

Fighting Intolerance with Data:

Technion's First-Ever 'Datathon' to Combat Hate Speech

An Enduring Commitment to Society

Beneath every scientific venture born from the Technion, exists an ongoing quest to improve the world through technology. The mission materializes in many forms. Sometimes, it manifests as transformative breakthroughs in the healthcare industry or through groundbreaking discoveries that unlock the secrets surrounding humanity, our planet and beyond. Other times, the knowledge gained through scientific exploration is applied to the social realm when technology is developed to promote a more inclusive and tolerant society, free from prejudice and bigotry. Recently, with the massive availability of data and the tools to deal with them, termed broadly data science, it is the data that transform the world as serves as an agent of change.

No matter the destination, the path to scientific achievement is lined with an innate yearning to envision and create a better world through technology. The Technion remains true to its commitment to society through producing not just the brightest minds and ideas in science and engineering, but also by inspiring them to become the socially conscious innovators of the future. To accomplish this, the Technion seeks new directions in curriculum, programs and events with a high priority placed on finding solutions to meet the neglected.

The Challenge of Content Moderation

The prevalence of hateful, discriminatory, and intolerant content on social media is a growing concern in modern society. Antisemitism, anti-LGBTQ, and other forms of hate speech are divisive and have the potential to incite violence. Incendiary content creates enemies, silences debates, and marginalizes individuals, groups, and minority populations. Left unhindered, toxicity on social media poses a threat to safety and security, peace, and tolerance.

As platforms expand their reach and increase their volume of traffic, traditionally manual forms of content moderation become impracticable. Human screening of massive online content is unreliable and often untimely: required to deal with torrents of content throughout the workday, content moderators employed by platforms often misread text or misinterpret context; similarly, content flagging by a site's visitors and users is necessarily unsystematic – leading to inconsistent applications of its policies and guidelines. Either way, cognitive biases are liable to taint the process of manual screening. Moreover, in many instances, after-the-fact deletion of already-posted materials does little to minimize their damage.

Automated machine learning-based moderation tools can, however, provide for a more dispassionate, thorough, and timely removal of volatile online content. Using data science techniques and big data analytics methodology, automated moderation solutions could efficiently and immediately parse contexts, identify, and isolate toxic material, and eliminate it from the platform in real-time – irrespective of the volume of online traffic.

With its 55 faculty members, Technion's Faculty of Industrial Engineering and Management trains students in a wide spectrum of activities – from applied engineering to mathematical modeling, to

operations research and statistics, and more. A recognized global leader in data science education and research, this extraordinary faculty is now intent on undertaking, and pushing the envelope of, automated content moderation.

Meeting of the Minds: Promoting Tolerance on Social Media

To join the fight against hate speech, the Technion proposes to host its first-ever “Datathon” – a massive brainstorming event where skilled participants are tasked with formulating a solution to a problem in a specified amount of time. The proposed Datathon will be designed with the intent to arm society with the technology required to effectively identify and eliminate hate speech from social media platforms.

To maximize the positive social impact of this Datathon, all participating teams agree to openly publish and share their models with organizers and any other organization interested in using it.

Approximately 150 students from all Technion faculties will take part in the 3-day event. To address the legal aspects of content filtering, students from the University of Haifa’s Faculty of Law will also be invited to join the Datathon.

The challenge will focus on topics such as disinformation, toxicity and terrorism. The students will be challenged with a massive social media dataset on which they will be required to implement data skills learned in the classroom and apply them to this real-world issue. The end result will consist of a complete end-to-end method (algorithms combined with a clear methodology for extracting data, comprehending the context, analyzing text and presenting outcome in a meaningful way) of identifying volatile content, offensive language, propaganda that aims at radicalizing youth, as well as misleading or inaccurate information spread deliberately to incite readers. The students will be presented with two specific challenges:

- **Challenge 1:** Build a model that can automatically classify text messages as hate speech. We hope for methodologies that not only focus on the word content of the message, but also consider the contextual factors and sentiment analysis related to hate speech.

- **Challenge 2:** Propose an action-oriented solution to the hate speech detected. This may include a full-pipeline architecture to handle hate speech, identifying and eliminating sources of toxic content, a methodology for balancing freedom of speech and hate speech (including methods for turning hate speeches to a version that satisfy free speech norms), and education plans for a world where hate speech will not be something to fight, but rather an act that is completely outside the borders of societal norms.

Groups will be evaluated based on the accuracy of their solution, as well as its innovation and its resistance to adversarial attacks. There will be two prizes. First place winners will receive 5,000 NIS, while second place winners will receive 2,500 NIS.

Contribution Requested

A one-time contribution of \$70,000 CAD will fully fund the Datathon – which will be scheduled to coincide with a future visit to Israel by the donors, at any time within the 2025 academic year (other than the examination period).

Datathon Budget 2025 Academic Year	
Item	Total (CAD)
Human Resources	

Mentors	4,200
Consultancy	1,800
Security and housekeeping services	6,300
Total Human Resources	12,300
Hosting Costs	
Meals, refreshments, snacks	23,700
Event site branding, furniture and set-up	5,000
Cloud computing	5,000
Prizes	3,200
Social/Extra-curricular activities	2,600
Total Hosting Costs	39,500
General infrastructure costs	8,100
Technion Canada Administrative Expenses	9,800
Total Project Cost	69,700

Donor Recognition

This generous gift will also be recognized in the following manner:

1. Within two months after the event, donors will be sent a report summarizing the event, its outcomes, and their expected impact.
2. The donation will be listed in the "President's Report," the official annual report of the Technion.
3. During their visit to Israel at the time of the Datathon, donors will have the opportunity to meet with participants, mentors, and organizers of event.