SEE WHAT LIES BENEATH THE SURFACE

ENERGY + TOPCODER DATA SCIENCE & ANALYTICS

LETTER FROM THE CEO

Every company wants to innovate, move faster, boost safety and evolve their organizations to get incredible results. The challenge is staffing the right people, accessing and leveraging emerging technologies, and adding change without disrupting operations. Our energy customers understand the need for next-generation solutions and experience without sacrificing quality. That's why they've used Topcoder to get groundbreaking design, data science, and development work done with crowdsourcing, a delivery model we've pioneered and honed over nearly twenty years.

Crowdsourcing allows companies to experiment, scale, and get high - end results from a global talent pool of developers, designers, data scientists, and testers — all available 24/7, on demand. With the skill, speed, scale, flexibility, and outcome-based pricing Topcoder provides, energy customers have been able to roadmap and accomplish more. Together, we can think bigger, aim higher, and create winning solutions. The future is here, and we're thrilled to be with you.

Mike Morris
Topcoder CEO

AN INTRODUCTION TO TOPCODER AND CROWDSOURCING

ON-DEMAND GLOBAL TALENT AND DIGITAL SOLUTIONS AT SCALE

As technology changes, so do expectations. But the time and technical expertise to exceed evolving customer expectations and get to market faster can be difficult to find. Topcoder, the leading crowdsourcing platform, makes it easy to quickly turn ideas and requirements into incredible digital solutions — with the help of the world's largest talent network. Not only do you pay only for results, rather than hours, but you also get a secure project management process, a dedicated project manager, and a team of experts available 24/7.

WHY CREATE SOMETHING GREAT WITH TOPCODER?

SEE HIGH-QUALITY RESULTS – FAST

With the ability to run concurrent workstreams on the same platform, get design, development, data science, and QA work done in record time.

GET A TEAM FOR THE PRICE OF A FREELANCER

Don't tie your project's success to one freelancer's time and expertise. Get a team of innovative designers and developers, as well as end-to-end project management.

PAY FOR OUTCOMES, NOT HOURS

No matter how many hours or iterations it takes to achieve success, you pay only for the result.

WHAT CAN YOU GET DONE WITH TOPCODER?

DESIGN AND DEVELOPMENT

One-of-a-kind designs and expert code for your next app or web project.

DATA ANALYTICS AND ALGORITHMS

Do more with your data. Quickly optimize algorithms and develop predictive models.

QUALITY ASSURANCE AND TESTING

Real-world testing and UI/UX suggestions from incredible testers — plus, best-in-breed QA engineers.

AI-POWERED SOLUTIONS

Everything from voice, text, and image analysis solutions to smart chatbots.

RAPID PROTOTYPING AND MVPS

Turn multiple solution concepts into designs, prototypes, and MVPs in a matter of days.

ENERGY, NATURAL RESOURCES & UTILITIES CASE STUDIES

OPTICAL CHARACTER RECOGNITION - MIXED FORMATS

MUD LOG DIGITIZATION

Converting paper records into digital format can be challenging, particularly when the records have various types of data, each with their own important meaning.

Mud logging, with its varied formats, diversity of graphed data, as well as hand-written annotations, presents a significant challenge for digitization efforts. However, a recent Topcoder customer approached us to assist them with exactly this challenge because they recognized that the valuable data contained in mud logs, if digitized, can unlock new analytical capabilities, speed-up decision-making, and increase confidence in geological interpretations.

