# Social and Ethical Responsibilities of Computing (SERC):

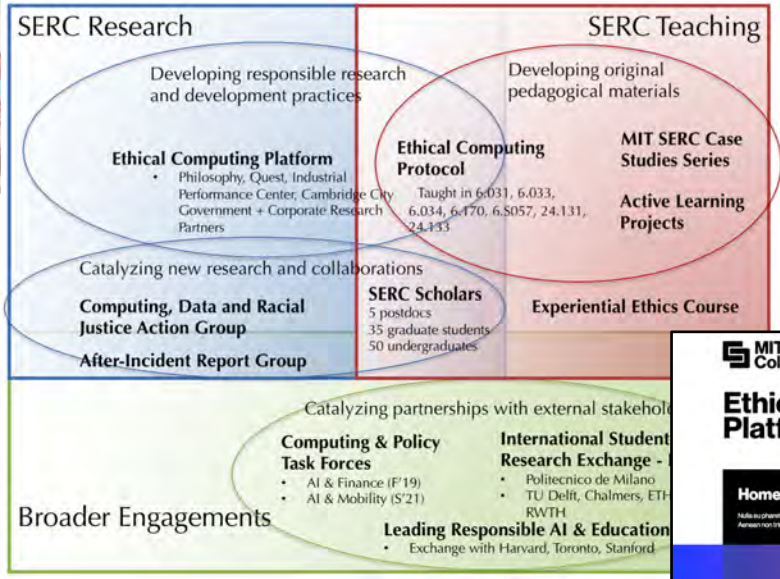
# New Activities at MIT

**David Kaiser**

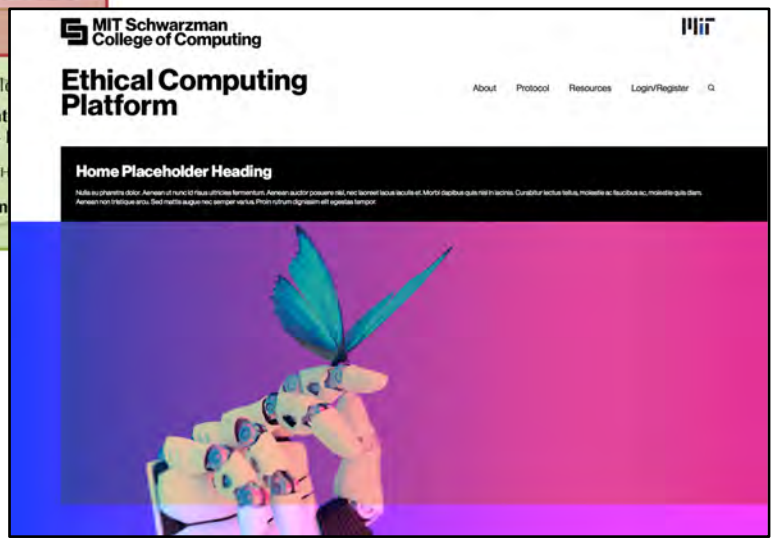
November 2021

# AI in the World

## SERC at MIT




## New Teams, New Tools



# Opportunities

HEALTH TECH

STAT+

## When AI is the opposite of sinister: An MIT researcher is held up as model of how algorithms can benefit humanity

By Rebecca Robbins Sept. 23, 2020

MIT artificial intelligence researcher Regina Barzilay is the in
Association for the Advancement of Artificial Intelligence.
MIT CSAIL

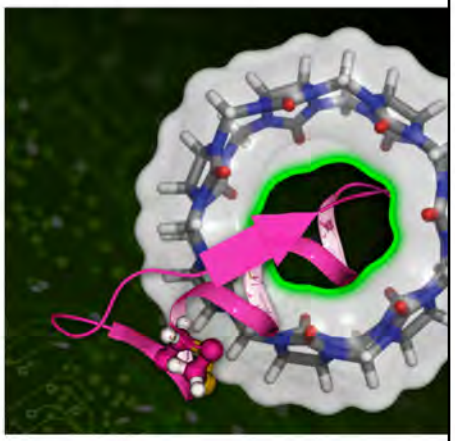
MIT News
ON CAMPUS AND AROUND THE WORLD

SUBSCRIBE

## Faster drug discovery through machine learning

New technique speeds up calculations of drug molecules' binding affinity to proteins.

