advance a broad array of information topics. CIPL has been a leader in AI matters for many years through its dedicated AI Project and specific Brazilian AI Project. Most recently, in Europe, CIPL has responded to the UK Information Commissioner's Office (ICO)'s consultations on Generative AI, and led various



forums on Accountable Governance of AI and AI Regulation in Brussels and the UK. Similarly, FPF is working on AI Governance and other responsible Gen AI initiatives.

5.9.4 Conclusion

Content authenticity is a very challenging topic without a silver bullet. Snap is very conscious of the issues and has implemented a number of measures (including relevant measures identified in the Commission's guidelines concerning elections). Snap continues to carefully monitor developments and industry practice, including regarding whether and how best to use prominent markings and other measures to distinguish content that falsely appears to be authentic or truthful.

We have concluded that the mitigations explained in this Section 5, reasonable, proportionate and effective for the risks identified for Snapchat's in-scope services.

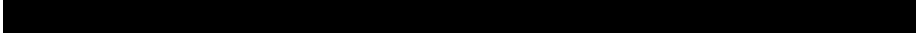
5.10 Trusted Flaggers

5.10.1 Trusted Flagger Program

Snap's Trusted Flagger Program was developed to help non-profits, non-governmental organizations (NGOs), select government agencies, and safety partners support the Snapchat community by leveraging a special channel to report content that violates Snapchat's [Community Guidelines](#). The program has been in operation for a number of years and has been developed in response to evolving risks and mitigations with respect to Snapchat.

Trusted Flaggers send a completed report form with details of the potential violation via email to a dedicated, confidential email address or via a webform. The email and webform reports feed into a high priority channel and reports are reviewed in less than 48 hours (with reports relating to the most serious harms prioritized and reviewed well within this timeframe). Once a decision has been reached, Snap informs the Trusted Flagger about the outcome of their report, including whether appropriate action has been taken. This channel can be used to report any type of Community Guideline violation or otherwise illegal content and is designed to supplement in-app reporting, which is still very much encouraged.

Our Trusted Flagger Program allows us to gain insight from the Trusted Flaggers over the types of harm they are encountering, and the behavior of victims in these circumstances. 


 In addition to providing a specific reporting channel, the Trusted Flagger Program also allows us to build strong relationships with Trusted Flaggers. Snap makes use of our strong relationship with Trusted Flaggers to give product safety updates, encourage the promotion of our safety tooling and provision of safety resources (like links to our [Safety Center](#)).



5.10.2 Onboarding a new trusted flagger

Snap receives reports from both Trusted Flaggers under the EU Digital Services Act and Trusted Flaggers who have entered into a direct agreement with Snap. Snap onboards the Trusted Flagger when the European Commission DSA Trusted Flagger web page is updated with the name of the appointed organisation and the contact email supplied. For the latter, typically, when we are considering accepting a new Trusted Flagger into our Snap program, we take into account geographic coverage, area of expertise, anticipated volume of reports, among other factors. Once a Trusted Flagger is appointed or accepted we send them an onboarding package, which includes an overview document of the program, including our commitment to review reports in less than 48 hours (with reports relating to the most serious harms prioritized and reviewed well within this timeframe); instructions on how to file reports to our dedicated and confidential email address; and contact information in case they have questions or concerns about the program.

When a Trusted Flagger wants to file a report, they leverage our instructions on which categories of information they should include in their reporting email. After we receive an email, Snap's Trust and Safety teams review the report and take any appropriate enforcement action, or request additional information if required for full investigation. Once a decision has been made, Snap will inform the Trusted Flagger of any enforcement action the Trust & safety team has taken in relation to the reported content or accounts.

Our teams remain in contact with the Trusted Flaggers, including when we need to discuss any issues with their reports. Our team evaluates the reports submitted based on the completion of the form, the accuracy of the information provided, and whether or not the report leads to enforcement or other action.

5.10.3 DSA Trusted Flaggers

We monitor the Commission's publication pursuant to Article 22(5) of the entities that have been awarded the status of 'trusted flagger' pursuant to Article 22(2) DSA. Upon receiving notice about a new DSA Trusted Flagger, we decide whether the trusted flagger is relevant to Snap and if so establish communication. Assuming all is well, we then follow the process outlined above. If our team identifies trends that are impacting the quality of the reports, we will communicate this information with the DSC to identify a resolution.

5.10.4 Trusted Flagger Program Trends

Europe including the UK: We continue to receive reports from and engage in cooperation with Trusted Flaggers. For the first half of 2024 and the second half of 2024 we saw 1,165 and 741 reports respectively from Trusted Flaggers in the UK, Norway and several Member States of the European Union. We have observed a reduction in reports from 2023, which we believe to be



due to our extensive efforts to combat financial sextortion and additional chat reporting options - and these have resulted in even fewer user escalations to trusted flaggers.

[REDACTED]

[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]



[illegible]



498



499

500



501



[REDACTED]

[REDACTED]

[REDACTED]

Snap has an existing, carefully managed Trusted Flagger Program with valued member organizations from a wide array of countries including many in the European Union. Snap looks forward to evolving its Program to incorporate organizations that have been awarded trusted flagger status under the DSA.

5.11 Dispute Settlement Bodies

We invest significant resources in our community support teams who work to resolve queries and complaints received from Snapchatters and others. In line with DSA requirements (Article 21), we inform users of their right to seek out-of-court dispute resolution if they are dissatisfied with the outcome of our internal complaint-handling process.



[REDACTED]

As of 14 August, we are aware of six bodies that have been established and certified by the DSCs. Recipients of the Snapchat service have the option to contact such out-of-court dispute settlement bodies to raise their case. The relevant out-of-court dispute settlement bodies are then able to reach out to Snap via our dedicated contact point (dsa-enquiries@snapchat.com) to start the out-of-court dispute settlement process. Snap will then engage, in good faith, with the selected certified out-of-court dispute settlement body with a view to resolving the dispute following Snap's policies and procedures.

5.11.2 Enquiries

From January 1st to June 30th 2025, Snap received a number of disputes submitted to certified Article 21 bodies. [REDACTED]

[REDACTED]

5.11.3 Further considerations

[REDACTED]

[REDACTED]

[REDACTED]



[REDACTED]

[REDACTED]

5.11.4 Conclusion

Snap remains committed to resolving user disputes effectively and in line with DSA requirements. Furthermore, we continue to support establishing an EU-wide settlement body or an EU portal for better user interactions. This approach would ensure consistent application of rules across all EU member states and provide a simplified, single point of access for operators.

As explained in Section 4, we have concluded that Snap’s current approach to Dispute Resolution, in combination with the other mitigations explained in this Section 5, is reasonable, proportionate and effective for the risks identified for Snapchat’s in-scope services.

5.12 Codes and Crisis Protocols

5.12.1 Cooperation

Snap highly values cooperation with other providers and industry experts as a way to share best practices and learning experiences that can enhance risk mitigation strategies. We are highly committed to industry partnership to steer progress in the fight against illegal and harmful content online. In particular, Snap is active member of the following groups:

- **EU Internet Forum** - Snap is an active member and contributor of the [EU Internet Forum](#) (EUIF), which provides a collaborative environment for EU governments, the internet industry, and other experts and partners to discuss and address the challenges posed by the presence of malicious and illegal content online. The EUIF aims at exploring possible responses against abuse and exploitation of online platforms by terrorists and violent extremists, as well as other malicious actors, including those that groom children for the purpose of sexual abuse and the production and dissemination of child sexual abuse material online. Earlier this year, the scope of the EUIF work was expanded to tackle also the issues of drug sales online and the trafficking of human beings.