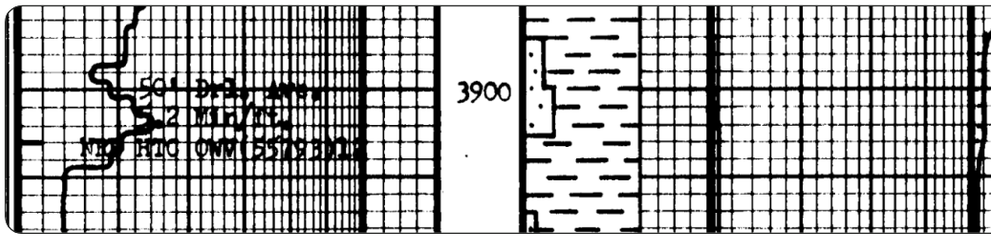
THE OUTCOMES

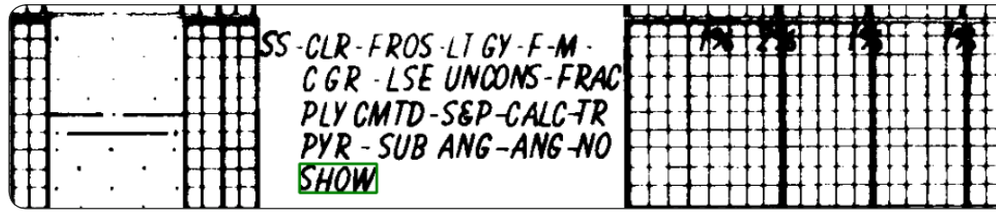
50+

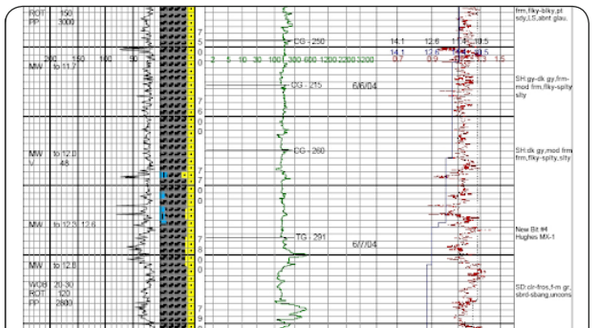
YEARS SAVED IN MUD LOG PROCESSING

The mixed-format of mud log data does not necessarily lend itself well to fully unassisted, automated approaches. To re-enter all data manually would be cost and time prohibitive (to say nothing of error prong, or other issues) so another hybrid approach was needed.

Through a series of competitions in the Topcoder community, a solution was built that combined human-assisted work, automated OCR, and machine learning algorithms to extract structured meaning from mud log records. This included digitized logging of shows, stains, traces, and negatives at respective drilling depths. By turning to Topcoder, this customer has saved 50+ years in mud log processing time and added millions in potential revenue through the value generated.







MACHINE LEARNING - OBJECTIVE OPTIMIZATION

AQUIFER BOUNDARY IDENTIFICATION

Cross contamination of water and wellbore fluids in oil wells is costly and unsafe for both workers and people living near the contamination sites. However, the process of identifying aquifer boundaries across hundreds or even thousands of wells is time-consuming and prone to human error, making it difficult to perform at scale. Our customer sought a method to rapidly identify aquifer boundaries in unidentified wells using algorithmic processing of gamma ray and resistivity logs measured at a range of depths within the well.

In Topcoder's long-form competition, known as a Marathon Match, contestants were provided with a dataset of well log data from 609 wells and asked to create an algorithm capable of identifying the depth of aquifer boundaries for eight different strata in a well at a given location.

With 349 algorithms submissions received, the algorithm results were compared with actual aquifer boundary locations to select the best-performing solutions. The winning submission combined 3D-positioning interpolation with random forest algorithms to estimate the depths of aquifer boundaries with an average depth error of less than one foot and false positive and false negative rates of just 1.38% and 2.02%, respectively, across all eight strata.

THE OUTCOMES

349 ALGORITHM SUBMISSIONS

1.38% FALSE POSITIVE RATE

2.02% FALSE NEGATIVE RATE

COMPUTER VISION - ACCIDENT PREVENTION

PROXIMITY HAZARD DETECTION

Struck-by and caught in hazards cause 3 out of every 5 oil and gas industry fatalities.[1] Yet the safety risks that are associated with oil and gas extraction activities are often preventable when appropriate measures are taken in advance. Even with thorough and routine re-certification on safety and accident prevention training, vehicle operation, material handling (crane, forklift, winch truck), and rig or equipment repair and maintenance remain the most common causes of fatalities in the industry.[2]

With the long hours and physical demands of work at drilling sites, it's no wonder that even experienced workers make mistakes. One Topcoder customer in the energy and utilities space sought to find a technology-driven solution to preventing accidents from occuring. They enlisted our help to create a computer vision tool to help determine the level of hazard under various vehicle and equipment operation scenarios.

