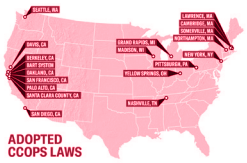
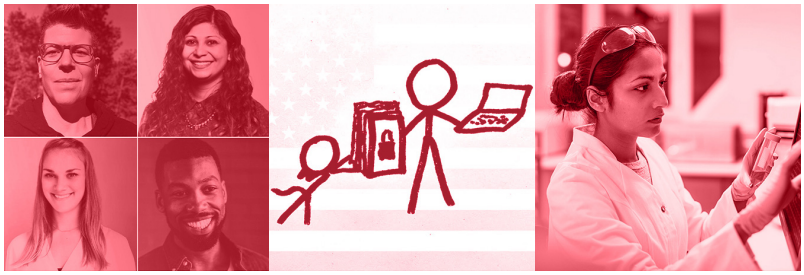


Powered by
Project Galileo



Tech Spotlight
Casebook

Spring 2021

HARVARD Kennedy School
BELFER CENTER
for Science and International Affairs
Technology & Public Purpose Project

Tech Spotlight
Casebook

Spring 2021

HARVARD Kennedy School
BELFER CENTER
for Science and International Affairs
Technology & Public Purpose Project

HARVARD Kennedy School
BELFER CENTER
for Science and International Affairs
Technology & Public Purpose Project

Technology and Public Purpose Project
Belfer Center for Science and
International Affairs
Harvard Kennedy School
79 JFK Street
Cambridge, MA 02138

www.belfercenter.org/TAPP

Design and layout by ccm.design

Faculty Director
Ash Carter

Technology and Public Purpose (TAPP) Project Director
Laura Manley

Tech Spotlight Project Lead
Ariel Higuchi

TAPP Core Team
Karen Ejiofor
Amritha Jayanti
Henry Kaempf
Mike Miesen

Student Research Assistants
Jenny Blessing
Kenny Chen
Gerard Connolly
Raj Gambhir
Jenna Hussein
Mark Karugarama
Jesse Lin
Nicholas Simmons
Lindsay Temes
Conor Truax

Table of Contents

Acknowledgements

The TAPP team would like to acknowledge and thank the Tech Spotlight Selection Committee, including Bill Anderson, Afua Bruce, Kade Crockford, Nancy Gibbs, John Holdren, Sheila Jasanoff, Chris Lynch, Travis McCready, DJ Patil, Kathy Pham, Katie Rae, Justin Sanchez, Nick Sinai, Steve Strassman, Nick Thompson, and Paul Waters for their leadership and decision making throughout this process.

The TAPP team would also like to thank Julie Balise, Josh Burek, Bennett Craig, Andrew Facini, Sharon Wilke, and Simone Worsdale for their content development and promotion efforts. The production and design of this report would not be possible without Claudio Mendonca and his team at CCM.Design.

The TAPP team is grateful to our diligent Student Researchers including Jenny Blessing, Kenny Chen, Gerard Connolly, Raj Gambhir, Jenna Hussein, Mark Karugarama, Jesse Lin, Nicholas Simmons, Lindsay Temes, and Conor Truax.

Finally the TAPP Team would like to recognize Bogdan Belei and Karen Ejiofor for their contributions to the original Tech Spotlight concept and bringing it to life.

About the Tech Spotlight

2020 tested our strength, our resilience, and our ability to break down barriers and innovate. Despite a year of much grief, anger, and loss, people around the globe have worked harder than ever before to improve the lives of others. Millions of people in government, business, civil society, and academia have participated in groundbreaking efforts to make our world more inclusive, safer, and fairer by leveraging tech for good. When the development and deployment of technology is grounded in the greater good of humanity, these efforts—which are often ignored in news headlines—should be acknowledged and encouraged.

The Tech Spotlight recognizes projects and initiatives that demonstrate a commitment to public purpose in the areas of digital, biotech, and future of work. Through a nomination process, the TAPP team evaluated entries based on their proven ability to minimize technological harms and protect public purpose values including, but not limited to:

Privacy

Safety and Security

Transparency and Accountability

Inclusion

Eligibility Requirements

- Nominations are accepted for projects and initiatives
- Must demonstrate a commitment to reduce societal harms and protect public purpose values such as privacy, safety and security, transparency, accountability and inclusion
- Must be related to one of the following priority areas: digital technology, biotechnology, or future of work.

SELECTION PROCESS

Nominations for the Tech Spotlight were open in Fall 2020. All qualifying submissions were researched and verified by the TAPP Team, including desk research, interviews, and a conflict of interest assessment. Each submission was then scored according to the following criteria:

Alignment

Minimizes technological harms and protects public purpose values, including privacy, safety and security, transparency and accountability, and inclusion.

Impact

Clear demonstration of impact, scalability, and geographic scope through specific, measurable examples.

Innovativeness

Novelty of strategy or approach and responsiveness to current issues.

Top submissions advanced to the Selection Committee to vote on the 2021 Tech Spotlight finalists and top recipients ■

***Please note:** *Selection Committee members with any affiliation to a nomination under consideration recused themselves from voting on said nomination.*

Tech Spotlight Recipients

2020 Census Disclosure Avoidance System
U.S. Census Bureau

Community Control Over Police Surveillance
ACLU

Project Galileo
Cloudflare

Runners-Up

***Privacy Not Included**
Mozilla Foundation

COVID Symptom Study App
COVID Symptom Study

COVID-19 Molecular Explorer
IBM Research

Dataset Nutrition Labels
The Data Nutrition Project

Garbage In, Garbage Out
Center on Privacy & Technology at Georgetown Law

Mind the Gap
Black and Brown Skin

Project Amelia
Probable Models

Project Lighthouse
Airbnb

Racial Disparities in Automated Speech Recognition
Stanford Computational Policy Lab

Safe House and Shelter Training Program
Operation Safe Escape

SmartNoise
OpenDP and Microsoft

Student Privacy Project
EPIC

Terms of Service Ratings
ToS;DR Association

Upsolve App
Upsolve

2020 Census Disclosure Avoidance System

U.S. Census Bureau

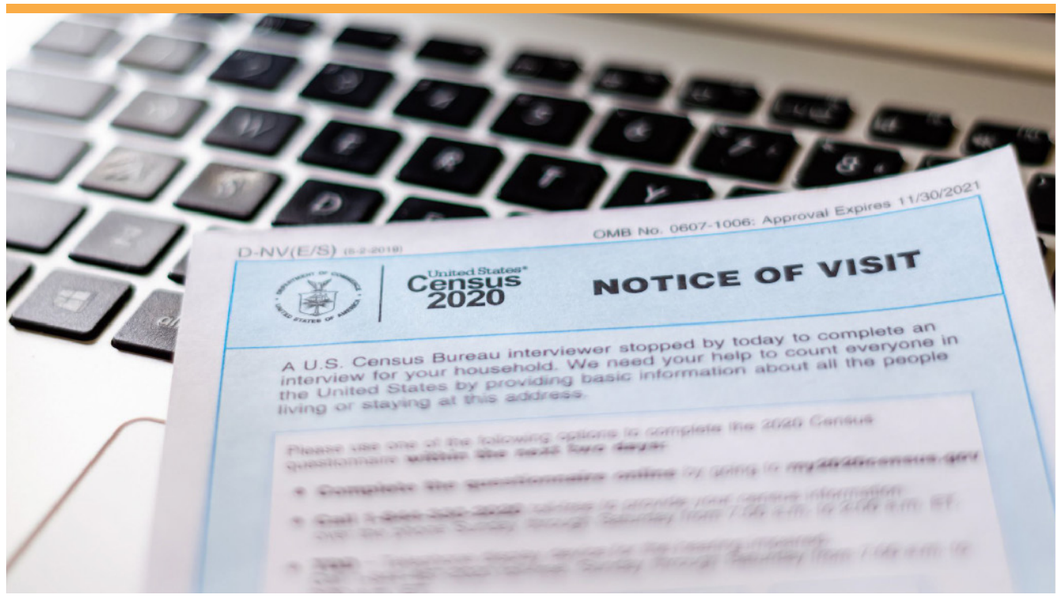
The 2020 Decennial Census used differential privacy to preserve the anonymity of U.S. residents.

Suitland, MD

OVERVIEW

During the 2020 Census, the U.S. Census Bureau used differential privacy in a powerful new 2020 Census Disclosure Avoidance System (DAS). The purpose of the 2020 Census DAS was to design a system that could withstand modern re-identification threats and improve privacy for U.S. Census participants. This initiative directly mitigates the growing threats to privacy that are caused by the increase in computing power and the proliferation of personal data online. It strikes a balance between preserving the privacy of census respondents while maintaining the availability and utility of published census data.

The Census Bureau previously took steps to anonymize data. This included injecting "noise" into the data to halt re-identification attacks. Historically, this was an effective way to maintain anonymity, but access to so many open source databases now allows nefarious actors to overcome previous data anonymization methods.



The 2020 Census DAS was implemented to address security concerns posed by the combined threat of modern computing capabilities and our data-rich environment.

THE CHALLENGE

The U.S. Constitution mandates the U.S. Census to take place every ten years. The census impacts election redistricting, federal and state funding allocation, research, and other government functions. When conducting the census, there is an expectation, and legal obligation, to protect the confidentiality of census respondents' data.

Modern computers and today's data-rich environment have rendered the Census Bureau's traditional confidentiality protection methods almost obsolete. Over the last few years, Census Bureau researchers simulated a re-identification attack on the published 2010 Census data. They were able to reconstruct individual responses for the entire population – without names or other identifiers. They were then able to match those reconstructed responses with publicly available commercial data that included names. They found that about 52 million people, or 17 percent of the 2010 Census population, were correctly re-identified. That was a best-case scenario. Using higher quality data, the number of people correctly re-identified would rise to about 179 million people, about 58% of the population.

The 2020 Census DAS tackles the complex problem of balancing census data anonymity against data accuracy.

ABOUT THE INTERVENTION

The 2020 Census DAS tackles the complex problem of balancing census data anonymity with data accuracy. The Census information must be usable for government actions and research while precluding bad actors from using the data to identify census participants.

The Census Bureau developed a new system that uses cryptographic principles to obstruct attackers from identifying the individuals behind published 2020 Census statistics. The system improves upon "legacy" methods of anonymizing the data. Starting with the 1990 Census, the Census Bureau began infusing "statistical noise" (controlled amounts of error) into data deemed most at risk for exposure. This was a relatively blunt technique designed to strike a balance between preserving overall data accuracy and reducing the risk of re-identification. As with all noise infusion techniques, these "legacy" methods produced data distortions. The nature and extent of those distortions are protected information to preserve the confidentiality of the underlying data. The Census Bureau is going to great lengths to ensure that the data is fit-for-use while maintaining confidentiality.

IMPACT & FUTURE PLANS

This system signals a concerted effort by the Census Bureau to protect the privacy of all 330 million people counted in the census. The first set of 2020 Census data products impacted by the new Disclosure Avoidance System will be the redistricting data released in August–September 2021. The system will be adapted to produce more detailed data products expected in 2022 and beyond.

Through the 2020 Census DAS, the Census Bureau is keeping pace with the challenges and opportunities posed by today's technology to protect the privacy of the population while serving the nation's critical information needs ■

Community Control Over Police Surveillance (CCOPS)

ACLU

CCOPS is a legislative advocacy and organizing initiative to increase transparency, oversight, and community influence over if and how local police are allowed to use surveillance technologies.

New York, NY

OVERVIEW

Technology has made it easier for police departments to surveil the communities they are charged with protecting. Some of the surveillance technologies themselves, such as facial recognition, are racially biased. Additionally, virtually every surveillance technology has been found – like policing itself – to be disproportionately deployed against communities of color.

The ACLU's Community Control Over Police Surveillance (CCOPS) is an advocacy and community organizing initiative developed in direct response to the harms of surveillance technologies and their growing use by police. CCOPS ordinances protect the privacy of communities and increase transparency and inclusion in the development of police surveillance practices. CCOPS laws help ensure the appropriate and responsible use of technology deployed by police. Furthermore, CCOPS laws allow communities to have a real and meaningful opportunity to reject the use of unwelcome surveillance technologies before they are acquired and deployed.

CCOPS legislative efforts have been at the forefront of responding to the dangers posed at the intersection of policing, surveillance technologies, and racial bias within American law enforcement.

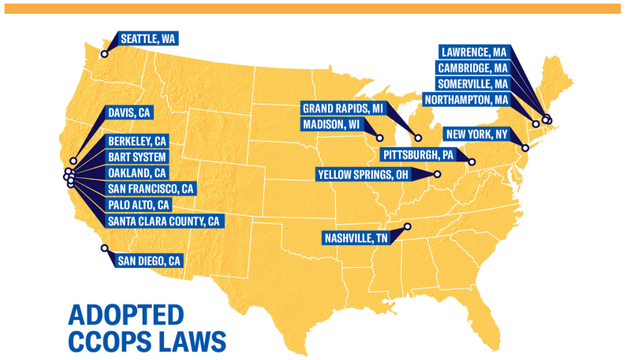
THE CHALLENGE

Police departments around the United States can acquire and deploy a vast array of surveillance technologies with minimal to no oversight. Facial recognition software, cell phone tracking devices, automatic license plate readers, and predictive policing software are among the technologies commonly used. But surveillance technologies do not pose an equal threat to everyone. Law enforcement officials use surveillance technologies in ways that magnify racial bias and stereotypes in the course of their work. Furthermore, the technology itself can be racially biased. Facial recognition is one increasingly popular tool among law enforcement that is significantly less accurate in identifying darker-skinned faces.[1] Predictive policing software, which relies on historically biased policing data to inform its predictions, also raises concerns over racial and socioeconomic bias.