- **Technology Coalition** - Snap is also a member of the [Technology Coalition](#), which is an alliance of global tech companies who are working together to combat child sexual exploitation and abuse online. The Tech Coalition coordinates industry's overall effort to combat child sexual abuse online. It provides resources, education, and capacity-building to tech companies, and serves as a resource for external stakeholders - from global policy-makers to members of the media - on what industry is doing to tackle this issue.
- **WeProtect Global Alliance** - Snap is a Board member of the [WeProtect Global Alliance](#), which brings together the private sector, government and civil society to drive positive change to help protect children from sexual abuse online.
- **Alliance to better protect minors online** - Until the recent discontinuation of the initiative in July, Snap was also a member of the [Alliance to better protect minors online](#). This self-regulatory initiative was aimed at improving the online environment for children and young people by steering debates and exchanges on the topic.
- **CIPL** - Snap is a member of The Centre for Information Policy Leadership (CIPL). This is a global privacy and data policy think and do tank based in Washington, DC, Brussels and London. We work with CIPL and other industry leaders, regulatory authorities and policy makers to develop global solutions and best practices for privacy and responsible use of data, including with respect to teenagers and young adults.
- **The Future of Privacy Forum** - Snap is a member of [The Future of Privacy Forum](#) (FPF). FPF is a non-profit organization that serves as a catalyst for privacy leadership and scholarship, advancing principled data practices in support of emerging technologies. We work with FPF on a range of matters, including developing best practices related to Augmented Reality (AR), Artificial Intelligence (AI), biometric data, children's rights, and more.
- **Centre on Regulation in Europe (CERRE)** - Snap is a member of [CERRE](#), which is a not-for-profit think tank based in Brussels. Its goal is to support and inform about regulation in Europe and beyond. We work with CERRE on in-depth reports and issue papers that address the major regulation challenges and high-quality, policy-oriented research undertaken by top-level academics in the tech, media and telecom sector.
- **Internet Watch Foundation (IWF)** - We are a member of the [Internet Watch Foundation](#), which is an independent, non-profit organisation that aims to prevent child sexual abuse online. We sit on their Funding Council and our Head of Public Policy, UK & Ireland, is an industry trustee on the IWF Board.
- **Thrive** - Snap is a member of this industry signal and best practice sharing initiative which is focused on suicide and self-harm content.

Other Cooperation with industry

Snap is actively involved in the work of a number of EU-based trade associations to contribute to the policy debate to support the development of a proportionate regulatory framework to promote online safety.

- **DOT Europe** - Coordination on EU privacy, security, safety, content policy issues
- **ITI** - Coordination on EU privacy, security and safety policy issues



5.12.2 Codes of Practice

The DSA establishes that the Commission and the European Board for Digital Services (“the Board”) shall encourage and facilitate the drawing up of voluntary codes of conduct at Union level to contribute to the proper application of the DSA (Article 45). [REDACTED]

Snap welcomes the opportunity to support industry-wide efforts to promote risk-mitigation practices in the form of voluntary codes. [REDACTED]

As a company with limited resources, Snap is constantly required to prioritize and ensure its resources and efforts are focused on where the biggest risks and challenges for the company are. As we advance in our learning curve from our DSA risk assessment, we will continue to prioritize interventions where we see the highest risks.



EU hate speech Code

As part of its long-standing commitment to fight harmful and illegal content, Snap signed onto the [EU Code of Conduct to counter illegal hate speech online](#) in 2018.

[REDACTED]

[REDACTED]

Since joining the code, Snap has successfully passed all the evaluations and in the course of the last [monitoring exercise](#) (2022), and for the 5th consecutive year, **Snap did not receive any notification** [REDACTED]

Additionally, in the course of 2022 Snap has worked closely with the European Commission and other signatories to further strengthen some of the Code commitments by reinforcing and better framing the existing cooperation between IT companies and CSOs, beyond the remit of the monitoring exercises. This work led to the publication of an [Annex to the existing code](#) in December 2022. Snap continued to engage with the European Commission team (DG Just) and regularly cooperated with other industry signatories on a further update to the EU Hate Speech code to bring it in line with the DSA.

Since the EU Hate Speech Code was updated in line with the DSA on 20 January 2025, Snap has complied with the Code as part of its DSA compliance program.



FSM Code of Conduct

In September 2017, Snap joined 'the Freiwillige Selbstkontrolle Multimedia- Diensteanbieter e.V. (FSM), an officially recognized voluntary self-regulation association for the protection of minors in online media. [REDACTED]

[REDACTED]

EU disinformation code

Snap has not yet signed up to be a member of the EU disinformation code. Being very resource constrained and considering our limited exposure to this type of risk, we have so far opted not to join. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



Article 28 DSA Guidance

- [illegible]



[illegible]

[REDACTED]



[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

Cooperation with external stakeholders is a very important element of risk mitigation for Snap. Knowledge sharing and best practice development with experts and peers are key to strengthen and increase the effectiveness of our internal risk mitigation measures. This is why the company has signed up to several voluntary codes and is actively engaged in many different international fora and associations to steer constructive debate and best practice development in areas like CSEAI, protection of Teens, and Hate Speech (and we work closely with our global trusted flagger network on these matters). We will continue to monitor our risks and prioritize interventions on the most severe risk areas. When it comes to dealing with unexpected events resulting in



heightened levels of risks for the platform, our Content Crisis Response Protocol plays an important role in providing a structure to our collaborative internal operations and efforts.

As explained in Section 4, we have concluded that Snap’s current approach to codes of practice and crisis protocols, in combination with the other mitigations explained in this Section 5, is reasonable, proportionate and effective for the risks identified for Snapchat’s in-scope services.



6. Ongoing Risk Detection and Management

Snap has developed a number of practices to detect and manage risks to Snapchat’s in-scope services. This includes: (1) the establishment of a Platform Risk Framework based on Snap’s product values, established international human rights principles, and risk-based metrics such as prevalence and severity analyses; (2) the designation of a senior, cross-functional team responsible for applying the framework and assessing its outcomes, including a DSA Governance Group and meeting; (3) development of a repository of internal resources to support the detection and management of risk—these include harm severity assessments; prevalence metrics; reporting data reviews and a library of Terms and policy resources; and (4) continual improvement and assessment through our Digital Well-Being Index (DWBI) Initiatives and Safety Advisory Groups (Safety Advisory Board and Council for Digital Well-Being).

This Section of the Report provides further details of these practices pursuant to Article 42.4.(b) (reporting on the mitigation measures relating to Article 35.1.(f)) and Article 42.4.(e) of the Digital Services Act.

6.1 Platform Principles-based Framework

[REDACTED]

[REDACTED]

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]



5

11

5

6.2 DSA Compliance Team and Cross-Functional Working Groups

6.2.1 Introduction

This section outlines Snap’s governance mechanisms relevant to compliance with the requirements of the DSA, in particular with regard to the requirements outlined in Articles 11, 12, 13, 41, and 43.



[REDACTED]	
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]

6.2.3 DSA Independent Compliance Function

Snap has established an Independent Compliance Function (ICF) as part of the Snap Legal team, separate from Snap operations, with sufficient authority, stature, and resources that is responsible for coordinating, overseeing, and implementing Snap's Privacy and Regulatory Program. The ICF provides oversight to ensure the necessary internal processes, resources, testing, documentation, or supervision are in place for compliance with the DSA and monitors Snap's compliance with the DSA, ensuring the identification and mitigation of risks associated with Snap operations.



In practice, Snap's compliance function extends well beyond our formally designated compliance officers. [REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED] [REDACTED] [REDACTED] [REDACTED]	[REDACTED]	[REDACTED]
[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]	[REDACTED]	[REDACTED]
[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]	[REDACTED]	[REDACTED]
[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]	[REDACTED]	[REDACTED] [REDACTED] [REDACTED]
[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]	[REDACTED]	[REDACTED] [REDACTED] [REDACTED]
[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]	[REDACTED]	[REDACTED] [REDACTED] [REDACTED]
[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]	[REDACTED]	[REDACTED] [REDACTED] [REDACTED]



[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]		

6.2.5 Compliance Officer Designation

Snap has designated Independent DSA Compliance Officers who report to the Management Body and fulfill the tasks outlined in DSA Article 41 for the head of compliance. Snap's compliance officers possess the professional qualifications, expertise, experience, and capabilities necessary to fulfill the designated responsibilities. The Independent Compliance Officers are also responsible for monitoring Snap's compliance with commitments outlined in the relevant codes of conduct or crisis protocols.

[REDACTED]
[REDACTED] The compliance function significantly overlaps with the Data Compliance and Data Protection Officer function, including the requirements to closely cooperate with the EC, responding to inquiries and monitoring compliance with the DSA, and conducting risk assessments. The head of the compliance function reports to individuals who are members of Snap's exec meetings and frequently [REDACTED]
[REDACTED]
[REDACTED] updates the Board and executives on DSA related matters. This satisfies the Article 41 requirement for the head of the DSA compliance function to report directly to the management of the company.

[REDACTED]
[REDACTED]
[REDACTED]

6.2.6 Compliance Officer Qualifications

Snap's Compliance Officers have numerous years of experience in data protection, privacy, compliance and governance.