With this in mind, a series of competitions were held to create a tool capable of analyzing video input and detecting changes in vehicle and equipment direction, changes in the presence and location of people, and the behaviors of workers in the vicinity (things like mobile phone usage, smoking, or other potential hazards and distractions).

In industrial settings, this tool can be configured to alert machine or vehicle operators of risks in real-time or to trigger emergency-stop procedures under specific circumstances. This way, when the new worker on the job makes a mistake or the sleep-deprived veteran slips up, the results will less often be fatal.

[1] https://www.osha.gov/SLTC/oilgaswelldrilling/safetyhazards.html#struck

[2] https://www.cdc.gov/niosh/docs/2017-193/2017-193.pdf?id=10.26616/NIOSHPUB2017193

MACHINE LEARNING - DATA LABELING

FAULT IDENTIFICATION

The identification of faults in oil and gas fields is critical to identifying promising locations and directions for drilling as well as avoiding unnecessary economic losses and environmental risks during operations. A customer approached Topcoder with an interest in both automating and enhancing the time-consuming manual work of labeling 3D seismic volumes as well as teasing out the relationships between intersecting faults.

To make this customer's fault identification desires a reality, Topcoder first turned to its community to build a visualization tool capable of taking 3D array inputs and displaying the probabilities of various labelings and interrelationships. After this a larger competition was held, using the customer's sample data along with the newly created visualizer to apply labels to each discrete crack found in the dataset and tease out their relationships.

With 150 submissions representing a variety of problem solving approaches, a winner was found by comparing competitors' work against existing ground truth labels. The top solution was adept at identifying multiple classes of faults and their relationships to one another at a speed and degree of accuracy that provides a significant benefit to our customer in the form of agility, decision-making, and profitability.

COMPUTER VISION - SAFETY COMPLIANCE

WORKER SAFETY PPE

Proper use of the right personal protective equipment (PPE) is critical to ensuring the safety of oil and gas workers as well as maintaining compliance with Occupational Safety and Health Administration (OSHA) requirements. With an occupational fatality rate more than three times greater than the overall rate for U.S. workers, finding any way to reduce the risk that workers face is essential.

With the rise of computer vision technology, new opportunities are available to reduce injuries and fatalities by creating systems that are ever vigilant of workers' proper PPE use. When an energy industry customer approached Topcoder to explore what could be done to improve the safety of their workers, we formulated an approach that relied on our large community to generate labeled data by taking existing video footage of industrial settings and creating detailed, minute-by-minute annotations of PPE usage.

A sizable, labeled dataset then allowed Topcoder to run a competition to generate computer vision algorithms capable of analyzing the number and location of workers present in an industrial setting as well as the kinds of PPE they are wearing or using. This solution gives oil and gas companies the means to continually monitor employees at job sites for PPE compliance, all the while preserving the workers' privacy and minimizing any interference or distraction from performing their jobs.

7 LAYERS OF DATA AND PROCESS PROTECTIONS

Andy LaMora
Global Director,
Crowd Analytics & AI

Universal internet access and the rise of the gig economy are delivering on the promise of as-needed, when-needed expert workforces. The benefits of these workforces are increasingly compelling; customers can expect their production capacity to flex beyond core teams with their real-time demand for it, and can access hard-to-hire skills instantly and as needed, instead of grappling with a job market that is increasingly difficult to access.

But with the power of a thousand minds on tap comes the risk of sharing your data and work with countless strangers. Each of these workers may see a slice of your data or strategic intentions. Concerns over IP sharing, IP contamination, disclosure, and privacy naturally follow. Our crowdsourcing platform was founded in 2001 and has dealt with these concerns every day since. Both the tools that we use and the methods we employ to control these risks change year by year as new tools emerge and ways of doing business change. We answer questions about these methods in every Q&A, and every deal cycle.