ABOUT THE INTERVENTION

CCOPS was launched as a nationwide initiative in 2016. CCOPS increases transparency and police oversight by encouraging the adoption of municipal legislation. The goal is to empower local communities to decide if and how their local law enforcement agency should use surveillance technologies.

CCOPS ordinances have been adopted by 19 cities around the United States, where they protect and empower over 16.2 million people.



The national ACLU organization developed a model CCOPS bill. ACLU state chapters work with local communities and organizations to tailor the model bill to the needs of each particular community. As a result, CCOPS ordinances vary from city to city, but the shared outcome is (1) increased transparency, oversight, and regulation of policing surveillance practices and increased community engagement and (2) increase community engagement and influence over decisions about if and how police may use surveillance technologies. The CCOPS ordinance in Cambridge, MA, for example, requires city council approval before Cambridge police can purchase or use various surveillance tools, as well as regular reporting on technology usage. The CCOPS law in San Francisco included the first municipal ban on the use of facial recognition in the United States. In 2020, the New York City Council enacted the Public Oversight of Surveillance Technology (POST) Act, another CCOPS ordinance that requires the NYPD to publicly disclose information about surveillance technologies in use.

While CCOPS's focus is on surveillance technology, the biggest concern of the initiative is racial justice. Police surveillance has disproportionately targeted communities of color. CCOPS addresses this issue by advocating for public disclosure of where and how local police would like to use surveillance tools. It also empowers the public to set rules for the surveillance technology's use or to reject its use altogether.

IMPACT & FUTURE PLANS

CCOPS transparency and community-focused legislative efforts have been at the forefront of responding to the dangers posed at the intersection of policing, surveillance technologies, and racial bias within American law enforcement. CCOPS ordinances have been adopted by 19 cities around the United States, including San Francisco and New York City, where they protect and empower over 16.2 million people. Community engagement is a cornerstone of CCOPS; the local communities have become substantially more engaged in the democratic process as part of the initiative. Going forward, the ACLU will continue to engage in legislative and community advocacy efforts in additional cities around the country, with the hope that these local efforts trickle up to the state and federal levels of government ■

1 Joy Buolamwini and Timnit Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," in ***Conference on Fairness, Accountability and Transparency*** (Conference on Fairness, Accountability and Transparency, PMLR, 2018), 77–91, http://proceedings.mlr.press/v81/buolamwini18a.html.

Project Galileo

Cloudflare

Project Galileo provides targeted cybersecurity services and protection to at-risk public interest groups and nonprofits.

San Francisco, CA

OVERVIEW

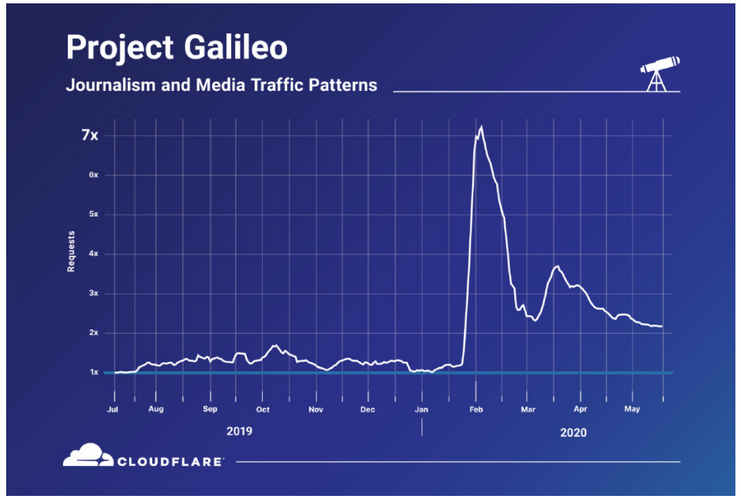
Cyberattacks attempting to take websites offline have become a 21st-century method of censorship. Nonprofits and public interest groups are uniquely vulnerable to such attacks but typically lack the resources and expertise to secure themselves. In 2014, Cloudflare founded Project Galileo to provide free cybersecurity protection to the most vulnerable groups on the Internet, including organizations that work in the arts, human rights, civil society, journalism, or democracy. Through ensuring that these groups will be able to stay online, Project Galileo facilitates a wide variety of critical work in the public interest.

Powered by
Project Galileo

THE CHALLENGE

Cyberattacks are becoming more common each year. As much of the world has been forced online due to the COVID-19 pandemic, attacks on Internet infrastructure are more concerning than ever. Denial of service attacks, in particular, are simple to perpetrate and enable attackers to take websites offline so that legitimate users cannot access them. Traffic from legitimate users can also cause a site to crash in cases where there is an abrupt increase in site visitors. For example, a journalism site publishing a story on government corruption can attract a sudden rush of millions of visitors to the website. Furthermore, the shift to remote work during the COVID-19 pandemic brought unexpected challenges. Unfortunately, many organizations struggled to enable remote access while maintaining a strong security posture, including encrypting sensitive data and ensuring secure remote access.

While any organization can be the victim of a cyberattack, public interest groups are among the most common targets but usually lack the resources to adequately prevent and respond to them. Advocacy groups promoting LGBT rights in the Middle East, political corruption trackers in Sri Lanka, independent Ukrainian journalists reporting on Russian military intervention, and many others are in vulnerable positions where they face powerful opposition to their work. Today, most governments have sufficient cyber offensive capabilities to launch attacks to take inconvenient information offline, especially in countries experienced in cyberwarfare like Russia and China.



During the pandemic, groups organizing around racial justice and COVID-19 were able to stay online because they were protected by Project Galileo.

ABOUT THE INTERVENTION

Cloudflare started Project Galileo in June 2014, after learning that an independent newspaper in Ukraine had come under a sophisticated denial of service attack while covering Russian military intervention in Crimea. With that, Cloudflare's CEO was inspired to start Project Galileo.

Project Galileo provides Cloudflare's upgraded security services for free to nonprofits and organizations acting in the public interest. Cloudflare collaborates with 41 civil society partners – including the ACLU, Amnesty International, and the Freedom of the Press Foundation – to identify organizations in need of Cloudflare's protection. Applications to join Project Galileo are reviewed by the 41 partners. If at least one of the 41 partners supports an organization's application, then that organization is automatically offered protection under Project Galileo.

As of April 2021, Galileo protects 1,480 domains in 100 different countries worldwide, ranging in work from human rights, environmental groups, artists, health organizations, education, and more. Examples of groups protected include Women's March Global, Vote America, The Trevor Project, and The Water Project.

IMPACT & FUTURE PLANS

Between January and June 2020, Galileo mitigated 2.4 billion cyberattacks on Galileo participants, for an average of 17 million each day. 60% of the organizations they protect experience cyberattacks every single day. In late May, Cloudflare saw a significant uptick in cyberattacks on racial justice groups and organizations fighting racism in the wake of the death of George Floyd. Between April to May, the number of cyberattacks on such groups increased by over 1,120 times to as many as 20 thousand requests per second for a single site. These organizations were able to stay online because Project Galileo protected them.

During the COVID-19 pandemic, Project Galileo protected many websites that promoted relief efforts, including organizations that developed COVID symptom trackers, platforms for up-to-date vaccine information, mental health hotlines for those struggling in lockdown, and more. Project Galileo also started to offer services for securing remote work, including access control and products that automatically block security threats like malware and phishing attacks. In doing so, Project Galileo provided a safer, more secure environment in the rapid transition to remote work.

As world events continuously shape how public-interest organizations operate, Project Galileo works with their 41 partners to adapt so they can best support those they are protecting. Project Galileo helps the organizations stay online, secure their internal teams, and focus on their mission of helping the greater good ■

*Privacy Not Included

Mozilla Foundation

**Privacy Not Included is a guide that helps consumers learn about the privacy and security features of tech gadgets.*

San Francisco, CA

OVERVIEW

*Privacy Not Included is a buyer's guide that informs consumers about the privacy and security of products connected to the internet. The Mozilla Foundation launched the guide in 2017 and has since reviewed over 180 products. Mozilla Foundation's technical experts evaluate each product against a set of criteria called the Minimum Security Standard. The standard focuses on encryption, automatic security updates, strong password requirements, managing system vulnerabilities, and the accessibility of privacy policies.

The guide also allows users to share their opinions on products' safety and security standards through a simple survey called the "Creep-O-Meeter."

Inspired by *Privacy Not Included, in December 2020, Mozilla hosted an event to explore the state of privacy and security in consumer tech gadgets.

THE CHALLENGE

Consumer privacy and security are under siege. Many connected devices and apps — from doorbells to watches — collect our data, then sell it, exploit it, or simply do not protect it. Meanwhile, consumers have few options to push back: privacy regulations are scarce, policies are indecipherable, and privacy-centric alternatives are not always well known.

Another part of the challenge is capturing, and building upon, user opinion. *Privacy Not Included believes it is critical that companies, and other consumers, see which products people think are safe and which products people feel are too invasive.

ABOUT THE INTERVENTION

*Each product featured in the *Privacy Not Included guide is assessed against the Minimum Security Standard. This standard was designed by Mozilla, Consumers International, and Internet Society in 2018 and focuses on encryption, automatic security updates, strong password requirements, managing system vulnerabilities, and the accessibility of privacy policies.

Additionally, products are evaluated based on how collected data is used, the ability of a user to control collected data, and a company's known track record on protecting user data over the past two years. This information, in combination with the Minimum Security Standard evaluation, determines whether a product will

**Privacy Not Included publishes an annual holiday ranking of the creepiest and safest connected devices.*

be tagged with a *Privacy Not Included warning label. The guide also reviews questions such as a product's use of AI, what data can be collected, and how creepy people think a product is. The guide also tailors product reviews to specific audiences or moments in time. For example, in 2020, *Privacy Not Included published their first review of video call apps to help users understand which apps were better than others at connecting them to loved ones while also protecting their privacy.

Ultimately, *Privacy Not Included is focused on informing consumers. As Ashley Boyd, vice president for advocacy and engagement at the Mozilla Foundation, stated in an interview with NBC News, "We know there's a lot of money being made by collecting and packaging our data. Our position is: Let consumers opt-in to that kind of data collection rather than opt-out. Our concern lies in the lack of transparency or even basic information about the data that's being collected."[1]

IMPACT & FUTURE PLANS

*Privacy Not Included has been translated into four languages (English, French, Spanish, and German) and written about in numerous outlets, including NPR, WIRED, the New York Times, and USA Today. The guide has helped hold companies accountable for their privacy and security policies and, through features like the Creep-O-Meter, has shown companies that consumers care about these issues. For example, after *Privacy Not Included published its edition on video call services, companies like Discord changed their policies to better protect consumers.

In 2020, Mozilla won a Webby Award for *Privacy Not Included in the "People's Voice Award for Activism" category, stating "Privacy is power. Demand it."

For the future, the Mozilla Foundation is constantly evaluating how to evolve the guide to better help consumers. For example, the *Privacy Not Included team plans to add privacy user manuals into the guide to help people know which settings on an app or device will help consumers most protect their privacy. The Mozilla Foundation will continue to publish and update *Privacy Not Included year-round ■

1 Herb Weisbaum, "Are the Smart Devices in Your Home Spying on You?," NBC News, accessed April 23, 2021, https://www.nbcnews.com/better/lifestyle/downside-connected-tech-are-smart-devices-your-home-spying-you-ncna1101906.

COVID Symptom Study App

COVID Symptom Study

The COVID Symptom Study App is a 1-minute daily mobile survey that allows researchers to gain critical information necessary to combat COVID-19 and the pandemic's secondary effects.

Boston, MA and London, UK

OVERVIEW

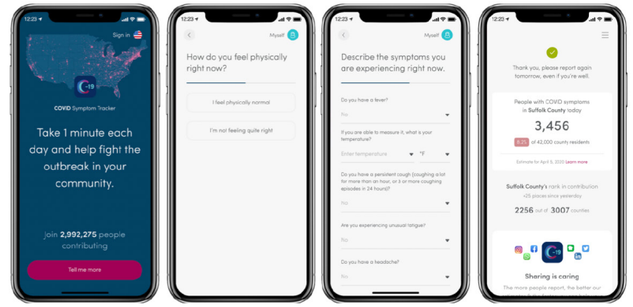
The COVID Symptom Study App was created by doctors and scientists at Massachusetts General Hospital, the Harvard T.H. Chan School of Public Health, King's College London, and Stanford University School of Medicine, co-developed with ZOE – a health science company. COVID-19 is still a relatively new disease, which makes new data collection especially critical in responding to the ongoing global pandemic. The data collected by the app helps scientists better understand COVID-19 symptoms and spread. It also helps

NIH Director's Blog

Predicting 'Long COVID Syndrome' with Help of a Smartphone App

Posted on March 23rd, 2021 by Dr. Francis Collins

*"As devastating as this pandemic has been, it's truly inspiring to see the many **innovative ways** in which researchers around the world have enlisted the help of **everyday citizens** to beat COVID-19."*



The Covid Symptom Study was recently featured in the NIH Director's Blog for its impact on our understanding of COVID.

identify COVID-19 risk groups and high-risk areas across the United States, the United Kingdom, and Sweden. As the pandemic evolves, the research scope of the COVID Symptom Study App has expanded to include how vaccines will influence the course of the pandemic and the impact of COVID-19 on diet and mental health.