- [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]



[REDACTED]

- [REDACTED]

6.2.7 Operation of the Independent Compliance Function

Responsibilities of the Independent Compliance Function

The Independent Compliance Function is responsible for the following activities:

- Cooperating with the relevant Digital Services Coordinator and the Commission for the purpose of DSA compliance;
- Providing oversight over the development of the Systemic Risk Assessment methodology and the conduct of the Systemic Risk Assessment, ensuring it is conducted on the basis of the best available information and scientific insights
- Ensuring that risks referred to in DSA Article 34 are identified and properly reported on and that reasonable, proportionate and effective risk-mitigation measures are taken pursuant to Article 35;
- Organizing and supervising the activities related to the independent audit pursuant to DSA Article 37;
- Providing oversight over the validation of controls leveraged to mitigate risks, evaluation of controls, and review of policies;
- Monitoring the compliance with DSA obligations;
- Reviewing and approving Transparency Reports;
- Informing and advising the management and employees about relevant obligations under the DSA;
- Where applicable, monitoring the compliance with commitments made under the codes of conduct pursuant to Articles 45 and 46 or the crisis protocols pursuant to Article 48 of the DSA.

Oversight and Monitoring of Snap's DSA Compliance

The Independent Compliance Function utilizes multiple ways to exercise oversight and monitoring over Snap's DSA compliance:

- [REDACTED]

| [REDACTED]

| [REDACTED]

Any relevant issues or observations are discussed within the Independent Compliance Function as well as with relevant stakeholders and escalated to the Management Body and the CEO as needed. The Independent Compliance Function investigates root causes of issues or observations, develops mitigation plans and works with stakeholders and control owners to



[REDACTED]

implement such corrective actions. The Independent Compliance Function may also escalate significant issues / observations to the Management Body and the heads of the respective function if the root cause is identified as relating to that function.

DSA Management Body

Snap has designated a Management Body which oversees and supports the independence of the compliance function and manages issues as escalated to the Body.

[REDACTED]

[REDACTED]

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



-
- | Response | Percentage |
|--|------------|
| U.S. should take action to address climate change | 95% |
| U.S. should take 'strong' action to address climate change | 90% |

[REDACTED]
 [REDACTED]
 [REDACTED]
 [REDACTED]



[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]

[REDACTED]	
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED] [REDACTED] [REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED] [REDACTED]
[REDACTED]	[REDACTED] [REDACTED] [REDACTED]
[REDACTED]	[REDACTED]

6.2.9 Points of Contact

Designation, Publication, and Change Management

Upon designation, Snap's cross-functional DSA governance team established a process for designating and updating in a timely manner the publication of points of contact for regulatory authorities, recipients of the service, and the legal representative in a location that is publicly available and easily accessible.



Snap's Cross-functional DSA Governance Team reviews Points of Contact and Legal Representative designation [REDACTED]

If a new Point of Contact or Legal Representation is appointed, the Data Compliance Officer or designee will work with the web team to update the above website accordingly.

Points of Contact are made publicly available in all official languages of the EU and are easily accessible on: <https://values.snap.com/privacy/transparency/european-union>

Point of Contact for the Authorities

Upon designation, Snap's cross-functional DSA governance team designated a single point of contact email address for communication with Member State authorities, the Commission, and the Board to enable direct and effective communication between Snap and the relevant authorities on matters related to the Digital Services Act.

Upon designation Snap's cross-functional DSA governance team ensured the easily accessible publication of the relevant information relating to the point of contact for regulatory authorities, including the languages to be used in such communications, on the Snap website.

Authorities can contact Snap at dsa-enquiries@snapchat.com, through our [Support Site](#), which supports all official languages of the EU, and at

Snap B.V.
Keizersgracht 165, 1016 DP
Amsterdam, The Netherlands

The European Commission also received contact details to communicate directly with the Head of Snap's DSA Compliance Function on all matters.

[REDACTED]

Law Enforcement can contact through mechanisms described here:
<https://www.snapchat.com/lawenforcement>

Point of Contact for Users

Upon designation Snap's Cross-functional DSA Governance team designated a point of contact for users to contact Snap regarding DSA-related matters that is user friendly.



Upon designation, Snap makes information on the customer service POC publicly available in a place easily accessible to the user in a place where they would be expected to be and up to date.

For general DSA inquiries, Snap can be reached through the dsa-enquiries@snapchat.com email address as well as by submitting a ticket through our [Support Site](#). The support site is available to support users in all official languages of the EU.

Legal Representative

Snap's Cross-Functional DSA Governance team has designated a legal representative (Snap B.V.) in one of the Member States where the service is provided for complying with DSA obligations and enforcement matters. This designation includes allocation of resources and authorities sufficient to cooperate with relevant authorities and comply with decisions issued by the European Commission in relation to the DSA.

We have notified our Digital Services Coordinator of the contact information for our legal representative, including the name, mailing address, phone number, and email, and ensured the publication of the information in a publicly available location.

Snap has appointed Snap B.V. as its Legal Representative for purposes of the DSA. Snap's legal Representation can be contacted at

Snap B.V.
Keizersgracht 165, 1016 DP
Amsterdam, The Netherlands

6.2.10 DSA Supervisory Fee

Once required Snap will pay the DSA supervisory fee following the instructions provided by the Digital Services Coordinator. In its latest assessment, the European Commission has determined that Snap does not meet the threshold for the supervisory fee, and Snap's current contribution is set at EUR 0.

6.3 Privacy and Safety by Design

6.3.1 DSA Risk Management

We recognize that Snap has the duty to assess new products and functionalities prior to deployment to determine whether such new products and functionalities are likely to have a critical impact on the risks identified (and therefore the mitigations specified to prevent them).



[REDACTED]

This is required under Article 34 of the DSA and reflected in other legal frameworks, including the GDPR.

Snap has a Privacy and Safety by Design review process. Privacy and safety by design is a cornerstone of Snap's approach to designing and launching its products, and is built into Snap's compliance program. Snap has an extensive privacy and safety by design review process to assess privacy and safety risks in the design and development of Snapchat.

As part of its privacy and safety by design program, Snap documents a review prior to new product and feature releases that materially affect the privacy, safety and/or security of its users. The privacy and safety by design process is a collaborative and cross-functional process, and stakeholders from Snap's legal and privacy engineering teams are embedded in key phases of the product's development.

At Snap privacy and safety by design decisions are typically made by cross-functional teams. We have a long standing cross functional team across Product, Eng, DSA Compliance Officers, Operations, Policy, Legal, Comms, Trust & Safety, and Privacy teams which meets regularly to address risks flagged through various mechanisms such as industry reports, current events/news, internal data analyses and investigations, and feedback from regulators to assess problems, prioritize them and agree on strategy and execution plans to resolve identified risks. Internally the team is called the Safety XFN. The team also meets in person and virtually quarterly to align on priorities for the next quarter, and reflects on safety improvements that were made in prior quarters. These findings are also presented to senior leadership on a regular cadence.

[REDACTED]

6.3.2 Privacy and Safety by Design review process

Privacy and safety by design is a cornerstone of Snap's approach to designing and launching its products, and is built into Snap's compliance program. As highlighted above, at Snap, our mission is to empower people to express themselves, live in the moment, learn about the world, and have fun together. We believe that privacy and safety are foundational to the success of our mission.

Snap already had an extensive privacy and safety by design review process to assess privacy and safety risks in the design and development of Snapchat prior to the DSA coming into force, and this continues to be the case. As part of its privacy and safety by design program, Snap documents a review prior to new product and feature releases that may materially affect the privacy, safety and security of our users. [REDACTED]

[REDACTED]

[REDACTED]



[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]

- [Redacted]
 - [Redacted]
 - [Redacted]
- [Redacted]
 - [Redacted]
 - [Redacted]
 - [Redacted]
 - [Redacted]
 - [Redacted]

[Redacted]

[Redacted]

- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]



6.3.3 Holistic Digital Risk Management

Formal risk assessments and mitigation obligations are being an increasingly common tool of digital service regulation. In the European Union, such obligations are not only imposed by the DSA, but also, for example, the GDPR (in the form of Legitimate Interest Assessments (LIAs) and Data Protection Impact Assessments (DPIAs)) and also in the UK and several EU Member States, the Age Appropriate Design Code (AADC) Assessments (or their equivalent). It is important that Snap is able to manage these, often overlapping European requirements (in addition to other global requirements) in an efficient, effective and operationalised manner.

In order to meet Snap's risk assessment obligations, including with respect to the DSA, Snap continues to use a Digital & Data Impact Assessment ("DDIA") framework that combines our privacy, safety and security obligations into a single risk assessment. Each DDIA also covers areas related to specific features, transparency, privacy compliance, security compliance, safety compliance, and training.