As our Global Director of Crowd Analytics & AI, I thought it would be helpful to share the latest basics on how Topcoder mitigates these concerns today with seven distinct layers of security.

LAYER #1: AGREEMENTS

It begins with agreements. When you're a customer of Wipro and Topcoder, we sign an agreement with you that sets the rules for what we can and can't disclose, as well as the process for disclosing it — exactly like any other prudent commercial transaction. These terms are typically handled in the MSA, and more stringent requirements can be layered on top when needed, SOW by SOW. For projects that require them, our contestants digitally sign NDAs as a condition of access to the challenge. There's sometimes a misconception that crowdsourcing is unique in this regard. In reality, customers experience a commercial relationship with us, complete with standard NDAs and contract terms.

LAYER #2: ATOMIZATION

Topcoder handles projects according to the skill types required — through a process called atomization. We take the project you'd like to build and break it down into bite-sized segments, which become separate challenges (e.g., app design, coding, etc.) that we run through our community. While this process was designed to allow us to control time and delivery, atomization also adds obvious protection. Think of it like this: members of our global crowd don't get to work on Voltron as a whole; they work on a single robot lion (or limb) at a time. Workers won't know there are other lion robots that assemble into Voltron unless you want them to. Atomization drastically and naturally reduces the number of people who see your entire project, which is already more protection than a traditional contractor engagement typically provides.

LAYER #3: PSEUDONYM

We don't disclose the identity of our customer to the Topcoder Community. We assign a pseudonym instead. Generally, it's the same pseudonym across all projects for any one customer, but we will also assign them project by project, or even component by component, as required for the project's security goals. So our members may not realize that two projects they're working on are even for the same customer.

LAYER #4: OBFUSCATION OF DATA

Obfuscation an important, very complex topic. Obfuscation is a best practice-driven scrubbing of personal identifiable information (PII) and other sensitive information in order to mask that data and reduce or eliminate the likelihood that a worker can correlate it with anything else in the field, or even who it's for. Obfuscation is always a partnership exercise, and either the data is treated before it's handed to us, or Wipro and Topcoder work with the customer to prepare it. We have adopted and developed several approaches for obfuscation. They range from simple scrambling of PII or key identifiers (e.g., product codes, warehouse IDs, etc.), to statistically rigorous replication of reference data to create a fabricated, but still relevant data set.

LAYER #5: METAPHORS

A metaphor transposes the domain. Metaphors have long played a role in gamification (see FoldIT and Play To Cure for examples), or abstracting the problem domain from the solution in order to find new approaches. They also help in protection. We'll apply metaphors when even the basic project domain or purpose shouldn't be exposed. To the extent that any position data is needed, Topcoder preserves relative but not exact spatial relationships while moving the scene to another continent or even planet, and might present the problem as a widget manufacturer instead. This way, we further distance the data and topic from its presentation to competitors on our platform. Together, Wipro and Topcoder then go on to unwind those metaphors when we return results to clients.

LAYER #6: DIRECT REVIEWS AND DIRECT TESTING

Our review process uses a two-pronged approach. One prong is direct, manual review performed by no fewer than two expert reviewers in our community — members who've proven to be not only technical masters, but also trustworthy on our platform. For critical code reviews, they inspect code line by line and complete lengthy scorecards, searching for best practices and security flaws. (Reviewers are unable to see the identity of the submitters.) A contestant must first get past those sentinels if they want a chance at victory. The second prong is technology. We also run the code across SAST when necessary, as well as IP detection platforms. Mike Morris, our CEO, wrote on this subject in relation to crowdsourcing as being more secure than traditional means of development.

LAYER #7: RING-FENCED CROWDS

There are times when the project or data simply cannot be shared in any form with the public crowd. Fortunately, this is quite rare. But in those cases, we are able to develop a subcrowd to work on projects. We will first qualify workers who are interested, available, and capable of working on the given solution. These workers are then asked to complete additional paperwork; past examples include data use agreements, network use agreements, even background checks. The worker pool may even include consultants from Wipro or the client's other trusted vendors. Only after this paperwork (and if necessary, location-specific) conditions are met are project details shared. When necessary, we can also set up virtual private clouds with I/O restrictions for their use. But it's worth noting that this is a last resort step; reducing the size of the addressable workforce always has an impact.