THE CHALLENGE

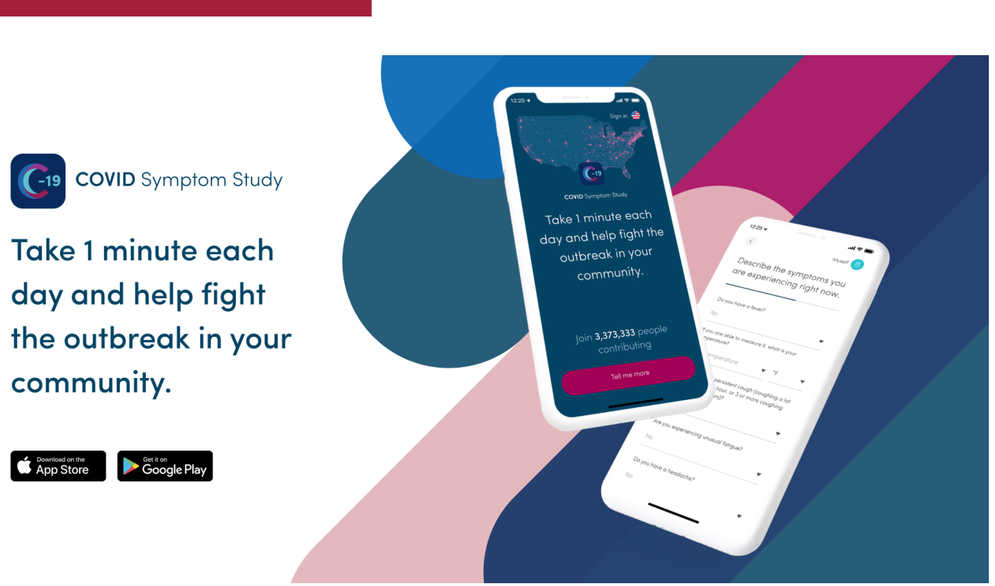
Public health officials have long struggled to bridge the lag between the spread of disease and the data needed in real-time to combat the disease. During the COVID-19 pandemic, in the absence of widespread testing, health officials worldwide struggled to track the spread of the coronavirus pandemic in real-time and assess the impact of mitigation efforts, such as lockdown measures and vaccines. Researchers also lacked data on risk factors, symptoms, and COVID-19's effects on other health conditions. This data is critical to combating the disease.

ABOUT THE INTERVENTION

The Covid Symptom Study App is a survey that quickly collects critical COVID-related data and extracts relevant insights. Through the app, the mobile survey has engaged over 4.6 million users to date, including many communities disproportionately impacted by COVID-19. The data that the mobile app collects is highly adaptable and changes depending on the needs of public health researchers.

The app has helped identify critical insights about COVID-19. These insights include the patterns and trends of COVID-19 spread and symptoms including loss of smell and taste. The research revealed that healthcare workers of color are five times more likely to contract COVID-19 than their non-Hispanic white counterparts. It has also helped explain the influence of social determinants of health on the disproportionate impact of COVID-19 on communities of color. Additionally, the study recently uncovered risk factors for long COVID, in which some people have symptoms that persist for a long time after they begin to test negative.

The COVID Symptom Study app is the largest community monitoring of COVID in the world.



IMPACT & FUTURE PLANS

The Covid Symptom Study's findings have been presented in well-regarded medical journals and highly cited in peer review literature. The Director of the National Institutes of Health recently highlighted the impact of the Study in his Director's blog. The study has been an effective way of tracking the COVID-19 infection rate; researchers have used the study data in algorithms to predict COVID-19 infections with nearly 80% accuracy among the 2.5 million people who used the app between March 21, 2020 and April 21, 2020. Using the app to detect the spread of the disease is particularly useful in the absence of widespread testing.

The next phase of research is focused on the COVID-19 vaccine response. This research will inform the public health community's understanding of vaccination rates, vaccine hesitancy, and nationwide uptake by demographics like race, socioeconomic class, and geographic location. Currently, nationwide vaccine uptake information is very fragmented because of challenges in systematic data collection.

Future research will also focus on the risk of developing COVID-19 infection or symptoms post-vaccination, the role of diet and lifestyle in COVID-19 risk, and the impact of COVID-19 on mental health and personal/socioeconomic disruption ■

COVID-19 Molecular Explorer

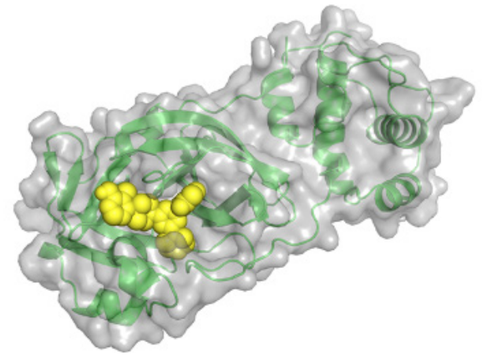
IBM Research

The COVID-19 Molecular Explorer is an AI-empowered platform that accelerates the discovery of potential new drug candidates.

Armonk, NY

OVERVIEW

IBM Research partnered with IBM Science for Social Good to explore the development and application of Generative AI models to help accelerate the discovery of potential new drug candidates by automating the molecule discovery process. In 2020, IBM Research applied this framework, consisting of a controllable deep generative model, to SARS-CoV-2 protein targets and identified novel molecules as potential COVID-19 therapeutic candidates. These molecules have been shared under an open license through an interactive tool known as the Molecular Explorer. Available online, biomedical and healthcare researchers and scientists can study the molecules and understand their characteristics and relationship to COVID-19 to identify candidates that might have the properties to be further pursued in drug discovery. The platform enables and promotes open innovation, transparent data and insight sharing, and trusted evaluation of new drug candidates at an unprecedented scale and pace.



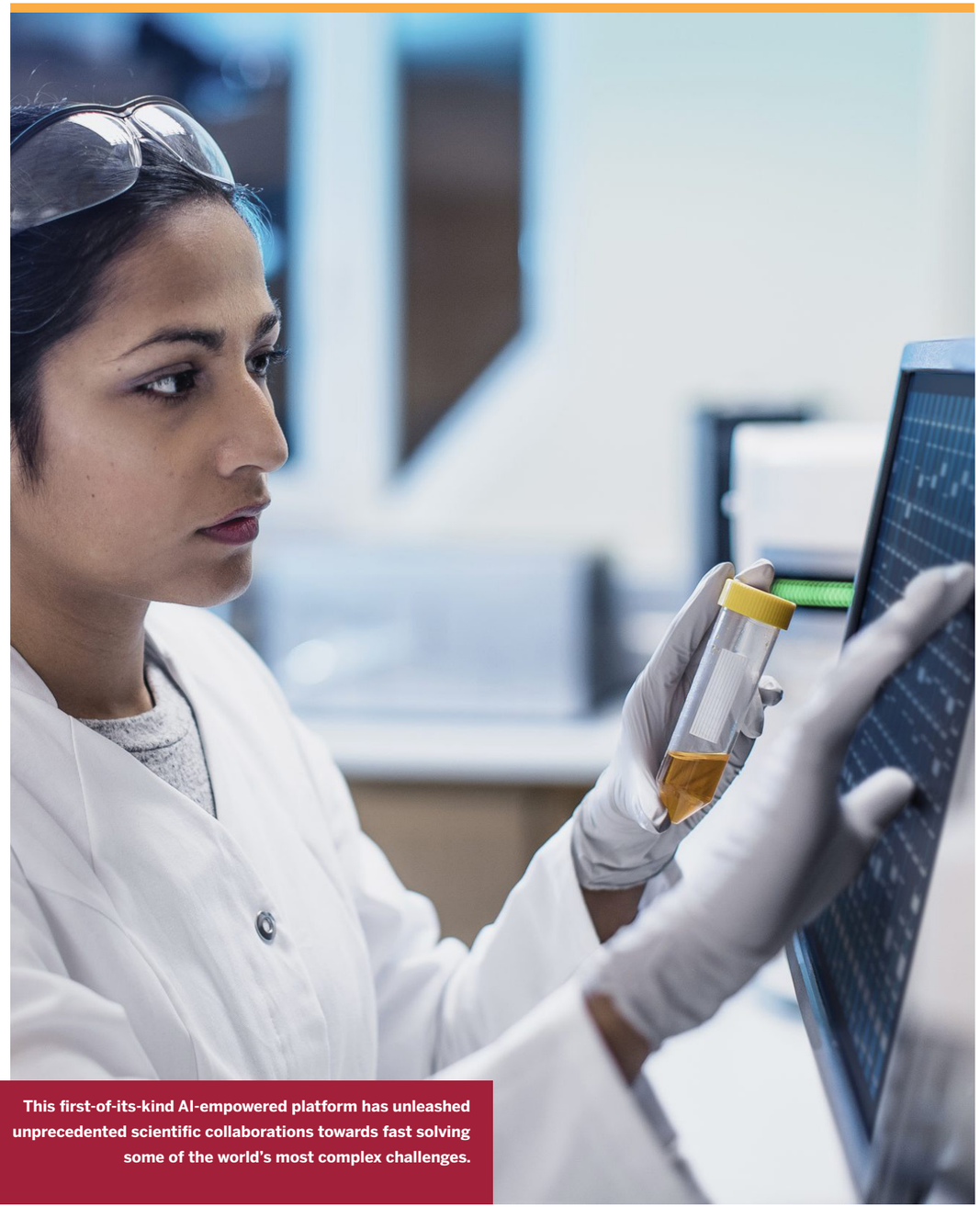
On average, creating a new drug costs as much as $2.6 billion and can take 12 to 14 years for the drug to reach the market.

THE CHALLENGE

Rapid drug discovery processes are critical to dealing with new viral outbreaks and epidemics. The COVID-19 pandemic required a global, collaborative, and rapid response. The traditional drug discovery pipeline is costly, time-intensive, and the research is often siloed. On average, creating a new drug costs as much as $2.6 billion and can take 12 to 14 years for the drug to reach the market. One-third of the overall cost and time is attributed to the drug discovery phase requiring the synthesization of thousands of molecules to develop a single pre-clinical lead candidate. Existing generative AI models have the potential to accelerate the design of novel drug candidates; however, several challenges exist.

ABOUT THE INTERVENTION

IBM's generative AI model has overcome the traditional challenges of applying AI models to accelerate the design of novel drug candidates. This has been accomplished by combining deep learning, generative modeling, novel sampling, and optimization methods. Overcoming the traditional challenges enables the generation of novel artifacts with desired properties. The same AI methods recently discovered novel antimicrobials that will be used in the research and development of antibiotics[1]. Applying this AI technology to three COVID-19 target proteins, IBM identified 3,500 small molecules as potential COVID-19 therapeutics and released these molecules on the COVID-19 Molecular Explorer platform under an open license. This interactive tool allows the user to select different biological targets and filter generated molecules by characteristic and relationship so that researchers can study the molecules and identify candidates that may have desirable properties to be further pursued in drug development.



This first-of-its-kind AI-empowered platform has unleashed unprecedented scientific collaborations towards fast solving some of the world's most complex challenges.

The COVID-19 Molecular Explorer helps researchers generate potential new drug candidates for COVID-19.

IMPACT & FUTURE PLANS

The COVID-19 Molecular Explorer has approximately 9,000 usages worldwide. This first-of-its-kind AI-empowered platform has attracted diverse stakeholders from the public and private sectors and unleashed unprecedented scientific collaborations and partnerships towards quickly solving some of the world's most complex challenges. Most recently, Oxford University and Diamond Light Source have been working together with IBM to screen AI-generated novel drug molecules exposed in the Molecular Explorer tool to find an effective antiviral against COVID-19. Diamond synthesized and tested molecules where they confirmed SARS-CoV-2 activity of these AI-designed novel molecules at an impressively high success rate, showcasing the Molecule Explorer platform's potential and the underlying AI techniques for accelerating the discovery of novel drugs and other molecules and materials. This work demonstrates the future of accelerated discovery, where AI researchers and pharmaceutical scientists can work together in an open and collaborative global community to rapidly create next-generation therapeutics aided by novel AI-powered tools ■

1 Payel Das et al., "Accelerated Antimicrobial Discovery via Deep Generative Models and Molecular Dynamics Simulations," *Nature Biomedical Engineering,* March 11, 2021, 1–11, https://doi.org/10.1038/s41551-021-00689-x.

Dataset Nutrition Labels

The Data Nutrition Project

The Dataset Nutrition Label provides targeted information about a dataset's quality.