The DDIA includes a template that serves as a vehicle to conduct the various risk assessments. This template supports the consideration of risk and mitigations related to a specific Snapchat product and includes guidance for the consideration of key factors and influencers on that risk, such as the performance of Snapchat recommender systems, the intentional manipulation of the platform, and regional and linguistic considerations. The template also requires specific consideration of recognised, international rights and risk assessment frameworks including the [UN Convention on the Rights of the Child](#) and the 5Cs risk categories set out by the OECD (Content, Conduct, Contact, Consumer, and Cross-cutting)²²⁵.

The DDIA template is embedded with our existing privacy and safety by design process and requires our cross-functional team to consider if a product change results in a significant impact on our existing consolidated DDIA assessment (including the DSA aspects). If so, this is required to be re-assessed before the product change launches. The new/updated/interim DDIA is completed dynamically depending on the nature of the Snapchat feature and impact of the change being assessed.

²²⁵ https://www.oecd-ilibrary.org/science-and-technology/children-in-the-digital-environment_9b8f222e-en



[REDACTED]	[REDACTED]
	[REDACTED]
	[REDACTED]
[REDACTED]	[REDACTED]
	[REDACTED]
	[REDACTED]
	[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED] is reviewed by Snap's cross-functional DSA Governance team on an annual basis.



6.3.4 Digital and Data Impact Assessment (DDIA) Template

In order to meet Snap’s annual and ongoing holistic risk assessment obligations, including with respect to the DSA, Snap has designed a new Digital & Data Impact Assessment (“DDIA”) framework that combines our privacy, safety and security obligations into a single risk assessment.

The DDIA templates are implemented at the product level and cover a range of requirements beyond the scope of the DSA. [REDACTED]

- [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]

- [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]
 - [REDACTED]

[REDACTED]

[REDACTED]



In each case, product reviewers use the template to highlight the aspects of the risk and mitigations particularly relevant to the feature to which the DDIA relates. Through these DDIA's, Snap considers the design of its critical recommender and algorithmic systems, terms and conditions, content moderation systems, enforcement, systems for selecting and presenting ads and considerations of data related practices. The analysis from each DDIA is then combined into this Report to enable a cross-platform view of risk.

The DDIA is embedded with our existing privacy and safety by design process and requires our cross-functional team to consider if a product change results in a significant impact on our existing consolidated DDIA assessment (including the DSA aspects). If so, this is required to be re-assessed before the product change launches. The DDIA is completed dynamically depending on the nature of the Snapchat feature and impact of the change being assessed.

If reviewers determine that the change does require an update to the DDIA, they will work with Snap's Legal team (and other cross-functional Compliance team that they engage as needed) to update the relevant DDIA accordingly.

6.3.5 DSA Critical Impact Check

Snap recognises that, as well as carrying out our annual DSA Risk Report, it must also re-assess risk prior to deploying functionalities that are likely to have a critical impact on the risks identified (and therefore the mitigations specified to prevent them). This is required by Article 34 of the DSA, but this is also an industry standard practice to ongoing risk management and found in many other laws requiring risk assessments (including guidance relating to DPIAs).

As part of the DDIA update review, product reviewers consider whether the change amounts to a critical impact. Generally the following criteria are considered in the determination of what is a critical impact:

- Products and Features that would materially change the likelihood or severity of the risks described in [Section 4](#) and related Mitigations in [Sections 5](#) and [6](#) of this Report
- Material impact to minors
- Material impact to advertising
- Material impact to algorithmic systems
- Material impact to DSA-relevant features

Product reviewers use their professional judgment to determine whether a material impact is expected based on the product-specific circumstances. They regularly consult with the DSA compliance officers, including when they have doubts whether a new product or new feature would have a critical impact on Snap's System Risk Assessment and related Mitigations.



6.4 Prevalence Testing

1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 2679, 26

-
- The image shows a document that has been almost entirely redacted with black bars. Only a few lines of text are visible at the bottom of the page, which appear to be a list or index of items, possibly names or titles, followed by what might be dates or page numbers. The rest of the document is obscured by solid black rectangles.

114

- _____



[REDACTED]

	T				

[REDACTED]

	T				
I					

[REDACTED]

These overall PVP metrics demonstrate that the effectiveness of our proactive detection mechanisms, agent training and other content moderation and enforcement efforts has increased significantly since our first 2023 DSA Report.

[REDACTED]



[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

	[Redacted]	[Redacted]	[Redacted]	[Redacted]
[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]
[Redacted]	[Redacted]	[Redacted]	[Redacted]	[Redacted]

[Redacted]

[Redacted]



[REDACTED]

[REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

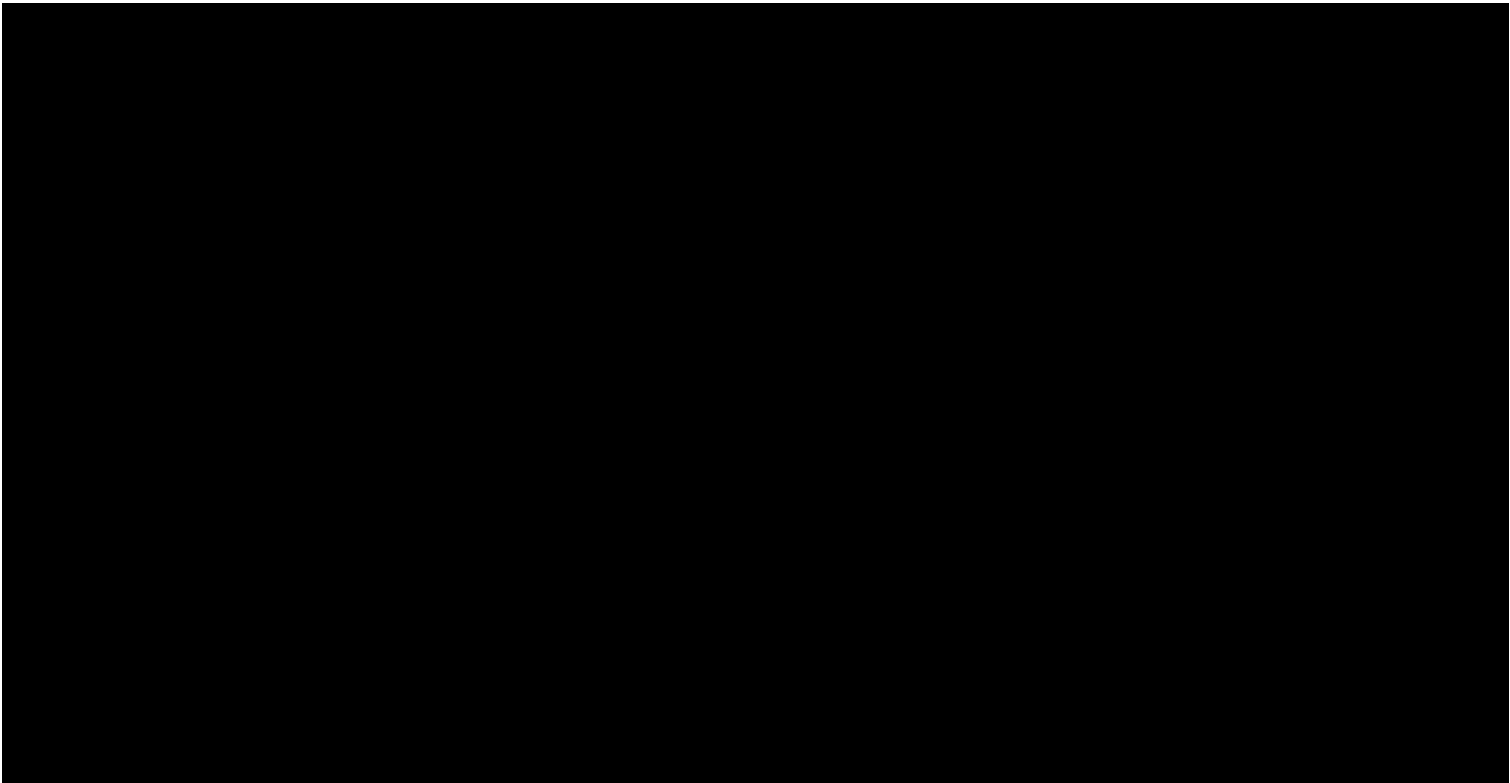
[REDACTED]

[REDACTED]