UP NEXT AT TOPCODER: DIFFERENTIAL PRIVACY

If concerns about the risks of data sharing are on the rise, so fortunately are methods for dealing with it. One promising technique for obfuscation on the ascent is called Differential Privacy ("DP"). DP seeks to replicate important data in a manner that both breaks the ability to triangulate data back to reality while also preserving key relationships.

To illustrate the point: imagine being able to replicate a data set of disease patients in an entire state in a manner where hundreds of data scientists can perform tests to seek precursor signals, without the risk that some bad actor can figure out patient or provider identities. Through our innovation contract with NASA and in partnership with NIST, Topcoder will be hosting a Differential Privacy challenge this November and are exploring methods to refine these techniques into our standard practice. If you're a data scientist and would like a chance to contribute to the solution, click here to stay in the loop and join the contest!

TOPCODER'S COMMITMENT TO PRIVACY AND SECURITY

As any security professional will tell you, privacy and security protections is a dynamic field that requires constant diligence. Rest assured, our methods of protecting our clients, our members, and ourselves, are always evolving.

Security is an intrinsic component of Topcoder's offering. It exists in all aspects of the business — from a customer's first interaction with the platform, to members registering and competing, to ultimately delivering solutions. Our platform enables collaboration yet preserves privacy, allowing for experimentation with limited risk.

CROWD-POWERED CONSULTING: ANALYTICS COE, ENERGY SECTOR

One of the world's largest independent upstream oil and natural gas companies needed to meet a growing demand from internal customers for data science, design, and solution development without adding in-house resources. The answer: a crowd-powered Analytics CoE that combines Wipro and Topcoder resources and expertise.

CROWDSOURCING ADVANTAGES

- Fast start — no waiting for hires
- No turnover — the crowd never quits
- On-demand access to hard-to-find skills (e.g., data science, development, and UI/UX design) with repeatable results

THE RESULTS

Since December 2017, the Analytics Consulting team has run over 318 competitions on 39 ideas — with an average opportunity ROI of 20x.