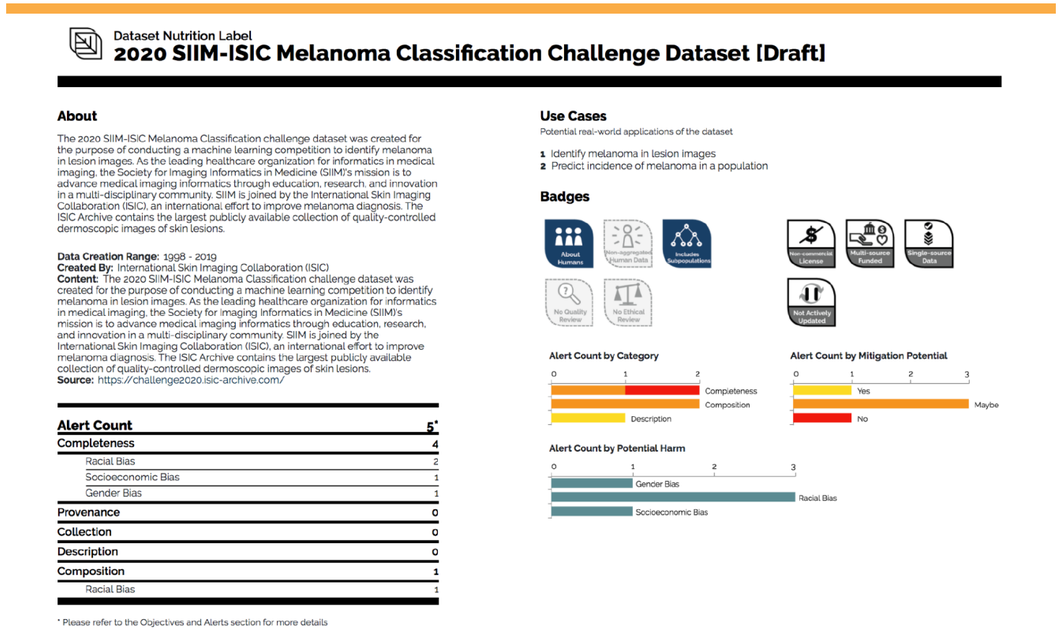
Jersey City, New Jersey

OVERVIEW

Data science is increasingly used by governments and businesses to make decisions that impact billions of lives worldwide. However, the models and algorithms developed by data scientists are only as unbiased as the data they are fed. Consequently, it is critical for data practitioners to build and train systems with as little bias as possible to develop equitable algorithms.

There are millions of publicly available datasets and no existing standards to rate a set's bias or determine its completeness. The Data Nutrition Project created the Dataset Nutrition Label to provide an at-a-glance evaluation of a dataset's quality. The Label cultivates a culture of transparency, accountability, and understanding in the realm of big data.

The Data Nutrition Project was founded in 2018 through Harvard University's Berkman Klein Center Assembly Fellowship.



Dataset Nutrition Label

2020 SIIM-ISIC Melanoma Classification Challenge Dataset [Draft]

About

The 2020 SIIM-ISIC Melanoma Classification challenge dataset was created for the purpose of conducting a machine learning competition to identify melanoma in lesion images. As the leading healthcare organization for informatics in medical imaging, the Society for Imaging Informatics in Medicine (SIIM)'s mission is to advance medical imaging informatics through education, research, and innovation in a multi-disciplinary community. SIIM is joined by the International Skin Imaging Collaboration (ISIC), an international effort to improve melanoma diagnosis. The ISIC Archive contains the largest publicly available collection of quality-controlled dermoscopic images of skin lesions.

Data Creation Range: 1998 - 2019
Created By: International Skin Imaging Collaboration (ISIC)
Content: The 2020 SIIM-ISIC Melanoma Classification challenge dataset was created for the purpose of conducting a machine learning competition to identify melanoma in lesion images. As the leading healthcare organization for informatics in medical imaging, the Society for Imaging Informatics in Medicine (SIIM)'s mission is to advance medical imaging informatics through education, research, and innovation in a multi-disciplinary community. SIIM is joined by the International Skin Imaging Collaboration (ISIC), an international effort to improve melanoma diagnosis. The ISIC Archive contains the largest publicly available collection of quality-controlled dermoscopic images of skin lesions.
Source: https://challenge2020.isic-archive.com/

Alert Count	**5***
Completeness	**4**
Racial Bias	2
Socioeconomic Bias	1
Gender Bias	1
Provenance	**0**
Collection	**0**
Description	**0**
Composition	**1**
Racial Bias	1

* Please refer to the Objectives and Alerts section for more details

Use Cases

Potential real-world applications of the dataset

1 Identify melanoma in lesion images
2 Predict incidence of melanoma in a population

Badges

About Humans
Non-aggregated Human Data
Includes Subpopulations
Non-commercial License
Multi-source Funded
Single-source Data
No Quality Review
No Ethical Review
Not Actively Updated

Alert Count by Category

0 1 2
Completeness
Composition
Description

Alert Count by Mitigation Potential

0 1 2 3
Yes
Maybe
No

Alert Count by Potential Harm

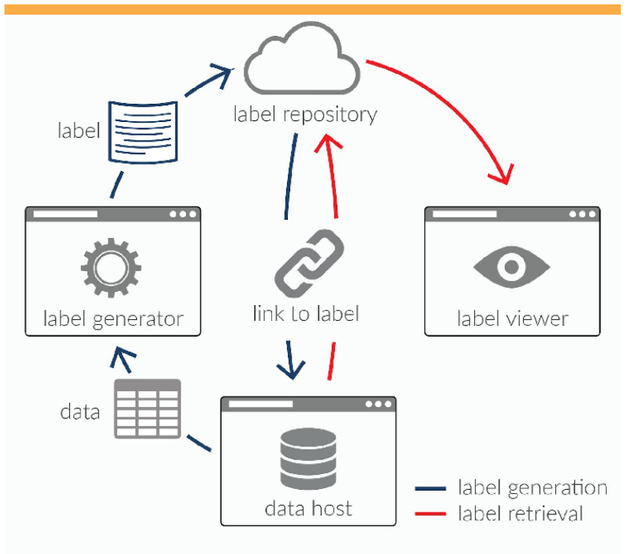
0 1 2 3
Gender Bias
Racial Bias
Socioeconomic Bias

The Data Nutrition Project aims to create a standard label for interrogating datasets.

THE CHALLENGE

Today, AI algorithms are used to inform decisions ranging from mortgage approvals to criminal sentencing to medical care, and more. The quality of AI algorithms is dependent on the quality of the data they are trained on. When the data is biased, the resulting algorithm is biased. For example, dataset bias has contributed to poorer medical attention and higher mortality rates for skin cancer patients with darker skin compared to those with lighter skin[1].

There is currently a missing step in the AI development pipeline: data practitioners cannot assess datasets against a standardized measure of quality that captures both the quantitative and qualitative attributes of a dataset. In short, there is no quick way of knowing the quality of a dataset. Thus the challenge in developing a tool that allows data practitioners to assess datasets quickly is twofold. The first challenge lies in developing the assessment criteria itself. The second challenge lies in expressing a label in a way that is easily understood by practitioners training AI models.



The architecture of the Dataset Nutrition Label ecosystem comprises two main components: a label maker and a label viewer.

ABOUT THE INTERVENTION

The Dataset Nutrition Label increases the transparency and understandability of datasets. Using the analogy of the Nutrition Facts Label on food, the Dataset Nutrition Label highlights the "nutrients" of datasets to concisely summarize how healthy a data set is for a particular use case.

The Label has four sections: About, Alert Count, Use Cases, and Badges. The About section contextualizes the dataset. The Alert Count measures any issues with a dataset's completeness, provenance, collection, description, and composition. Alerts are classified into three categories based on whether they can, cannot, or might be able to be mitigated. The Use Case section describes potential appropriate uses of the dataset. Lastly, the Badges communicate a variety of information like whether the dataset has undergone quality or ethical review. Alerts are also tallied and expressed graphically to show their categorical frequency, their mitigation potential, and their potential harm.

The Dataset Nutrition Label is helpful to dataset owners and data practitioners. For dataset owners, the Label provides scaffolding in the form of questions and processes to surface relevant information about a dataset. For data practitioners, the Label helps inform whether and how to use a dataset. The code and framework are entirely open-source.

IMPACT & FUTURE PLANS

After two years of developing a robust assessment framework, The Data Nutrition Project worked with a consortium of hospitals to create Dataset Nutrition Labels pertaining to melanoma and rare diseases long associated with biased medical datasets.

The Data Nutrition Project is looking to broaden its impact. In the short term, the organization is developing 10 to 20 new labels on several Harvard datasets popular for benchmarking in academia and industry. The Data Nutrition Project is also working with organizations like Humans in the Loop to mitigate problems experienced by refugees perpetuated by biased datasets. Additionally, they have partnered with AI Global and the World Economic Forum to develop a certification for healthy AI using a data rubric based on the Dataset Nutrition Label schema. Longer-term, The Data Nutrition Project is focused on scalability. The Project is investigating how to automate the production of the Label to make the Label as accessible and impactful as possible ■

1 Angela Lashbrook, "AI-Driven Dermatology Could Leave Dark-Skinned Patients Behind," The Atlantic, August 16, 2018, https://www.theatlantic.com/health/archive/2018/08/machine-learning-dermatology-skin-color/567619/.

Garbage In, Garbage Out: Face Recognition on Flawed Data

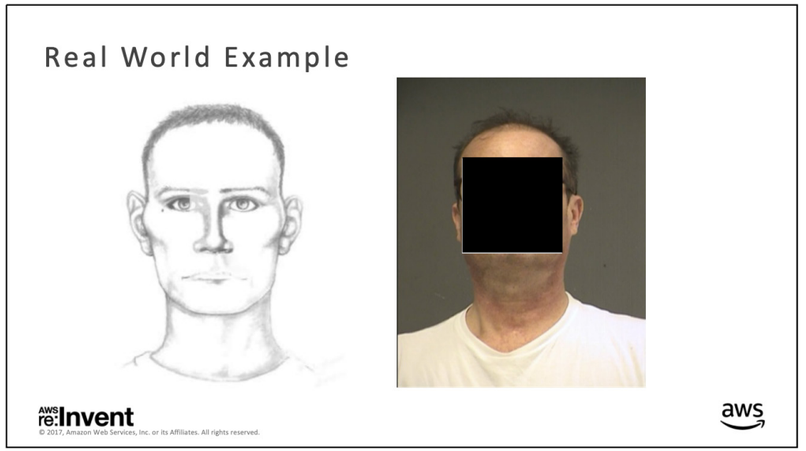
Center on Privacy & Technology at Georgetown Law

"Garbage In, Garbage Out: Face Recognition on Flawed Data" is a report that uncovers the inconsistent and unregulated uses of face recognition by law enforcement.

Washington, D.C.

OVERVIEW

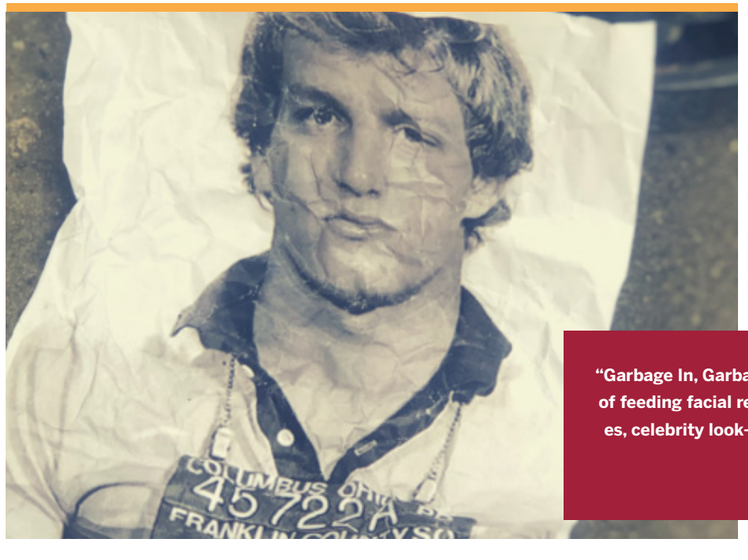
"Garbage In, Garbage Out" is a 2019 report published by the Center on Privacy & Technology at Georgetown Law, one of several Center publications exploring law enforcement's use of facial recognition technology. These reports aim to document the corrosive effects of facial recognition technology on privacy, civil rights, and civil liberties, including defendants' due process rights. Additionally, the paper suggests substantive reforms to protect these rights against such incursions. The Center's 2016 report "The Perpetual Lineup" uncovered the extent of police use of facial recognition. In contrast, "Garbage In, Garbage Out" specifically explores troubling uses of facial recognition technology in criminal investigations, such as locating suspects by feeding facial recognition programs heavily edited images, police sketches, or celebrity look-alike photos.



The practice of using face recognition and artists' sketches is highlighted by Amazon Web Services in a case study about the capabilities of its face recognition software, Rekognition.

THE CHALLENGE

As described in "Garbage In, Garbage Out," when attempting to locate a suspect, police officers can input "probe photos" of unknown individuals into face recognition algorithms for comparison against photographs in government databases. Absent any rules about which images may be used, police run face recognition searches on police sketches, low-quality stills from surveillance camera footage, social media posts, and even celebrity doppelgängers. If images are of poor quality or do not show the subject's full face, officers may heavily edit the photos before running facial recognition, going so far as to mirror partial faces or replace open mouths with closed ones pulled from Google Images. Although most law enforcement agencies do not consider face recognition matches as positive identification sufficient to make an arrest, the lack of clear guidance on law enforcement's use of face recognition technology means that police will sometimes make arrests almost solely on the basis of face recognition alone. Furthermore, when suspects are charged on this basis, they receive little or no information about the role face recognition played in their arrest.[1] The widespread and growing use of face recognition technology by law enforcement absent strong guidelines on appropriate use has opened the way for the erosion of due process in the American criminal justice system.



"Garbage In, Garbage Out" explores the troubling use of feeding facial recognition programs police sketches, celebrity look-alike photos, and other low-quality inputs in criminal investigations.