[REDACTED]	[REDACTED]		[REDACTED]		[REDACTED]		[REDACTED]	
	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

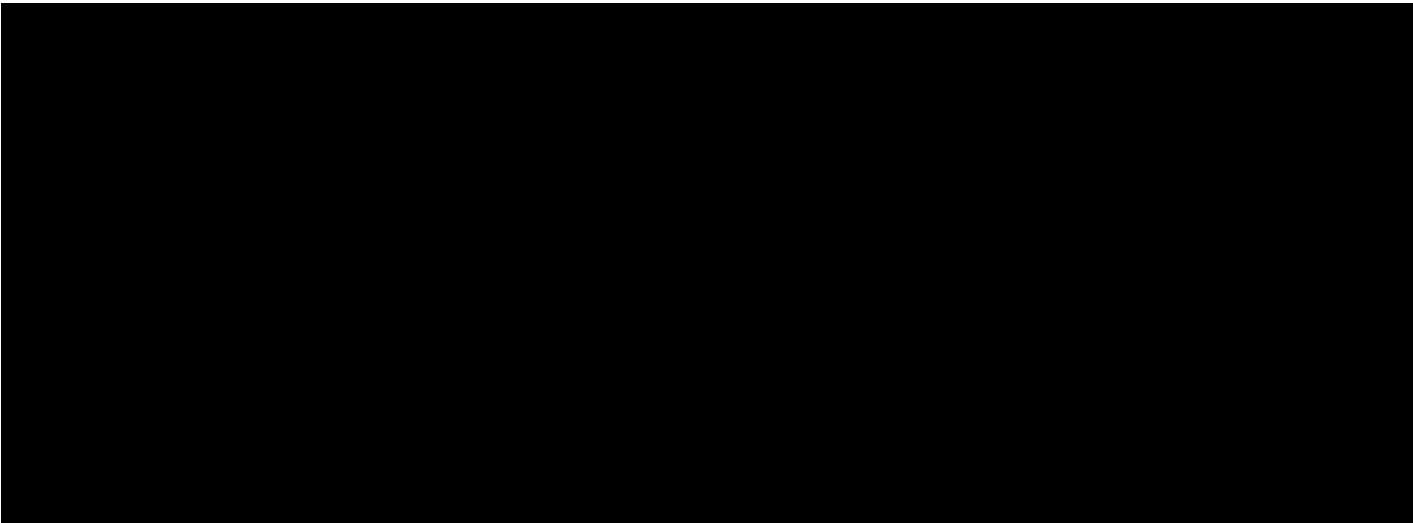
[illegible]

[REDACTED]



[Redacted text block]

[Redacted text block]



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

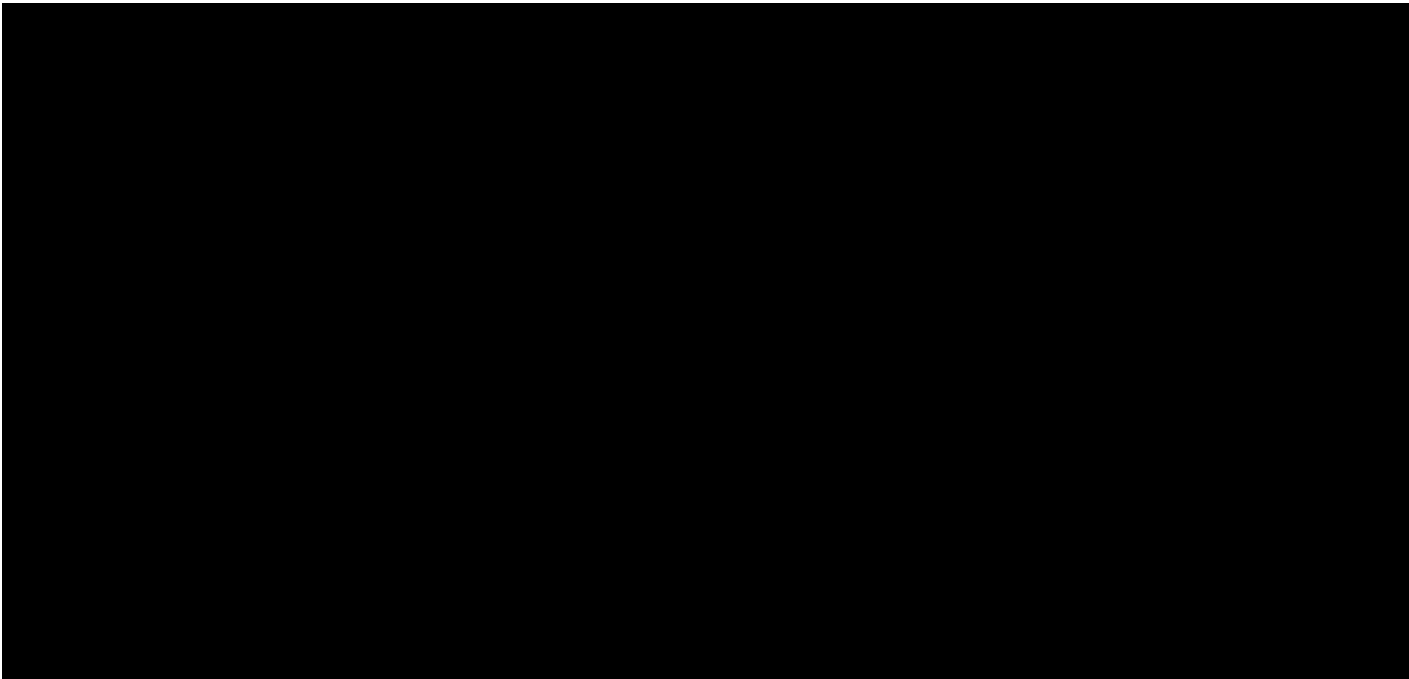
[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]





[Redacted text block]

[Redacted text block]

[Redacted text block]

[Redacted text block]

[Redacted text block]



[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]



	■ [REDACTED] ■ [REDACTED] ■ [REDACTED]
[REDACTED]	■ [REDACTED] ■ [REDACTED]

An example of the further mitigations we have taken since our 2023 Report (and as mentioned in our Report of 2024) is an improvement to the efficacy and reduced latency of content enforcement mechanisms. [REDACTED]

Conclusion

In conclusion, prevalence testing continues to be an extremely valuable measure for our ongoing detection and management of content risks. Our ongoing efforts to improve our prevalence testing and our mitigations to reduce prevalence of illegal and other violating content has resulted in further significant decreases in PVP rates [REDACTED]

[REDACTED] While our prevalence rates from our sample testing of Public content services on Snapchat are now extremely low, there is always more to do as we ultimately aim to reduce prevalence across all our violating content categories to as close to zero as possible.

6.5 External Request Monitoring and Review

As noted in our 2024 Report, we produce a semiannual (every 6 months) [Transparency Report](#), that captures our Community Guidelines enforcement data, law enforcement operations data, and copyright & trademark data. The goal is to provide insight into our content moderation data, as well as our work with law enforcement and governments, in terms of how we work to keep our users safe. As we produce the Report, we recognize shifts in our metrics (e.g., spikes or decreases in content and account reports and enforcements) and utilize these to inform heightened awareness from our moderation teams. Internally we also continue to review additional breakdowns of this data and, in preparation for our DSA compliance, we continue to review data relating specifically to the European Union’s individual Member States.



We also continue to monitor advertising review rejections, advertising reporting and enforcements, ‘privacy, data protection and DSA’ requests and general community support requests. Since the DSA came into force for Snapchat on 25 August 2023, we have also monitored queries relating to compliance received via our dedicated dsa-enquiries@snapchat.com email address. This dedicated contact point is published on our website here, pursuant to Articles 11 and 12 of the DSA.

As with our 2024 Report, we have continued to review and use this external request data to support the conclusions reached in this Report.

6.6 Digital Well-Being Index (DWBI) Initiative

In the Spring of 2022, Snap launched a research project designed to gain insight into how Generation Z teens and young adults are faring online. Our inaugural Digital Well-Being Index ([DWBI](#)), a measure of Generation Z’s online psychological well-being, was announced on Safer Internet Day 2023. The study asked about the risks and potential harms teens and young adults are encountering online across all platforms, services and devices, not just Snapchat. We conducted the research in six countries – Australia, France, Germany, India, the UK, and the U.S, which includes three of the largest European countries, two of which are in the EU) – and also included parents of teenagers between the ages of 13 and 19.

In our 2024 Report, we explained that we had repeated and expanded this research in 2024 (“Year Three”). For more about Snap’s Digital Well-Being Index and research, see: Our [website](#), as well as this [explainer](#), the [full research results](#), and each of the six country infographics: [Australia](#), [France](#), [Germany](#), [India](#), the [United Kingdom](#) and the [United States](#). We took account of this and the previous year’s research when conducting our assessment of risk and mitigations as highlighted in our 2024 Risk Report, and have continued to do so in this Report.