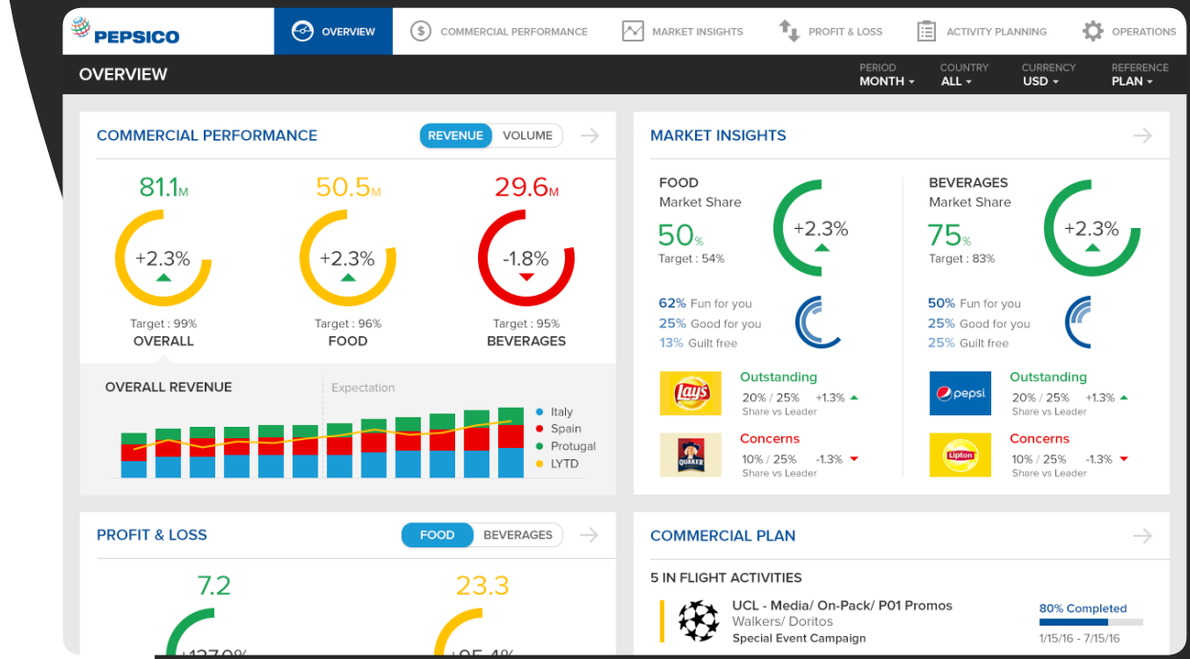
ADDITIONAL TOPCODER CASE STUDIES

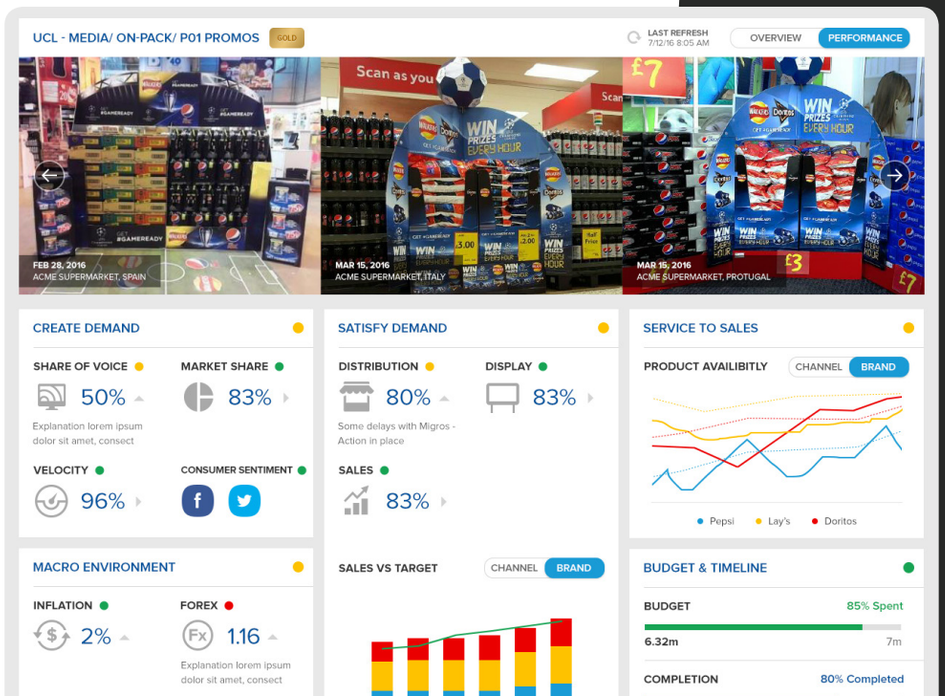
DATA VISUALIZATION - EXECUTIVE DASHBOARD