ABOUT THE INTERVENTION

"Garbage In, Garbage Out" leverages meticulous research to uncover the dangerously inconsistent and unregulated uses of face recognition by law enforcement. It is the first report of its kind, presenting the general public with a detailed and highly readable look at facial recognition in the criminal justice system. The report systematically documents multiple instances in which major police departments have utilized police sketches, computer edited images, and celebrity look-alike photos in conjunction with facial recognition databases in order to apprehend suspects.

Importantly, the report also shows that there are multiple instances where police officers have apprehended suspects solely or almost solely on the basis of a possible match from a facial recognition system – a practice which often flouts departmental and agency regulations.

IMPACT & FUTURE PLANS

As a result of this report, some police departments have changed their policies to restrict the use of heavily edited photos or forensic sketches in face recognition systems. "Garbage In, Garbage Out" has contributed greatly to legislative efforts focused on protections for due process in the age of facial recognition. The Utah State Legislature passed a law that enhances internal checks against misidentification in the face recognition process and requires prosecutorial disclosure, which may help alleviate endemic due process issues. A number of other states are considering similar legislation, or placing moratoria or bans on the use of face recognition outright. Armed with the findings of "Garbage In, Garbage Out," the Center's 2016 report "The Perpetual Lineup," and trainings provided by the Center on Privacy & Technology, defense attorneys have begun challenging the use of face recognition in their clients' criminal cases. The Center is also advocating for the institution of comprehensive federal legislation delineating limits on the use of facial recognition and guarantees to due process rights in the face of this new technology ■

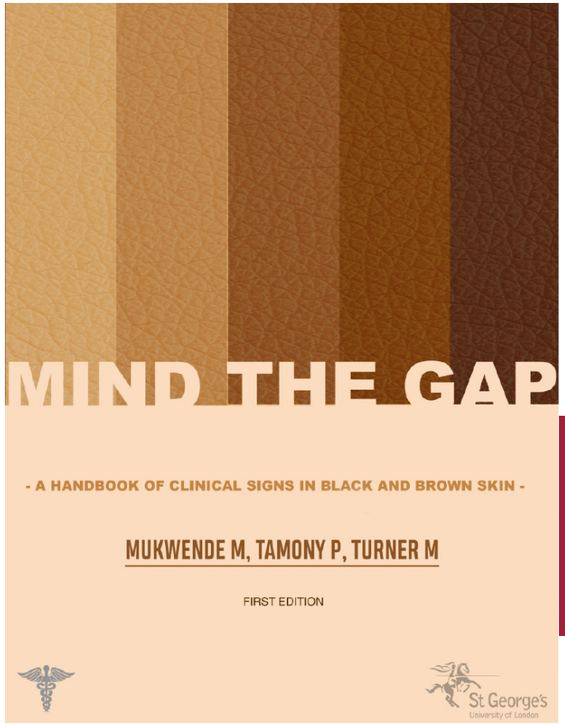
1 Clare Garvie, "Garbage In. Garbage Out. Face Recognition on Flawed Data," Garbage In. Garbage Out. Face Recognition on Flawed Data, May 16, 2019, https://www.flawedfacedata.com.

Mind the Gap

Black and Brown Skin

Mind the Gap is a clinical handbook of signs and symptoms in people with black and brown skin.

London, UK



Mind the Gap is a dynamic web-based living document that crowdsources a growing database to address gaps in medical descriptors and imagery.

OVERVIEW

Mind the Gap is a free clinical handbook that describes signs and symptoms of skin disease as they present in people with black and brown skin. The handbook aims to address a long-standing problem: the majority of medical texts worldwide underreport how symptoms manifest on darker skin. This bias increases the risk of errors and misdiagnoses, leading to poorer health outcomes for people of color.

The first iteration of Mind the Gap was published online in August 2020 and included details on 24 skin conditions. Unlike a traditionally printed handbook, Mind the Gap is a dynamic web-based living document that can be constantly updated with additional images and data. The website allows users worldwide to submit labeled images of diseases on black and brown skin, effectively crowdsourcing a growing database used by various stakeholders to address gaps in medical imagery.

THE CHALLENGE

The lack of diversity in the medical field is a well-known issue that has had pernicious effects on health outcomes, medical education, and the experiences of healthcare professionals. Most medical textbooks that teach about diagnosing skin disorders do not include images of skin conditions as they appear on people of color. Because many key characteristics of skin disorders appear differently on different complexions, this lack of representation can lead to underreporting, misdiagnosis, and unnecessary suffering.

For example, African Americans have the lowest survival rate for melanoma out of any racial group in the U.S., with a five-year survival rate of 66% compared to 90% for white patients[1]. The relative lack of information regarding, and experience diagnosing melanoma in dark skin reduces the likelihood of accurate and early diagnoses that might improve a patient's survival rate.

The recent COVID-19 pandemic has further illustrated the extent of racial health disparities, with some studies finding people of color to be four times more likely to be negatively affected by COVID-19 than white people. Common approaches to identifying COVID-19 patients have included looking for blue discoloration in lips and skin tone, as well as blistered digits known as "COVID toes." These symptoms present differently in people with dark skin, but these details have not been well-captured. For example, when dermatologists started an international registry to catalog skin manifestations of COVID-19, they compiled conditions in more than 700 patients, of which only 34 were Hispanic and 13 were Black[2]. These factors contribute to the fact that COVID-19 has been less accurately diagnosed and thus more likely to spread among communities of color[3].

ABOUT THE INTERVENTION

Mind the Gap originated as a staff-student partnership project at St. George's University of London. Between December 2019 - May 2020, the team developed the first iteration of the Mind the Gap handbook, with the intent for it to serve as a tool for addressing the absence of details, imagery, and terminology around skin conditions in darker skin. They started by trying to compile images and information on conditions known to have different symptoms in lighter vs darker skin. For example, they compiled information on Kawasaki disease, which presents with a red rash on white skin, but is less conspicuous on dark skin tones. However, they encountered substantial difficulty in finding and collecting appropriate images—a challenge shared by many others in the field. To address this, they created a website to publicly crowdsource labeled images, so that clinicians, medical students, and patients could actively contribute their own images and descriptions. In June 2020, they published the first edition of Mind the Gap, which included information on 24 skin conditions.

The majority of medical texts worldwide have been significantly biased toward lighter skin leading to poorer health outcomes for people of color.

IMPACT & FUTURE PLANS

Since its release, Mind the Gap has received exuberant global attention and has been featured across numerous international platforms. As of March 2021, the handbook had been downloaded over 150,000 times by people in at least 106 countries. It has also been added to the curricula and recommended reading lists of at least 20 universities and hospitals. The work has been featured in the House of Lords, Sky News, the Washington Post, the BMJ, NBC News, Medscape, Al Jazeera, Bloomberg News, CBC, Fox 5, and ITV News.

The team is in conversation with a wide variety of potential partner organizations, including research institutions, governments, and technology companies. They are continuing to collect crowdsourced images through their website, and plan to release a second edition of the Mind the Gap handbook in both print and digital formats later in 2021 ■

---

1 Carolyn McMillan and U. C. Newsroom, "The Skin Care Myth That Harms People of Color," University of California, July 17, 2019, https://www.universityofcalifornia.edu/news/skin-color-clinics-aim-end-health-care-disparities-dermatology.

2 Roni Caryn Rabin, "Dermatology Has a Problem With Skin Color," *The New York Times*, August 30, 2020, sec. Health, https://www.nytimes.com/2020/08/30/health/skin-diseases-black-hispanic.html.

3 Jill C. Muhrer, "Risk of Misdiagnosis and Delayed Diagnosis with COVID-19," *The Nurse Practitioner* 46, no. 2 (February 2021): 44–49, https://doi.org/10.1097/01.NPR.0000731572.91985.98.

Project Amelia

Probable Models

Project Amelia uses immersive theater to help people grapple with the challenges and consequences of tech companies' most pressing ethical dilemmas.

Pittsburgh, PA

OVERVIEW

Project Amelia is an immersive theater production that invites audiences to learn about issues in data privacy, AI ethics, and corporate governance. Audience members participate in a fictional company's troubled launch of a superintelligent AI assistant that uses the participants' real-world data. The show was developed by technologists, researchers, and AI practitioners who wanted to provide an accessible means for people to better understand critical technology issues. The show uses various technologies, including RFID bracelets, custom smartphones, indoor localization, and facial recognition to engage audiences and expand their imagination of what can be done with tech products. Research conducted in partnership with Carnegie Mellon University has shown the immersive theater medium to be an effective method of tech pedagogy.[1]

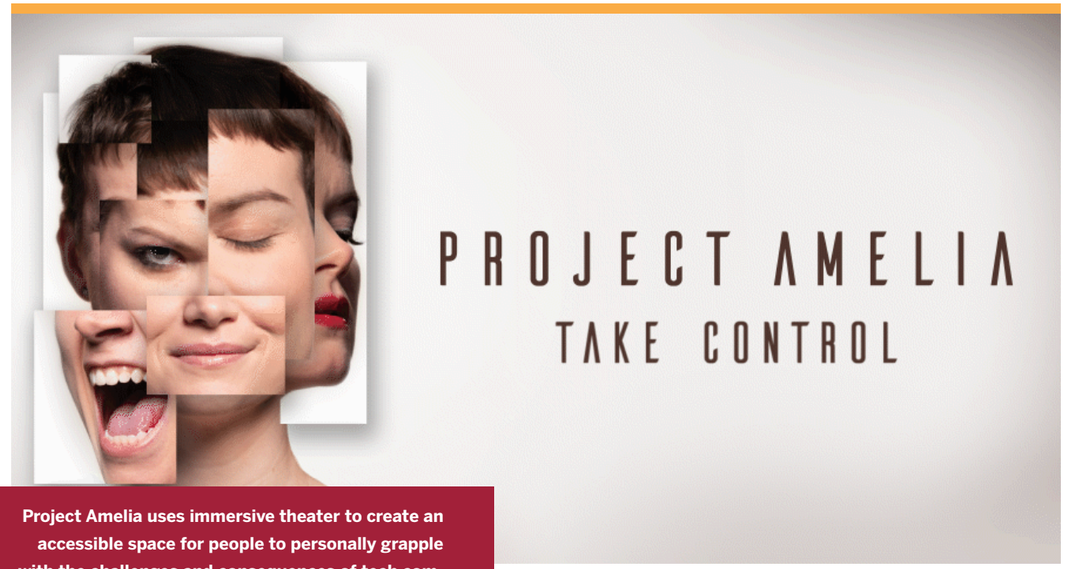
Project Amelia simulates what it might look like to engage with a powerful near-future tech company whose superintelligent AI assistant does not function as intended.

THE CHALLENGE

Today, many people recognize that there are deeply concerning issues in how technology companies use their personal data, but feel helpless to do anything in response. Even industry experts often find it difficult to comprehend, navigate, and meaningfully discuss key ethical challenges around technology. Furthermore, conversations regarding technological dilemmas that do happen tend to be siloed between industries and issues, limiting their reach and impact. The result of these dynamics has been a deep erosion of trust, which has only continued to worsen with each subsequent crisis and controversy. In order to rebuild trust and find consensus, it is necessary to have spaces where diverse stakeholders can come together and engage in critical conversations about how best to shape future ethical outcomes.

ABOUT THE INTERVENTION

Project Amelia was initially conceived and piloted in 2016 under the title "Quantified Self," as part of the writer's doctoral research on technology ethics and science pedagogy. This research and other studies showed compelling evidence that fictional storytelling could be a more effective method for educating and engaging people in issues around technology ethics compared to other media, such as news articles or terms of service agreements.[2] Continued interest and inquiry around the topic



Project Amelia uses immersive theater to create an accessible space for people to personally grapple with the challenges and consequences of tech companies' most pressing ethical dilemmas.

led to a full production of Project Amelia in 2019, through a partnership between the technology ethics firm, Probable Models, and the nonprofit Bricolage Production Company, with funding from several foundations and corporate sponsors.

The resulting experience used an immersive theater format to create a safe and accessible space for people to personally grapple with the challenges and consequences of tech companies' most pressing ethical dilemmas. Through its dynamic plot, multiple endings, and many participatory scenes and mechanisms, Project Amelia simulated what it might look like to engage with a powerful near-future tech company as an employee, board member, activist, journalist, consumer, or other stakeholder. Up to 60 participants were assigned one of several roles, based on data collected from them before the show. The participants would then be equipped with a corresponding smartphone and RFID wristband which they could use to interact with each other and the immersive set. Throughout the show, participants would encounter difficult choices and would often need to work with actors and participants alike to achieve their desired outcomes or endings.

Participants were encouraged to stay after the show for a debriefing session to discuss their questions, experiences, and reflections with the actors, producers, and other participants. For many, this debrief would prove to be the most powerful aspect of their learning experience.

IMPACT & FUTURE PLANS

Project Amelia premiered in Fall 2019 for a 12-week sold-out production in Pittsburgh, PA. The show was primarily open to the public. The show was also delivered for some private audiences as a corporate training exercise. Project Amelia reached over 5,000 people, receiving overwhelmingly positive reviews and critical acclaim.

Collaborative research with Carnegie Mellon University studied the effectiveness of immersive theater as a pedagogical method. Among participants who opted to respond to a post-show survey, most reported a significant increase in their understanding of, concern around, and desire to act on data privacy and technology ethics issues.

Project Amelia also worked with several international artist-technologists to co-design a variety of interactive exhibits, disguised as product demos of the fictional company. These exhibits were designed as easily constructed, portable installations to facilitate mini-experiences at other venues, such as conferences, schools, and public settings.