Daniel Ackerman | MIT News Office
March 15, 2021



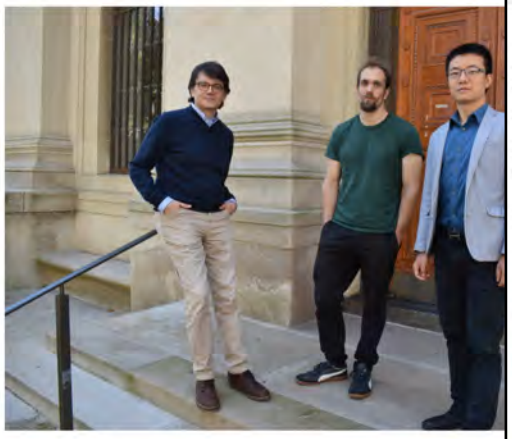
MIT News
ON CAMPUS AND AROUND THE WORLD

SUBSCRIBE

## Taming the data deluge

A National Science Foundation-funded team will use artificial intelligence to speed up discoveries in physics, astronomy, and neuroscience.

Sandi Miller | Department of Physics
October 29, 2021



MIT News
ON CAMPUS AND AROUND THE WORLD

## At the crossroads of language, technology, and empathy

With a double major in linguistics and computer science, senior Rujul Gandhi works to surmount language and cultural barriers, globally and on campus.

Alli Armijo | MIT News correspondent
October 13, 2021

# Challenges

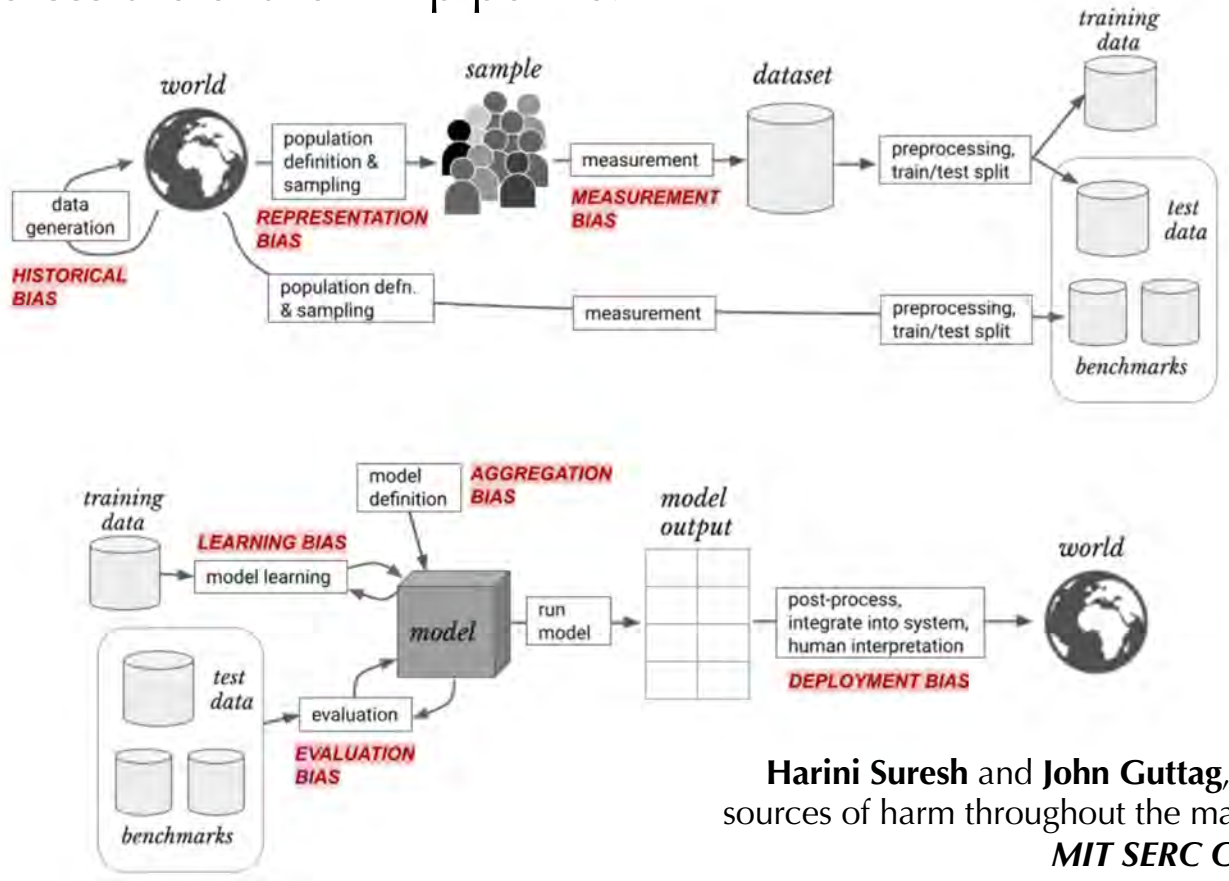
The interface of **artificial intelligence** (AI) + **machine-learning** (ML) techniques with ***people*** — both individuals and groups — presents **special challenges**.