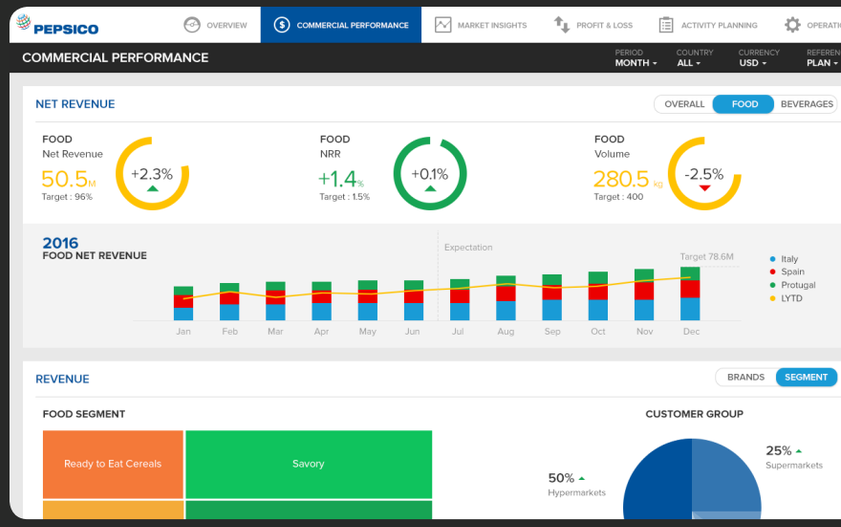
PEPSICO 3D MARKET LEADERSHIP COCKPIT

Like many large organizations, PepsiCo measures the state of its business across several KPIs to give executives important up-to-date information and allow them to make the best-informed decisions possible. Since business performance data from a range of markets and product categories needed to be consolidated in one place, a new approach to data visualization was requested from Topoder.

PepsiCo executives needed to be able to quickly scan the information, view key drivers of the business, and take action based on what the information was telling them. Because of this any dashboard for displaying the variety of indicators had to be consistent, clear, and concise. When PepsiCo turned to Topcoder, they received all of this and more.