Lastly, it is important to note that on March 19, 2025, we hosted an event to share the latest findings from the *Digital Well-Being Index (DWBI)*, followed by an informal discussion with subject-matter experts. The event received positive feedback and generated thoughtful recommendations for future iterations of the research. In particular, participants underscored the need to further investigate teens’ exposure to AI-generated child sexual abuse material, as well as the evolving role of AI tools, such as chatbots, in shaping young people's online experiences.

Seven experts from four prominent organizations—5Rights, Eurochild, Missing Children Europe, and ThinkYoung—joined the conversation, offering valuable insights and perspectives that enriched the dialogue.



6.7 Snap Advisory Groups

We have continued to progress our work with the Snap Safety Advisory Board, the Snap Council for Digital Well-Being, and our regular outreach to civil society organisations. We have established a new dedicated workshop with civil society organizations to focus specifically on our DSA risk assessment. We have provided more detail on each of these groups below.

6.7.1 Safety Advisory Board

Snap launched a new [Safety Advisory Board](#) (SAB) in April 2022 with the aim of growing and expanding membership to include a diversity of geographies, safety-related disciplines and areas of expertise. In doing so, we initiated an application process, inviting experts and individuals from around the world to formally express their interest in providing guidance and direction to Snap on “all things safety.”

The SAB Board was developed to educate, challenge, raise issues with, and advise Snap on how best to keep the Snapchat community safe and counterbalance the online harms-dominated external landscape. When appropriate, the SAB provides feedback on new products, features, policies, and initiatives before they are launched or released. The SAB and its individual members do not act as a representative or spokesperson for Snap, but rather as a collection of independent voices. The initiative helps to shape Snap’s approach to important safety issues and provides Snap with strategic safety-related advice and guidance as Snap grows.

Our Advisory Board currently stands at 19 members, based in 11 countries and representing 12 different geographies and regions. Four of our members are based in the UK. The board comprises 16 professionals from traditional online safety-focused non-profits and related organizations, as well as technologists, academics, researchers, and survivors of online harms. Members are experts in combating significant online safety risks, like child sexual exploitation and abuse and lethal drugs, and have broad experience across a range of safety-related disciplines. In addition, the Board has three members who are young adults and youth advocates. We selected these applicants to ensure the Board has ready access to the all-important “youth voice” and viewpoint; to make certain a portion of the Board includes committed Snapchat users; and to seek to balance professional views with practical perspectives from a core demographic of the Snapchat community. The SAB meets three times annually: twice virtually and once in-person at Snap headquarters for an in-depth strategy session to help prepare Snap for the coming year’s planning.

In June 2025, at the third in-person board meeting of the SAB, members were divided into working groups to discuss various topics such as age assurance, screentime, and financial sextortion. They were also given updates on some new efforts, including our youth-focused program (see Snap Council for Digital Well-Being below). The SAB’s input and feedback was memorialized and shared with Snap executives and others that were unable to attend in-person.

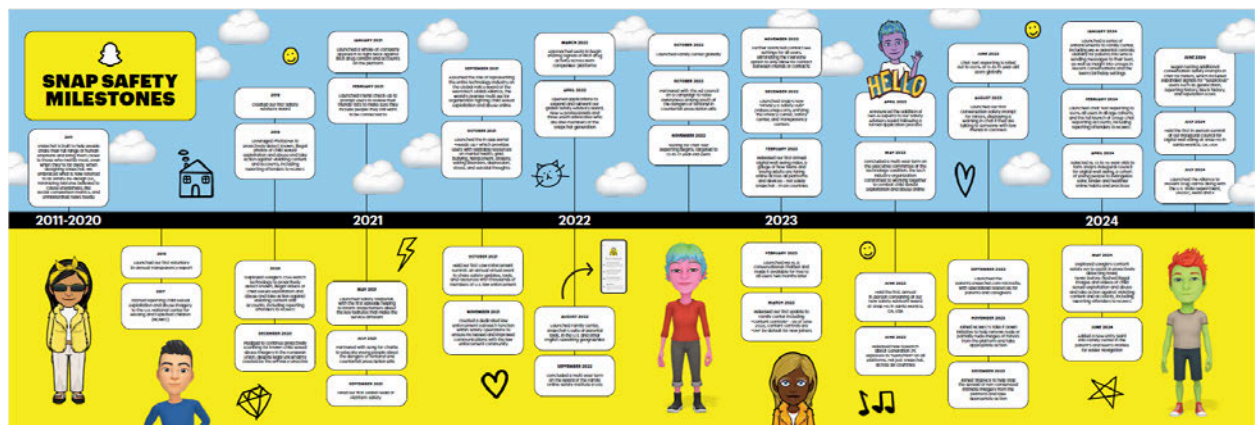


6.7.2 Snap Council for Digital Well-Being

Snap formally launched a new youth-focused program in January 2024 - Snap's first “Council for Digital Well-Being (CDWB)”. The Council is a pilot program in the U.S. designed to encourage safer online habits and practices among teens with the aim of having these young people champion their knowledge and insights in their schools and communities.

For the first interaction of this program, we opened [applications](#) from U.S.-based teens, aged 13 to 16. The inaugural cohort was selected and [announced](#) in May. Following two virtual monthly meetings, the group traveled to Snap HQ in Santa Monica, California, USA, in July 2024 for the first [Council Summit](#). The 2-½-day program consisted of breakout sessions for both the teen cohort and their parents/chaperones, full-group discussions, and guest speakers. The teens also got a glimpse into working at a technology company, as they were treated to a 90-minute “speed-mentoring” session with 18 Snap employees representing different roles and teams.

The Summit yielded interesting conversations and insights on topics such as online pitfalls, parental tools, and the differences and similarities between digital and in-person social dynamics. By the end of the Summit, the full group, chaperones included, was extremely motivated to be more involved in their own local communities and to act as ambassadors for online safety. We shared with them some of our outreach material to aid their efforts, including the below infographics on reporting and Snap Safety Milestones:





Safety Concerns

2024

Reporting a Safety Concern

Snap empowers people to express themselves, live in the moment, learn about the world, and have fun together. One of the most important things Snapchatters can do to help keep the service free of bad actors and potentially harmful content is to **reach out** to us when you encounter something that makes you uncomfortable. All you need to do is **press and hold** on the piece of content or the chat message and a menu will appear. Then, tap **"Report"** to see a menu of options.

Our safety teams work 24/7 to review reports made on Snapchat or through our Support Site and, once reviewed, our safety teams will take action on content and accounts that violate our [Community Guidelines](#) or [Terms of Service](#). It's important to remember that reporting is confidential and the account-holder you reported won't be told who reported them. If you encounter anything that appears to be illegal or dangerous, or if you have reason to believe someone is at risk of harm or self-harm, **contact local law enforcement** immediately and then report it to Snapchat, as well.

You can read through Snapchat's [Community Guidelines](#) and our [Terms of Service](#) to familiarize yourself with what content is permitted on Snapchat. A good rule of thumb: if what you're saying could create an unsafe or negative experience for someone, it's better left unsaid.

Also, if you see something you don't like on Snapchat, but it may not violate the [Community Guidelines](#), you can choose to **unsubscribe**, **hide the content**, or **unfriend** or **block the sender**.

Your Common Questions Answered

Is reporting on Snapchat confidential?

Yes. We do not tell other Snapchatters when you make a report.

Who reviews my submitted report?

When you report a concern on Snapchat, you receive a confirmation that your report has been submitted. Behind the scenes, our safety teams work 24/7. If the teams' review confirms a violation of our [Community Guidelines](#) or [Terms of Service](#), the content will be removed and we may even lock or delete the account, and report the offender to authorities.

Does Snapchat tell the reported account who reported them?

No. The account holder that you reported won't be told who reported them. All reports are strictly confidential.

If I block or remove someone, will they know?

When you block or remove someone from your Friend List, they are not formally notified, but they may be able to infer this when their messages are no longer reaching you.

Does Snapchat alert me if someone reports me?

If we take action on your content that was reported, you may be alerted in our app or via email.

I reported something on Snapchat but it wasn't taken down. Why is this?

Not all reported content is removed. We remove content that violates our [Community Guidelines](#) or [Terms of Service](#). If you see content that you don't like, but is permitted according to our [Community Guidelines](#) or [Terms of Service](#), you can avoid seeing it by adjusting your privacy settings, hiding the content or blocking and removing the sender.