The producers hope to replicate and scale the production, pending recovery from the COVID-19 pandemic. There has been interest in touring the show across cities and setting up permanent residency in certain cities. Any such iterations would certainly facilitate more public discourse and collaborative opportunities. ■

---

1 Maggie Oates, "Privacy in Unusual Contexts: A Case Study of A Theater Company," 2019, https://www.usenix.org/conference/pepr19/presentation/oates.

2 Michael Warren Skirpan, Jacqueline Cameron, and Tom Yeh, "More Than a Show: Using Personalized Immersive Theater to Educate and Engage the Public in Technology Ethics," **in** *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18 (New York, NY, USA: Association for Computing Machinery, 2018), 1–13, https://doi.org/10.1145/3173574.3174038.

Project Lighthouse

Airbnb

Project Lighthouse is a multistakeholder effort to quantify and help prevent discriminatory practices by users on the Airbnb platform.

San Francisco, CA

OVERVIEW

Designed with leading civil rights organizations like Color Of Change, Upturn, AAJC, NAACP, and more, Project Lighthouse is an anonymized research approach to help determine when and where racial discrimination happens on the Airbnb platform and the effectiveness of policies that fight it. Project Lighthouse maintains the privacy of sensitive demographic data to prevent its exploitation while providing Airbnb users with agency over whether they participate in the program. Through its novel privacy and inclusion-centric approach, Project Lighthouse aspires to elevate building for equity as a standard for responsible innovation in the minds of technologists.

Project Lighthouse is a groundbreaking initiative launched in the United States to uncover, measure, and help overcome discrimination when booking or hosting on Airbnb.

THE CHALLENGE

Following a civil rights audit in 2016, Airbnb took steps to reduce discrimination on their platform, including the creation and enforcement of a strict Non-Discrimination Policy, removal of guest profile photos from the booking process, and the increase of listings available for instant booking, among other changes. Since 2016, over 1.4 million people have declined to agree to Airbnb's mandatory Community Commitment and Non-Discrimination Policy, and have been removed from the platform. However, because Airbnb does not ask its users for racial demographic data, the company struggles to measure the effectiveness of these policies or how else people may be discriminated against on their platform.

Airbnb, like online platforms and marketplace rental services in general, needs data and tools to combat discrimination effectively. But studying demographic data opens a risk of exploitation by external actors with nefarious intentions or internal actors inadvertently causing harm.

Consequently, the main challenge in developing a system to study discrimination on the platform is twofold: (1) designing a privacy-centric process that utilizes anonymized perceived race data and (2) trying to ensure that the resultant data can be used to accurately measure potential experience gaps in the product.

Project Lighthouse was designed with input from leading civil rights and privacy organizations to thoughtfully measure discrimination while respecting individual privacy.

ABOUT THE INTERVENTION

Project Lighthouse was developed with input from leading civil rights groups and privacy organizations, including Color Of Change, AAJC, Center for Democracy & Technology, The Leadership Conference on Civil & Human Rights, LULAC, the NAACP, National Action Network, and Upturn. In designing a privacy-centric process that utilizes anonymized perceived race data, Project Lighthouse employs p-sensitive k-anonymity to measure experience gaps by perceived race.

K-anonymity means that there are at least k instances of each unique set of values (e.g. for columns number_of_accepts, number_of_rejects used in the technical paper for discussion) in Airbnb's dataset, which can be achieved in a number of ways, such as by averaging comparable number_of_accepts. For instance, number_of_accepts values of 2, 3, and 5 could all be represented as 3.33, thereby removing a unique identifier that could be used to infer the perceived race of a user. K refers to the minimum number of users represented by each unique set of values in the anonymized dataset, or the degree of anonymization. In the above example, k=3 as the 3 individuals' data can no longer be distinguished after the K-anonymization process. In complement to k-anonymity, p-sensitive k-anonymity means that each unique set of values in Airbnb's dataset has at least p distinct perceived race values among them. P-sensitive k-anonymity further protects against the risk of uniquely identifying data in the case of all k users having the same perceived race.

This privacy-centric approach works in concert with A/B testing and other research to support the understanding and mitigation of discrimination that may occur on Airbnb's platform in a way that protects user-perceived demographic data. Additionally, Project Lighthouse grants users agency over their privacy as a user can opt-out at any time.

IMPACT & FUTURE PLANS

Project Lighthouse was in development for two years before being publicly launched in the United States in June 2020. Since September 2020, the Project Lighthouse team has been analyzing collected data to inform future design iterations of platform facets.

The ultimate aspiration of Project Lighthouse is to create a genuinely equitable experience for all users on the Airbnb platform and to elevate the concept of building for equity across all tech companies as a core component of responsible innovation. As a first step, Airbnb has shared its methodology behind Lighthouse in a publicly available technical paper in the hopes that the paper can be a starting point for further innovations that rigorously measure discrimination while upholding user privacy. However, Airbnb acknowledges that Lighthouse is tailored to its platform and not agnostic to the size and structure of other technology companies. Therefore, Airbnb is exploring inventing additional methodologies to provide directional information for technology companies of various sizes to combat discrimination ■

Racial Disparities in Automated Speech Recognition Systems

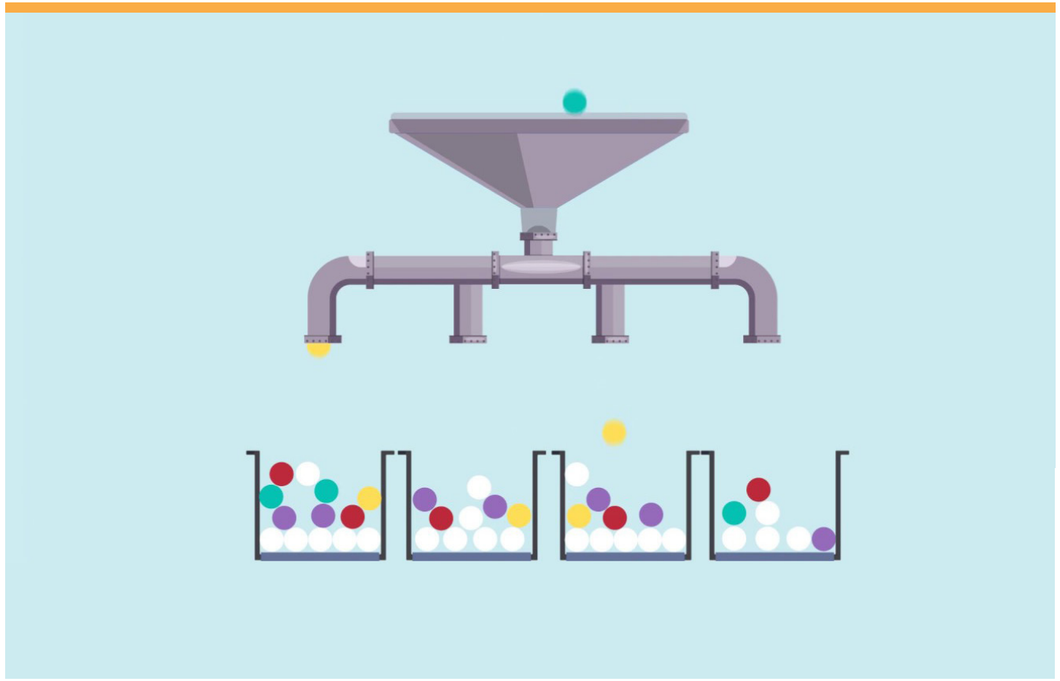
Stanford Computational Policy Lab

"Racial Disparities in Automated Speech Recognition" (ASR) is a report that analyzed racial disparities in the performance of five popular commercial ASR systems.

Stanford, CA

OVERVIEW

"Racial Disparities in Automated Speech Recognition" is a report that analyzed how accurately automated speech recognition systems (ASRs) interpret audio by white and Black speakers. In 2019, the project investigated the leading speech-to-text systems built by Amazon, Apple, Google, IBM, and Microsoft. The researchers fed thousands of audio samples from 42 white and 73 Black men and women and analyzed resulting error rates. The paper found significant disparities by race that were compounded by gender. Furthermore, the report revealed that the technology tested performed far worse for speakers who used more linguistic features characteristic of African American Vernacular English. The disparities were consistent across all five firms studied.

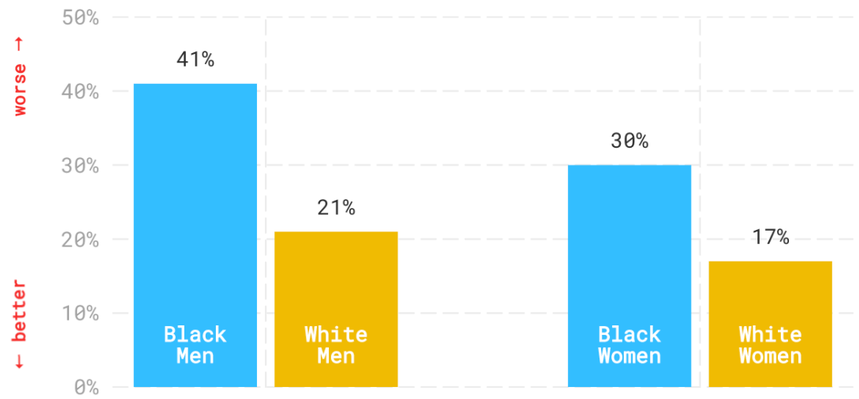


To ensure that ASR technology is inclusive and available to all users, it is critical that academic researchers and industry professionals develop more diverse datasets.

THE CHALLENGE

AI is rapidly moving into everyday life and influencing decision making in various industries such as healthcare and law enforcement. However, as summarized in "Racial Disparities in Automated Speech Recognition," researchers have uncovered that applications in machine learning suffer from racial bias in many facets of automation. Those groundbreaking studies consistently revealed that the underlying machine learning technology is based on incomplete and unrepresentative datasets. In the case of ASR systems, "Racial Disparities in Automated Speech Recognition" concluded that "machine learning technology underlying speech recognition systems likely relies too heavily on audio samples of white Americans."[1]

ASR technology can benefit everyone and range from everyday convenience to life-changing aids. Unfortunately, at present, not everyone can equally take advantage of these powerful new tools.



The report found that error rates in ASRs for Black speakers nearly double those for white speakers.

ABOUT THE INTERVENTION

"Racial Disparities in Automated Speech Recognition" was authored by researchers affiliated with the Stanford Computational Policy Lab. The report was published in the Proceedings of the National Academies of Sciences of the United States of America.

The report demonstrated that for every hundred words spoken, the systems studied made 19 errors for white speakers compared to 35 errors for Black speakers. The word error rates were worst for Black men. For every hundred words spoken, the systems studied made 41 errors for Black men, 21 errors for white men, 30 errors for Black women, and 17 errors for white women.

The report also considered the full distribution of error rates across the populations of white and Black speakers which revealed starker disparities. For example, less than 2% of recordings of white speakers had an error rate above 50%. In contrast, more than 20% of recordings of Black speakers had an error rate of at least 50%.

The report concludes that closing the ASR performance gap will require datasets that reflect the full diversity of accents and dialects of all Americans. Furthermore, the researchers assert that speech recognition tools should be regularly assessed and academia/industry should publicly report progress in making ASRs more broadly inclusive.

IMPACT & FUTURE PLANS

"Racial Disparities in Automated Speech Recognition" received significant media attention. The report was mentioned in 44 news stories by 36 news outlets including The New York Times, Brookings, Medium, World Economic Forum, Forbes, Scientific American, and more. The paper scored in the top 5% of all research ever tracked by Almetric, a measurement of research impact online that has tracked over 32.1million research outputs.

The report is the basis of the Stanford Computation Policy Lab's Fair Speech project[2], which currently hosts an interactive piece of journalism on racial disparities in ASR. "Racial Disparities in Automated Speech Recognition" also inspired "Voicing Erasure,[3]" a moving recorded poem produced by the Algorithmic Justice League featuring leading scholars on race, gender, and technology. Additionally, the research team continues to advise and help the groups that reach out to them in regard to how ASR systems can be adjusted to account for the discrepancies identified in this report.

In the future, the report's lead author, Allison Koenecke, would like to see more research on disparities in automation in other domains of technology ■

1 Allison Koenecke et al., "Racial Disparities in Automated Speech Recognition," *Proceedings of the National Academy of Sciences* 117, no. 14 (April 7, 2020): 7684–89.
2 Sharad Goel et al., "The Race Gap in Speech Recognition," accessed April 23, 2021, https://fairspeech.stanford.edu.
3 Joy Buolamwini, Voicing Erasure – *A Spoken Word Piece Exploring Bias in Voice Recognition Technology*, 2020, https://www.youtube.com/watch?v=SdCPbyDJtKo.

Safe House and Shelter Training Program

Operation Safe Escape

The Safe House and Shelter Training Program teaches and provides computer, physical, operations, and crime prevention through environmental design (CPTED) security to domestic violence safe houses and shelters.

Glen Burnie, MD

OVERVIEW

When domestic violence victims escape their abusers, they often turn to safe houses and shelters (SHASs). Unfortunately, new forms of technology have enabled abusers and made it more difficult for victims to successfully escape. The Safe House and Shelter (SHAS) Program by Operation Safe Escape helps victims of domestic violence and related crimes escape abusive relationships by providing SHAS staff with computer, physical, operations, and CPTED security training and services at no cost. Since 2016, through the SHAS Program, Operation Safe Escape has helped over 3,000 individuals successfully escape their abusers.