USABILITY TESTING

AMERICAN PUBLISHING COMPANY

07
DAYS

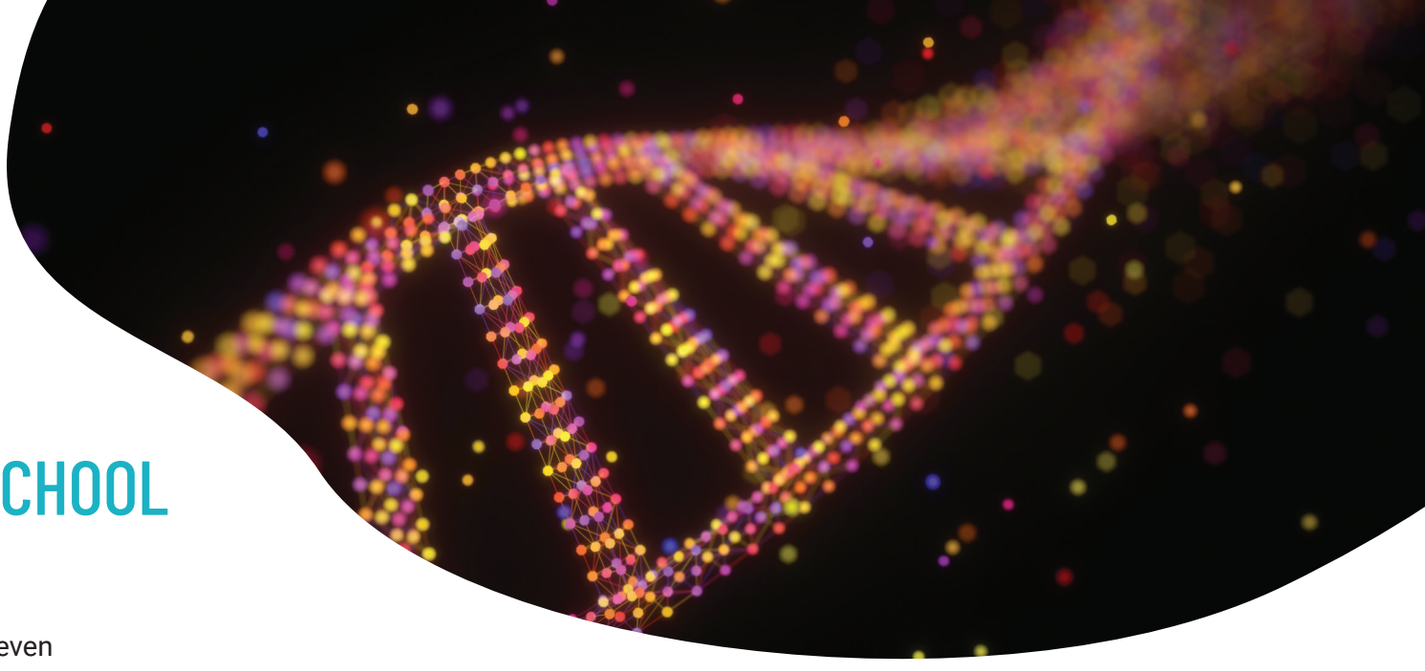
34
QA ENGINEERS

150
ISSUES IDENTIFIED

81
ISSUES ACCEPTED

Media, news, and publishing is an ever-changing industry, particularly when digital… especially when the greater user experience and critical workflows are in question. A leading American publishing company based in Sacramento, California, needed to test their digital subscription process across multiple browsers on various devices and verify usability of workflows surrounding subscriptions, e-editions, and rewards for four news publications.

They came to Topcoder to run a crowdsourced QA competition. Topcoder instantly onboarded a global QA community for testing the applications in real-world conditions — on real devices in real networks, using real user accounts to sign up and subscribe — to uncover defects and insights into end-user behavior. In just seven days, 34 expert testers found 81 critical issues and suggested 28 UI/UX enhancements to help move the publishing company forward into new target markets.



DNA SEQUENCING ALGORITHM

HARVARD MEDICAL SCHOOL

Studying the human genome is the world's largest project, but the work can be slow—even with computers.

Renowned for its innovation in medical research and genomics, Harvard Medical School wanted to speed the process of standard DNA sequencing, which is essential for making precise, high-throughput readouts of the immune system. A full-time employee had worked for a year to optimize an algorithm that calculates the distance between DNA strings, but they wanted to see if more data scientists working on the problem could deliver even better results.

With an Analytics Starter Pack from Topcoder, Harvard received on-demand access to more than 120 data scientists who worked to optimize the algorithm for two weeks. Competitors submitted more than 650 possible solutions using 89 unique approaches to the problem. The final result delivered by Topcoder was extraordinary, increasing the speed of their algorithm from 260.4 minutes to just 16 seconds — 976 times faster. This extreme value outcome not only enabled Harvard to accelerate their study of genetics as a unified way to extract organizing principles, but it also shifted the way they approach genetics research.

"Topcoder surpassed expectations: a two-week competition led to code that was just as good but almost three orders of magnitude faster... Hard to imagine beating that."

Ramy Arnaout MD
DPhil Associate Director Clinical Microbiology, Department of Pathology, BIDMC

122	654	89	2 WKS
DATA SCIENTISTS	CHALLENGE SUBMISSIONS	UNIQUE APPROACHES	PROJECT DURECTION

PIPELINE THREAT DETECTION

With millions of miles of pipeline across the country, protecting critical fuel supplies is more difficult than ever.

The U.S. government wanted to develop an algorithmic solution to detect and classify objects within a certain range of energy pipelines—and make monitoring more efficient and secure. Thanks to advances in satellite capabilities and drone technologies, the necessary aerial imagery was abundant. But automatically differentiating between downed tree limbs and enemy vehicles in those images required a new approach to the problem.

With an Analytics Starter Pack from Topcoder, the government client received on-demand access to more than 100 data scientists who worked to develop algorithmic solutions that detect and evaluate potential risks. A crowd-powered delivery team managed all logistics during the three-week project, including testing the 500+ possible solutions submitted.

The winning solutions rapidly process tens of thousands of images and tag objects with an appropriate threat level. Today, the final algorithmic solution delivered by Topcoder is also being used to drive other government research in planetary satellite classification, Mars reconnaissance, federal disaster response and recovery, and beyond.

topcoder

IF YOU'RE READY, WE'RE READY!

I'm Minal Kanungo, your point person for all things Energy, Natural Resources, and Utilities at Topcoder. Questions? Ideas? Projects? Reach out and let's get started.

Minal Kanungo
Director North America ENUC
minal.kanungo@topcoder.com

topcoder

THE SIMPLEST WAY TO ACCESS AND EXECUTE WITH INCREDIBLE DIGITAL TALENT

topcoder.com