The inaugural U.S. program will conclude this summer, and we are fashioning an "alumni network" to continue to update the teens on major safety initiatives and developments and to encourage their continued advocacy.

We are now replicating these efforts in Europe. We recently finalized the selection process for the first European (including the UK) Council for Digital Well-Being ("E-CDWB") and held our first European Teen Council for Digital Well-Being Summit in Amsterdam at the end of June 2025. The E-CDWB is made up of 14 young people aged 13 to 16 and three members are from the UK. The Summit hosts the teens and their parents/chaperones to hear from about their online life, the things they enjoy, the value it brings to them, the worries they have and their ideas for the future.



We will be looking to incorporate their thoughts into future plans for Snapchat, and our risk assessments. Similar efforts are also underway in Australia.

6.7.3 Regular External Engagement

In addition to the deep investment we are making in our SAB programme Snap also regularly engages with external experts and civil society organisations. Snap periodically consults a cadre of some 50 safety experts from around the world on new product features and functionality, policies, and initiatives. Snap also conducts periodic internal trainings and learning-sessions, inviting external experts to help inform and educate Snap personnel working in a variety of safety disciplines about the overall risk landscape and Snap's potential exposure. This includes collaborations with the U.S.-based National Center for Missing and Exploited Children (NCMEC) on the topics of sextortion and improving CyberTip reports to NCMEC, as well as smaller, executive-attended sessions with WeProtect Global Alliance, IWF, Thorn, and others. Snap will continue to invest in these and other external partnerships and relationships to help bolster internal knowledge and awareness of the overall risk landscape.

6.7.4 Dedicated DSA Risk Assessment Workshop

On July 10, 2024, Snap held our first dedicated DSA Risk Assessment and Mitigation workshop with civil society organisations (CSOs) in Brussels and online. The session provided a detailed overview of Snap's 2024 DSA Risk and Mitigation Assessment Report and aimed to collect feedback from relevant privacy and safety CSOs based in Brussels and across EU Member States. This initiative also served as a follow-up to the European Commission's workshop on systemic risk mitigation held on May 7, during which several stakeholders expressed interest in more structured opportunities to provide input to platforms.

The workshop brought together 19 experts representing 14 NGOs. *Participating CSOs:* Child Focus, COFACE-Families Europe, E-Enfance (3018), ECPAT International, Eurochild, European Parents' Association, European Schoolnet, FAFCE, Media Council for Children & Youth (SIC) (Medierådet for Børn og Unge), Missing Children Europe, Offlimits, Point de Contact, Saferinternet.at, ThinkYoung.

Participants appreciated the opportunity to engage deeply with Snap's risk assessment process and shared constructive feedback that will inform future reports. This included in particular:

- Requests for additional clarity on the scope of risk categories - for example with respect to well-being risks affecting minors and adults. In response to this feedback, we have included a more expansive harm description for each harm in [Section 4](#).
- Request for more risk-specific mitigation highlights. In response to this feedback, we have merged the 'highlights' paragraphs into the specific mitigations table under the 'Snap's mitigations' heading for each harm in [Section 4](#). This streamlines the Report and



[REDACTED]

makes it easier for readers to locate the corresponding detail we have provided about our specific mitigations in [Section 5](#).

- Request to explain how we reference frameworks such as the UN Charter on the Rights of the Child in our risk assessments. In response to this feedback, we have updated [Section 6.3](#) (Privacy and Safety by Design) to be clear that Snap has already embedded well-known, internationally recognised frameworks into its Privacy and Safety by Design assessment processes, including the [UN Convention on the Rights of the Child](#) and the 5Cs risk categories set out by the OECD (Content, Conduct, Contact, Consumer, and Cross-cutting)²²⁷. Consideration of these frameworks is specifically required by our holistic Digital & Data Impact Assessment templates used to assess critical product updates.
- Request to ensure we have explained how we have assessed and mitigated the risks relating to harmful AI-generated content, including content affecting body image. We already cover this in [Section 5.9](#) (Content Authenticity). This will become clearer to civil society organisations when our 2024 Report, and this Report, are published.

Participants praised Snap's use of data to intelligently target safety interventions and support mechanisms. We were recommended to build on this strong element of the company's risk mitigation approach by further expanding in-app support resources and reminder features. This is something that we will be reviewing during the following year.

We found this workshop helpful, with the vast majority of participants focused on constructive dialogue and developing practical solutions to achieve common goals. Snap intends to continue and build on this engagement with participating CSOs to inform future DSA risk and mitigation assessments and provide transparency on how their input is reflected.

6.8 Audit

We completed the first external DSA audit of Snap's compliance with its obligations under Chapter 3 of the Digital Services Act for the audit period between July 1st 2023 and June 30th 2024 pursuant to Article 37. [REDACTED]

[REDACTED]

[REDACTED]

²²⁷ https://www.oecd-ilibrary.org/science-and-technology/children-in-the-digital-environment_9b8f222e-en



7. Conclusion

This Report has been prepared to meet Snap’s obligation under Article 42(4) of the DSA and sets out the results of: (i) the risk assessment conducted by Snap pursuant to Article 34(1); and (ii) the review of the specific mitigation measures that Snap has put in place to assess whether they meet the requirements of Article 35(1) DSA.

The risk assessment conducted by Snap identified, analyzed and assessed in accordance with Article 34(1) DSA any systemic risks in the European Union stemming from the design, functioning or use of Snapchat’s in-scope services. Snap has also reviewed the specific mitigation measures that it has put in place to ensure they are “reasonable, proportionate and effective” for the specific systemic risks identified by its risk assessment as required by Article 35(1). The results of the risk assessment and mitigation review are set out in [Section 4](#) of this Report. The specific mitigation measures put in place by Snap are further detailed in [Section 5](#) and [6](#) of this Report as required by Article 42(4).

The Report shows that we have reasonable, proportionate and effective mitigation measures in place and we continue to monitor a few areas to confirm that if additional measures are required Snap will act accordingly, as follows:

- 1. Dissemination of illegal or violating content:** Since our previous Reports, we have observed a further substantial reduction in the prevalence of content that is illegal or otherwise violating Snap’s Terms being disseminated on Snapchat’s online services in general [REDACTED]. We have observed PVP rates for this content to now be at a very low level [REDACTED] compared to the prevalence of this content on websites and other online spaces. As we have previously noted, historical studies suggest that around 4% of websites have adult sexual content which is much higher than the level of this content on Snapchat’s in-scope services. This year, all of our illegal or other violating content categories were observed from our [testing](#) to have a prevalence rate to fall within our



lowest relative prevalence category (which is a rate of 0.049% or below). Within this very low level of dissemination in general:

- a. We continue to categorize two dissemination risk areas as falling within Level 1 risk prioritization for Snapchat's in-scope services: (i) child sexual abuse material, and (ii) sale of drugs. We have again confirmed we have reasonable, proportionate and effective mitigation measures for all two of these categories. As a result of these ongoing measures, these two categories are still assessed to fall within our lowest likelihood category of the risks identified by Snap.
- b. We have continued to categorize five dissemination risk areas as falling within Level 2 risk prioritization for Snapchat's in-scope services: (i) Sale of Weapons, (ii) Terrorist Content, (iii) Adult Sexual Content, (iv) Harassment and Bullying and (v) Self-Harm and Suicide, due to the risk of serious harm each may cause. We have again confirmed we have reasonable, proportionate and effective measures in place for all five of these categories. As a result of these measures all five are assessed to fall within our lowest likelihood category of the risks identified by Snap. This is a change from our 2024 report, when bullying and harassment was assessed to fall within our medium relative likelihood category. **In the case of terrorism content, we have continued to observe a slight increase in prevalence since our previous reports. It is normal to expect some natural fluctuation when the PVP percentages are so low, as they are for terrorism content, and the increase is not considered statistically significant at this time. We have also made some improvements and expanded our prevalence testing to cover new areas that may have resulted in a slight increase in the volume of terrorism cases identified during our prevalence testing. The PVP percentages for terrorism remain extremely low relative to other harms and we continue to closely monitor prevalence levels for this harm as one indicator of new or increased terrorism threats on Snapchat.**
- c. We have continued to categorize eight dissemination risk areas as falling within Level 3 risk prioritization for Snapchat's in-scope services: (i) illegal Hate Speech, (ii) sale of prohibited products or services (excluding Drugs and Weapons); (iii) intellectual property infringements, (iv) other Adult Sexual Content,²²⁸ (v) Violent or Dangerous Behaviour, (vi) Harmful False Misinformation, (vii) Fraud and Spam and (viii) content relating to Other Illegal Activities. We have again confirmed we have reasonable, proportionate and effective measures in place for all eight of these categories. As a result of these measures all eight of these risk areas are assessed to fall within our lowest likelihood category of the risks identified by Snap. This is a

²²⁸ With regard to the Adult Sexual Content category, sexual crimes are treated as Level 2, and other forms of Adult Sexual Content are prioritized as Level 3. See [Section 4.1.6](#) for more detail.



change from our 2024 report, when fraud and spam and adult sexual content were assessed to fall within our medium relative likelihood category.