Since 2016, Operation Safe Escape has helped over 3,000 individuals successfully leave their abusers.

THE CHALLENGE

According to the CDC's National Intimate Partner and Sexual Violence Survey, in the U.S., over 1 in 3 (43.6 million) women and about 1 in 3 (37.3 million) men experience contact sexual violence, physical violence, and/or stalking by an intimate partner during their lifetime[1].

Growing evidence suggests that during the COVID-19 pandemic, domestic violence has become more common and often more severe. Stay-at-home orders, limited mobility, and increased financial/social distress have exacerbated domestic violence cases. The United Nations has referred to the global rise in domestic violence as the "shadow pandemic" within the COVID-19 health crisis.[2]

Moreover, new technology has enabled abusers. Historically, abusers may have monitored a victim by checking a car's mileage or manipulating neighbors into reporting a victim's activities; today, abusers have access to a plethora of legal location tracking apps and illegal stalkerware that severely limit a victim's freedom of movement. Abusers are also able to leverage social media platforms to teach, support, and share how to best leverage technology against a victim's will. Finally, advances in tracking devices and increasingly deceptive cybersecurity attacks pose growing security risks to the victims and SHAS staff /clients.

One of the most dangerous moments a domestic violence victim faces is when they attempt to flee an abuser. Victims often turn to SHASs for advice and support. SHASs are typically staffed by non-security specialists, such as retired healthcare workers, social workers, and local concerned citizens. While the staffers provide critical services to victims, they generally lack security awareness training to combat the growing threat posed by technology-equipped abusers. Without proper security support, victims may be threatened into returning to the abusive relationship. In fact, as quoted in Time Magazine, according to

Cassie Mecklenberg, Executive Director of the domestic violence support group Sheltering Wings, "on average, survivors return to the abusive relationship seven times before they leave for good."[3] The SHAS Program works to address the security deficits that may threaten a successful escape.

ABOUT THE INTERVENTION

The SHAS Program focuses on teaching and providing critical security expertise at no cost. Like a customer service model, the SHAS Program's services are tailored to an organization's needs. Services include computer, physical, operations, and CPTED security. Below are three examples of the SHAS Program security support services:

- Computer Security: With limited resources and cybersecurity expertise, several SHASs have chosen to build their websites and communication channels through free or inexpensive services that can easily be infiltrated or attacked. To protect clients and staff, the SHAS Program works with organizations to set up secure networks, websites, and communication channels and teach staff how to practice computer security.
- Physical Security: The SHAS Program has helped organizations secure their technical hardware such as wiring closets and routers from attack. They've also helped sweep SHAS facilities to remove technical hardware devices planted by abusers such as listening, monitoring, and tracking devices.
- Digital Evidence Collection and Protection: In many cases, SHAS employees become aware of digital evidence that would help to obtain a restraining order and the prosecution of an abuser. The SHAS Program helps staffers learn how to handle and protect digital evidence properly.

Additionally, in 2019, Operation Safe Escape hosted the 2019 Domestic Violence Safety and Security Conference. The conference taught approximately 250 SHAS staff attendees relevant security issues. Through this conference, a shelter staff attendee finally cracked a mystery regarding stalkerware used on a client that the police had not solved for years.

It is unrealistic to expect overworked retired healthcare professionals, social workers, and other concerned citizens who serve as SHAS staff to also be cyber, physical, and operation security experts. Moreover, these under-resourced organizations should not be expected to pay large sums of money to security companies for training and support. And yet, as abusers become more tech-savvy, a security background becomes increasingly necessary to help domestic violence victims escape. That's why Operation Safe Escape leverages its 100+ thoroughly vetted security professional pro-bono volunteers to provide critical security expertise at no cost. This allows SHASs to focus their limited resources towards other pressing needs such as healthcare, re-housing, and helping victims establish financial independence.

The SHAS Program is just 1 of 4 initiatives supported by Operation Safe Escape. Altogether, the four initiatives (1) educate victims, SHAS staff, law enforcement, and industry (2) help victims escape their abusers, and (3) empower survivors with the technical support they may need.

IMPACT & FUTURE PLANS

The SHAS Program has 12 enduring SHAS partnerships and serves many more SHASs on an as-needed basis across the United States and internationally in Canada, Australia, Ghana, and elsewhere.

Since launching in 2016, Operation Safe Escape has helped over 3,000 individuals successfully leave an abuser. They have distributed nearly 1,000 TAILS (a privacy software that allows for encrypted communication) to help victims who do not have access to secure networks.

Operation Safe Escape is continuing to scale the SHAS Program. There is currently minimal information sharing between shelters and safe houses. Opening regional communication channels would be critical to sharing threat intelligence information and best practices as a victim attempts to escape their abuser. Consequently, Operation Safe Escape is currently working on a formal information sharing and collaboration platform for shelters and safe houses all over the country.

Additionally, Operation Safe Escape is creating a comprehensive security guide tailored for domestic violence safe houses and shelters ■

---

1 S.G. Smith et al., "The National Intimate Partner and Sexual Violence Survey : 2015 Data Brief – Updated Release" (Center for Disease Control and Prevention, November 2018), https://stacks.cdc.gov/view/cdc/60893.

2 "The Shadow Pandemic: Violence against Women during COVID-19," UN Women, accessed April 24, 2021, https://www.unwomen.org/en/news/in-focus/in-focus-gender-equality-in-covid-19-response/violence-against-women-during-covid-19.

3 Jeffrey Kluger, "Domestic Violence and COVID-19: The Pandemic Within the Pandemic," Time, February 3, 2021, https://time.com/5928539/domestic-violence-covid-19/.

SmartNoise

OpenDP and Microsoft

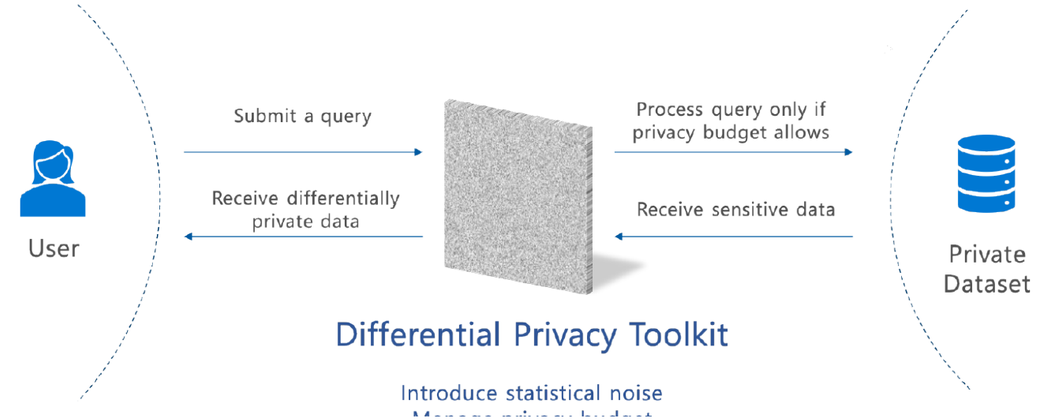
SmartNoise is a toolkit that uses differential privacy techniques to enable privacy-protective evaluation of sensitive data for scientific research and exploration in the public interest.

Cambridge, MA

OVERVIEW

Big data informs policymaking. But because big data is also highly sensitive, there are significant risks to individual privacy. SmartNoise aims to advance privacy-protective analyses by building a community around developing open-source software for differential privacy.

SmartNoise was established as a collaboration between Harvard's Institute for Quantitative Social Science (IQSS) and Microsoft. SmartNoise's platform ensures data is kept private while enabling researchers from academia, government, non-profits, and the private sector to gain new insights that can rapidly advance human knowledge. As an open-source initiative, researchers can use the platform to make their own data sets available to other researchers worldwide.



The toolkit is designed to be a layer between queries and data systems to protect sensitive data.

THE CHALLENGE

Insights from datasets can help solve complex problems in areas such as health, the environment, economic inequality, and more. Unfortunately, because many datasets contain sensitive information, legitimate concerns about compromising privacy currently prevent the use of informative data. By injecting random noise into statistics computed on the dataset, differential privacy techniques make it possible to extract valuable insights from datasets while safeguarding the privacy of individuals.

Unfortunately, using differential privacy can be highly complex. As argued by Feldman et. al, when using differential privacy, "the trade-off between accuracy and privacy is not straightforward. Indeed, a major area of research in differential privacy is figuring out techniques that can improve that trade-off, making it possible to give stronger privacy guarantees at a given level of noise, or less noise at a given level of privacy." The high complexity surrounding the use of differential privacy has made these tools cost-prohibitive for civil society and non-profit organizations to implement. Consequently, practical adoption of differential privacy remains slow despite increasing demand from government agencies and research communities.

The toolkit injects noise into data to prevent disclosure of sensitive information and manage exposure risk.

ABOUT THE INTERVENTION

SmartNoise created a trustworthy suite of differential privacy tools that serve as a public resource for any organization wanting to use differential privacy. SmartNoise focuses on supporting scientifically-oriented research and exploration in the public interest through enabling archival data repositories to take on sensitive data, allowing safe sharing of data from government agencies and companies with researchers, and increasing the robustness of research findings. SmartNoise seeks to make data more accessible and usable and empower researchers to simply and confidently deploy differential privacy tools.

The collaboration between industry and academia allows SmartNoise to create scalable frameworks and tools for companies to share data with university researchers in a secure and privacy-preserving way.

IMPACT & FUTURE PLANS

SmartNoise implemented its tools in both businesses and government agencies. Microsoft has deployed differential privacy tools in technologies such as Windows and Workplace Analytics. Companies, including Facebook, have also been more willing to make datasets available to researchers through the OpenDP platform.

SmartNoise has been used to study student outcomes by granting researchers access to education data from California and Texas. Broadband usage data has also been unlocked to help the Federal Communications Commission and policymakers expand internet access to under-served communities and bridge the digital divide.

To alleviate the cost of adoption, SmartNoise launched an early adopter acceleration program aimed to help civil society and non-profit organizations leverage differential privacy tools without the need to build their own code or platform. SmartNoise experts will provide technical assistance to incorporate differential privacy into the data-sharing processes of projects where unlocking data or insights will benefit society. ■

Vitaly Feldman et al., "Differential Privacy: Issues for Policymakers," June 29, 2020, https://simons.berkeley.edu/news/differential-privacy-issues-policymakers.

Student Privacy Project

Electronic Privacy Information Center (EPIC)

The Student Privacy Project utilizes research, advocacy, and litigation to uphold and expand the privacy rights of students.

Washington, D.C.

OVERVIEW

The Student Privacy Project is an initiative of the Electronic Privacy Information Center (EPIC), a premier public interest research center focused on civil rights issues in the digital era. In light of the expanding use of digital tools in the classroom, EPIC's Student Privacy Project has utilized a combination of research, litigation, and multi-stakeholder advocacy to defend and expand the privacy rights of students for over a decade. For instance, their efforts contributed to New York state's recent moratorium on school-place use of facial recognition technologies. As many students have switched to online instruction in light of the COVID-19 pandemic, the Student Privacy Project has shifted its focus to demanding more transparency and less expansive data collection policies from online proctoring services. This effort to hold online proctoring companies accountable includes a December 2020 complaint filed with the Attorney General of the District of Columbia against the five largest test proctoring companies.

A handful of proctoring companies have gained access to a trove of student personal and biometric data with few internal or external restrictions.

THE CHALLENGE

The last major piece of federal student privacy legislation was passed in 1974, almost five decades ago. Since that time, the landscape of student privacy has changed drastically: student records are mostly digitized, digital tools from third-party vendors are frequently deployed in classrooms, and educators routinely utilize test proctoring software that collect and process student biometric data. These trends have only intensified with the onset of the COVID-19 pandemic and the widespread transition to online learning. A handful of proctoring companies have gained access to a trove of student personal and biometric data with few internal or external restrictions.

Over the past year and a half, the mass deployment of these proctoring tools has posed privacy and equity threats to students. Data collected and disclosed to third parties by some of these widely used proctoring services include information relating to gender identity, citizenship status, and disability status. Much of this data is unnecessary to the proper operation of these proctoring services. Furthermore, the algorithms utilized by many of these services to scan students' faces to determine whether a student is cheating on an exam systematically fail to work on students of color and non-male students.

In the absence of comprehensive federal privacy legislation or updated student privacy legislation, EPIC has sought to protect and expand student privacy rights by utilizing existing statutes.

The Student Privacy Project's advocacy fights to ensure that student privacy is not a casualty of the digital age.

ABOUT THE INTERVENTION

In December of 2020, EPIC filed a complaint with the Attorney General of the District of Columbia alleging that the collection and processing of student data by the five most popular online proctoring services violate D.C. consumer protection law. Through meticulous research, the Student Privacy Project revealed the extent to which these companies collect and process student personal and biometric data. The complaint attempts to hold the five companies accountable for their practices by demonstrating how the data collection and processing practices may violate existing law. The complaint concludes by enumerating specific changes that these companies may implement to operate in concert with existing law and general expectations about privacy and equity.

IMPACT & FUTURE PLANS

The work of the EPIC Student Privacy Project will continue to focus on the nexus between student and consumer privacy. The proctoring services investigated by EPIC will continue to be used in classrooms long after the COVID-19 pandemic ends. Therefore, EPIC's consistent and principled advocacy will continue to play a significant role in ensuring that student privacy is not a casualty of the digital age.■

Terms of Service Ratings

ToS;DR Association

Terms of Service Ratings is an initiative that rates and labels website terms and privacy policies to help inform consumers of their rights.