2. **Negative effects on EU Fundamental Rights:** We continue to categorize: (a) three risks to fundamental rights as falling within the Level 1 priority category for Snapchat's in-scope services: (i) Human Dignity, (ii) Data Protection and (iii) Children's Rights; (b) one risk as falling within the Level 2 priority category: Private Life; and (c) three risks as falling within the Level 3 priority category: (i) Right to Freedom of Expression, (ii) Right to Non-Discrimination and Freedom of Religion and (iii) Right to Consumer Protection. We have again confirmed we have reasonable, proportionate and effective measures in place for all of these categories. **We continue to assess Snapchat's in-scope services against the Commission's newly published guidance on Art 28 to ensure a high level of privacy, safety and security for minors to assess if further industry measures are needed to address risks to child rights.**
3. **Negative effects on Public Security:** We continue to categorize three risks to public security within the Level 3 priority category for Snapchat's in-scope services: (i) Negative Effects on Democratic and Electoral Processes; (ii) Negative Effects on Civic Discourse and (iii) Negative Effects on Public Security. We have again confirmed we have reasonable, proportionate and effective measures in place for all of these categories.
4. **Negative effects on Public Health:** We continue to categorize: (a) two risks to Public Health within the Level 1 priority category for Snapchat's in-scope services: (i) Negative Effects on Children; and (ii) serious negative consequences on Physical and Mental Well-Being; (b) one risk within the Level 2 priority category for Snapchat's in-scope services: Negative Effects on Gender-Based Violence; and (c) one risk within the Level 3 priority category for Snapchat's in-scope services: Negative Effects on Public Health. We have again confirmed we have reasonable, proportionate and effective measures in place for all of these categories. **As above, we continue to assess Snapchat's in-scope services against the Commission's newly published guidance on Art 28 to ensure a high level of privacy, safety and security for minors to assess if further industry measures are needed to address risks to the protection of minors.**

It is Snap's mission to reduce and maintain a lower prevalence of illegal and otherwise violating content on Snapchat's inscope services. [REDACTED]

[REDACTED] We have been successful in increasing the granularity of data that we rely on, which we have incorporated into this Report. We will continue to progress this action to ensure that risks can be tracked with even greater precision across in respect of each of Snapchat's in-scope services.



Regarding our [Mitigations](#), since our previous Reports, we have worked to evaluate our mitigation measures against new guidance from the Commission (in particular the Commission's Guidelines on Article 28, the Guidelines for providers of VLOPs and VLOSEs on the mitigation of systemic risks for electoral processes and the updated EU Code of Conduct to counter illegal Hate Speech online). We have also updated the information on our measures to reflect the increased use of generative AI technology such as updates to our content moderation policies (note that many of these measures relate to out of scope services on Snapchat). Where required by the DSA, we have made available our mitigations in all of the languages of the European Union and in all other cases in every language in which Snapchat is available.

Since our previous Reports, we did not identify any deployed functionalities that were likely to have a critical impact on our assessment of risks and mitigations pursuant to Articles 34 and 35 of the DSA. As described in the [Ongoing Risk Management](#) section above, our DSA Governance Team continues to regularly evaluate the effectiveness of its measures as we look to maintain or further reduce prevalence, detect any new risks, assess any deployed functionalities for critical impacts and determine whether further mitigating measures might be required.

In summary, we have carried out our third annual risk assessment of Snapchat's in-scope services required by Article 34(1) of the DSA. We have observed further significant reductions in the prevalence across our illegal and otherwise violating content categories. We continue to conclude that we have reasonable, proportionate and effective mitigation measures, tailored to the specific systemic risks identified, as required by Article 35(1) of the DSA. However, there are two high risk categories that we continue to monitor to ensure their corresponding mitigations continue to be effective.



8. Final Words

This is the third year in which VLOPs have had to produce a report on their assessment of risks and the specific mitigation measures they have put in place. Snap has continued to take a comprehensive approach to the obligations in Articles 34, 35 and 42. As in our 2024 Report, although there is still no settled legal definition of ‘systemic risk’, we have again adopted the position that all the risks identified in the DSA are systemic to online platforms (which is why they have been identified in the DSA). We are still then looking to ensure we have appropriate platform wide measures in place in general, taking additional steps for specific risks, certain services and high priority risks as necessary.

As noted in our 2024 Report, our position reflects Snap’s own internal approach to risk management and our core values to be kind, smart and creative. We have always taken the assessment of privacy and safety risks and mitigations seriously and this is demonstrated again in this Report which concludes that Snapchat represents an even lower risk profile than identified in our 2024 Report. This is due to its unique design and function, but also in particular to the efforts of our cross-functional teams who have worked hard to provide high levels of privacy, safety and security for all our users and further substantial falls in the risks specifically referred to in the DSA. We are particularly proud of the further **46% decrease** in the overall prevalence rate for our illegal and harmful content categories.

We are pleased to have confirmed that we meet all of the recommendations for risk assessment and mitigation reports provided by the Commission in its workshop on 7 May (as set out in the [Foreword](#)). We have also introduced some enhancements to further improve our approach. We look forward to again receiving feedback from the Commission on this third Report, as well as the publication of our second Risk Assessment and Mitigation Report and our Audit Report.



Annex

Contents

Annex..... 552

 Community Guidelines..... 553

 Overview..... 553

 Community Guidelines: Explainer Series..... 555

 Sexual Content..... 555

 Threats, Violence & Harm..... 557

 Hateful Content, Terrorism and Violent Extremism..... 559

 Harassment and Bullying..... 561

 Illegal or Regulated Activities..... 563

 Harmful False or Deceptive Practices..... 565

 Severe Harm..... 568

 Snapchat Moderation, Enforcement, and Appeals..... 568



[Redacted]

Community Guidelines

[Overview](#)

[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]



[Redacted]

[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]

[Redacted]

[Redacted]
[Redacted]

- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]



Community Guidelines: Explainer Series

[Sexual Content](#)

Community Guidelines Explainer Series

Updated: February 2025

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]



[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]



[Redacted]

Threats, Violence & Harm

Community Guidelines Explainer Series

Updated: February 2025

[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]

[Redacted]
[Redacted]
[Redacted]

[Redacted]

- [Redacted]
[Redacted]
- [Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]
- [Redacted]
[Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]
- [Redacted]



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



[Hateful Content, Terrorism and Violent Extremism](#)

Community Guidelines Explainer Series

Updated: February 2025

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



[Redacted]

[Harassment and Bullying](#)

Community Guidelines Explainer Series

Updated: February 2025

[Redacted]

[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]

[Redacted]

[Redacted]
[Redacted]
[Redacted]

[Redacted]

[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]



[Redacted text]

[Redacted text]

[Redacted text]
[Redacted text]
[Redacted text]
[Redacted text]
[Redacted text]
[Redacted text]

[Redacted text]
[Redacted text]
[Redacted text]



[Redacted]

[Illegal or Regulated Activities](#)

Community Guidelines Explainer Series
Updated: May 2025

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

- [Redacted]

- [Redacted]

[Redacted]



[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]

[Redacted text]



[Redacted]

Harmful False or Deceptive Practices

Community Guidelines Explainer Series

Updated: February 2025

[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]

- [Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]
- [Redacted]

[Redacted]

[Redacted]
[Redacted]

[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]



[Redacted]

- [Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]
[Redacted]
- [Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]

[Redacted]

[Redacted]
[Redacted]

[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]



[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]



[Redacted]

[Severe Harm](#)

Community Guidelines Explainer Series

Updated: December 2023

[Redacted]

[Redacted]

- [Redacted]
- [Redacted]
- [Redacted]

[Redacted]

[Snapchat Moderation, Enforcement, and Appeals](#)

Community Guidelines Explainer Series

Updated: March 2025

[Redacted]



[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]



[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]

- [REDACTED]
- [REDACTED]
[REDACTED]
[REDACTED]
- [REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]