Global

OVERVIEW

Terms of service agreements are standard methods for businesses online to give notice and obtain consent from users that they have permission to legally handle and collect personal data. But some of the most popular online services have terms of service agreements over 10,000 words long with complex legalese that discourage reading and reduce comprehensibility. Terms of Service, Didn't Read (ToS;DR) was founded in 2012 to fix the "biggest lie" on the internet, "I have read and agree to the Terms of Service."

ToS;DR's Terms of Service Ratings rate and label website terms & privacy policies from very good, Class A, to very bad, Class E. The project relies on a group of volunteers to (1) identify specific phrases from terms of service that impact users, (2) simplify the language, (3) label those phrases into green, orange, red, and gray marks, and (4) combine those marks to create a rating. Through the Terms of Serving Ratings, ToS;DR hopes to educate users on their digital rights online.

Through the Terms of Serving Ratings, ToS;DR hopes to educate users on their digital rights.

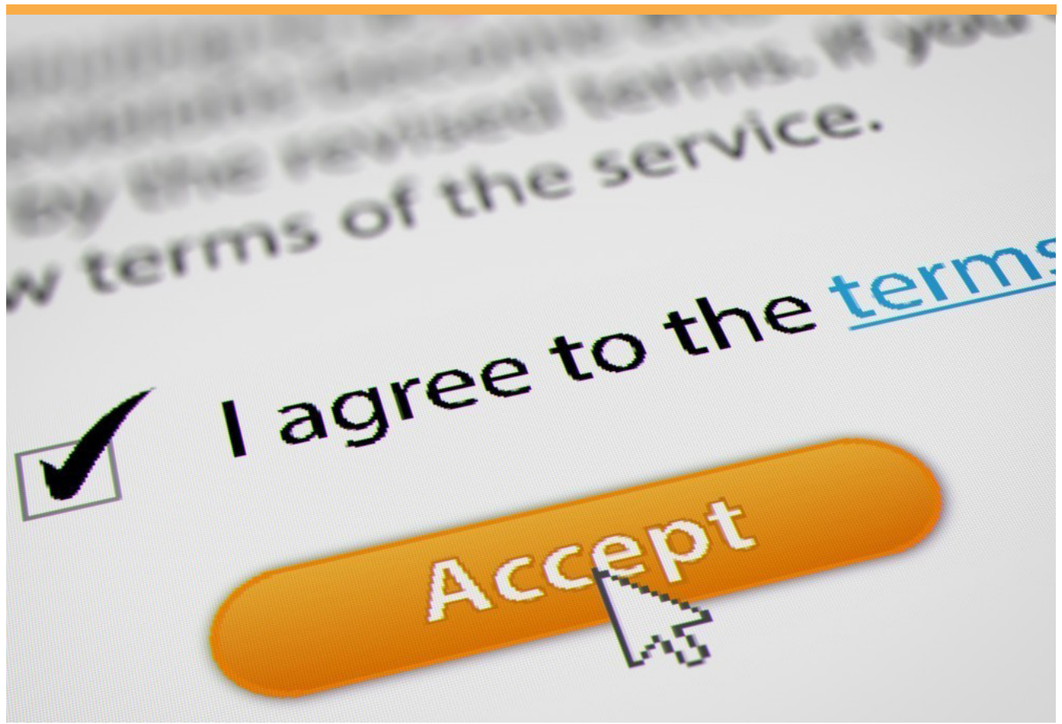
The name ToS;DR is inspired by the internet acronym TL;DR which stands for "Too Long; Didn't Read." TL;DR is frequently used in blogs and emails when summarizing a very long block of text.

THE CHALLENGE

A 2008 study by Carnegie Mellon researchers found that the average internet user encounters almost 1,500 privacy policies a year, each about 2,500 words in length.[1] Further research demonstrates that users don't read terms and might not understand them even if they did.[2] Users may misinterpret their purpose, assuming that the agreements offer a level of data protection, when they do not guarantee user privacy. It is clear that even if users would like to understand the terms before using a service online, they are too long to read and too complicated to understand. To create a more transparent process between businesses and users, terms of service need to be simplified and shared with current and potential users.

ABOUT THE INTERVENTION

ToS;DR established a community where volunteers worldwide can help identify and review terms of services. Terms of Service Ratings are open-source and maintained through their online forum. To start a ratings process, desired terms of service agreements are uploaded to their database. Volunteers review terms using a computer program that automatically searches documents on the Web. Reviewers then highlight specific phrases before attributing a score, ranging from green (good), orange (bad), red (blocker), and gray (neutral). Once an individual



Research demonstrates that most users don't read terms of service and might not understand them even if they did.

terms of service has enough scores, the score is averaged and assigned a grade from A, the best terms of services, to E, a terms of service which raises serious concerns.

For example, phrases from terms indicating that data will be stored even if a user did not interact with the service will be classified with a red (blocker) score. In contrast, phrases indicating that the service provides a complaint mechanism for the improper handling of personal data will be given a green (good) score.

IMPACT & FUTURE PLANS

ToS;DR has published and reviewed the terms of hundreds of services, including Facebook, Google, Reddit, Twitter, and more. The Terms of Service Ratings are reflected on their browser extensions which users can download to be informed of a service's grade before accepting the term. Terms of Service Ratings are also integrated with DuckDuckGo's Privacy Essentials browser extension.

The Terms of Service Ratings project has raised awareness about problems with notice and consent models like terms of services online. The terms are typically offered as "take it or leave it." Declining the terms frequently results in being denied the product or service. The Terms of Service Ratings project has also affected policy change in Europe, with the General Data Protection Directive (GDPR) including a provision that asks the European Commission to come up with standards to simplify the information in privacy policies.

Because terms of services are updated frequently, future plans include creating an automated document annotator and reviewer. ToS;DR is also looking to grow the community of reviewers and raise further awareness about problems associated with terms of services ■

1 Aleecia M McDonald and Lorrie Faith Cranor, "The Cost of Reading Privacy Policies," *I/S: A Journal of Law and Policy for the Information Society*, no. Privacy Year in Review (n.d.): 22.
2 Craig Wigginton, Mike Curran, and Terrence Karner, "2017 Global Mobile Consumer Survey: US Edition" (Deloitte, 2017).

Upsolve App

Upsolve

The Upsolve App is a tool that helps individuals file for bankruptcy for free.

New York, NY

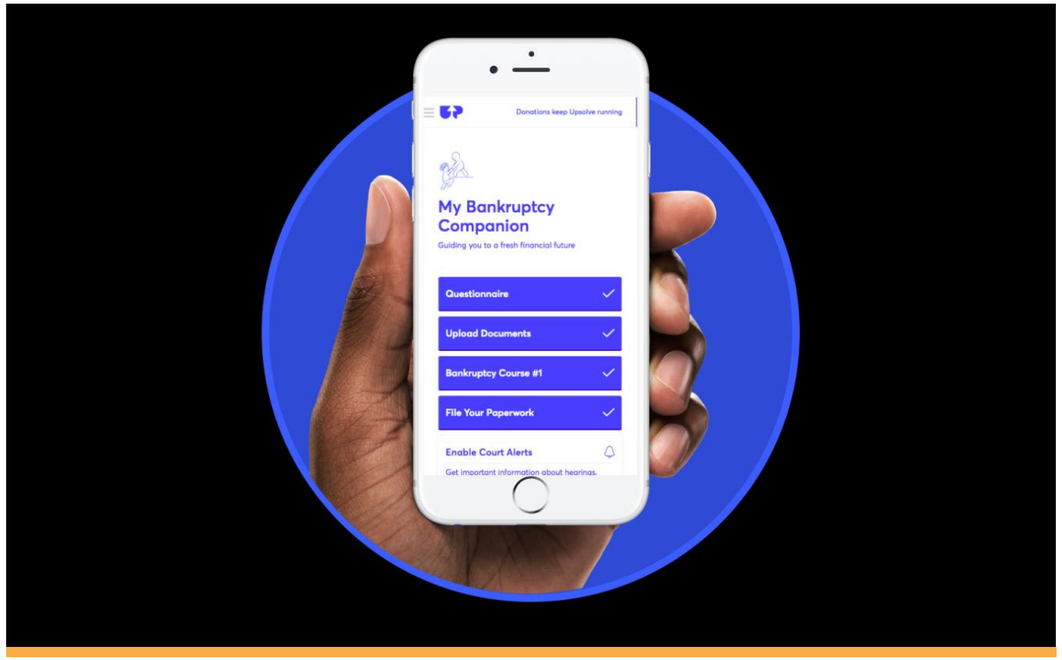
In 2020 The Upsolve App was named one of TIME magazine's 100 Best Inventions that make the world a better place.

OVERVIEW

The Upsolve App helps low-income families and individuals who cannot afford lawyers file for bankruptcy for free. In many cases, filing for Chapter 7 bankruptcy can be an efficient way to eliminate debt. Paradoxically, filing for bankruptcy is expensive to do - it can cost up to $1,300 in fees, a significant burden for people who are already struggling financially. In addition to the exorbitant fees, the filing process can be confusing and complicated to navigate alone. Resulting filing errors can be costly to fix. Upsolve simplifies and accelerates the bankruptcy filing process by walking people through the necessary steps with an online web app that is simple and easy to use. The app allows the user to easily generate and complete the bankruptcy forms required to alleviate debt. Upsolve, a non-profit, is not predatory and simply helps users successfully navigate a time in their lives that can feel challenging, frustrating, and isolating. Upsolve is working towards becoming the defining brand in America for low-income families in financial distress who need access to their legal and financial rights.

THE CHALLENGE

Millions of low-income families and individuals are trapped in debt due to medical issues, job loss, or other crises. The COVID-19 pandemic has exacerbated the financial challenges that many families were already facing. According to consumer credit reporting company Experian, in 2020, the U.S. consumer debt balance increased by 6% -- the highest annual growth recorded in over 10 years.[1] In many cases, this debt has pushed people over the edge to the point where filing for bankruptcy is one of the few steps that they can take to help them regain control of their financial lives. Filing for bankruptcy allows families and individuals to wipe out certain kinds of debt. For some, filing for bankruptcy can even prevent evictions, foreclosures, or repossessions. Filing can help people restart with a clean slate and provide necessary financial relief.



The Upsolve App helps low-income families file bankruptcy for free.

ABOUT THE INTERVENTION

The Upsolve App, also known as the "Chapter 7 bankruptcy tool," walks people through how to complete the necessary forms to file for bankruptcy. Users can easily navigate through the necessary steps to file, get access to frequently asked questions, and are invited to join a community of people who discuss bankruptcy-related topics and support one another with tips and advice. The app is designed to spread positivity and encourage users to regain control of their financials. For example, the app has cheerful graphics that inform users that famous individuals like Walt Disney and Henry Ford also filed for bankruptcy at some point in their lives, and then went on to achieve great things. Many Upsolve users have become trapped in debt because of various personal and family crises including unexpected medical expenses or natural disasters that affected family businesses. These same users have leveraged the Upsolve App on their path back to financial security, new careers, and improved livelihoods.

IMPACT & FUTURE PLANS

Nearly 7,000 people have confirmed filing for bankruptcy with the Upsolve app, helping to alleviate over $300M in debt. In addition, Upsolve has educated over 2.5 million people about their legal and financial rights through free bankruptcy-related education on the Upsolve website. As of April 2021, Upsolve has 150,000+ members.

In 2020 The Upsolve App was named one of TIME magazine's 100 Best Inventions that make the world a better place ■

1 Stefan Lembo Stolba, "Average U.S. Consumer Debt Reaches New Record in 2020," Experian, April 6, 2021, https://www.experian.com/blogs/ask-experian/research/consumer-debt-study/.

About the Technology and Public Purpose (TAPP) Project

The arc of innovative progress has reached an inflection point.

Technological change has brought immeasurable benefits to billions through improved health, productivity, and convenience. Yet as recent events have shown, unless we actively manage their risks to society, new technologies may also bring unforeseen destructive consequences. Making technological change positive for all is the critical challenge of our time. We ourselves – not only the logic of discovery and market forces – must manage it. To create a future where technology serves humanity as a whole, we need a new approach.

To this end, Harvard Kennedy School's Belfer Center for Science and International Affairs launched the Technology and Public Purpose (TAPP) Project in 2018. Led by Belfer Center Director, MIT Innovation Fellow, and former Secretary of Defense Ash Carter, the TAPP Project works to ensure that emerging technologies are developed and managed in ways that serve the overall public good.

We aim to create a set of conditions that leaven today's technological change across three domains: digital, biotech, and the future of work. TAPP leverages a network of experts from Harvard University, MIT, and the Greater Boston Area, along with leaders in technology, government, business, and civil society to work on the following priorities

Training & Mentorship – Training today's practitioners and tomorrow's leaders in the responsible development and management of new technologies.

Convening Stakeholders – Convening leaders in tech, policy, academia, and civil society to develop solutions to the societal dilemmas of emerging technologies.

Publishing Leading Edge Research – Conducting world-class research on high-risk technologies and frameworks for effective development and governance ■

For more information, visit:
www.belfercenter.org/TAPP

HARVARD Kennedy School
BELFER CENTER
for Science and International Affairs
Technology & Public Purpose Project