These challenges are **exacerbated** when AI + ML techniques move ***beyond research settings*** into ***real-world situations***.

# Challenges

The interface of **artificial intelligence** (AI) + **machine-learning** (ML) techniques with ***people*** — both individuals and groups — presents **special challenges**. These are often described as "**algorithmic bias**," though there exists a ***whole range*** of challenges, across the entire ML pipeline.

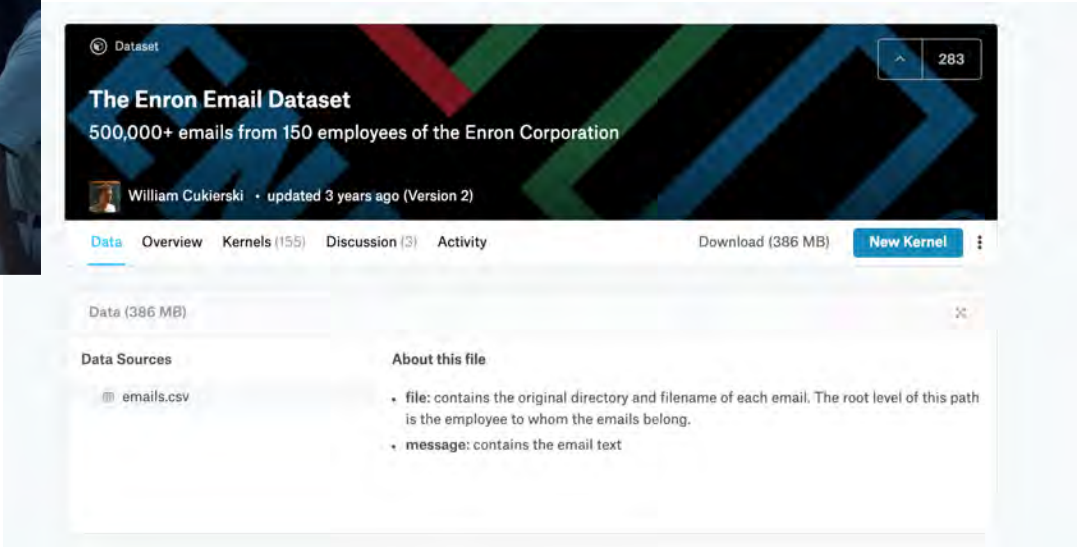



**Harini Suresh** and **John Guttag**, "Understanding potential sources of harm throughout the machine learning life cycle," ***MIT SERC Case Studies*** Issue 2 (2021)

These challenges are **exacerbated** when AI + ML techniques move ***beyond research settings*** into ***real-world situations***.

# Challenges

The interface of **artificial intelligence** (AI) + **machine-learning** (ML) techniques with ***people*** — both individuals and groups — presents **special challenges**. These are often described as "**algorithmic bias**," though there exists a ***whole range*** of challenges, across the entire ML pipeline.


These challenges are ***exacerbated*** when AI + ML techniques move ***beyond research settings*** into ***real-world situations***.

# Facial Recognition Technologies


The **2019 NIST report** (third in a series) analyzed 189 algorithms from 99 (mostly commercial) developers. They assessed performance both for **1:1 matching** (for verification of a known individual) and for **1:*n* matching** (for attempted matches within a whole dataset). [Main report: 75pp, plus 1200pp technical appendices.]

# Facial Recognition Technologies

NISTIR 8280

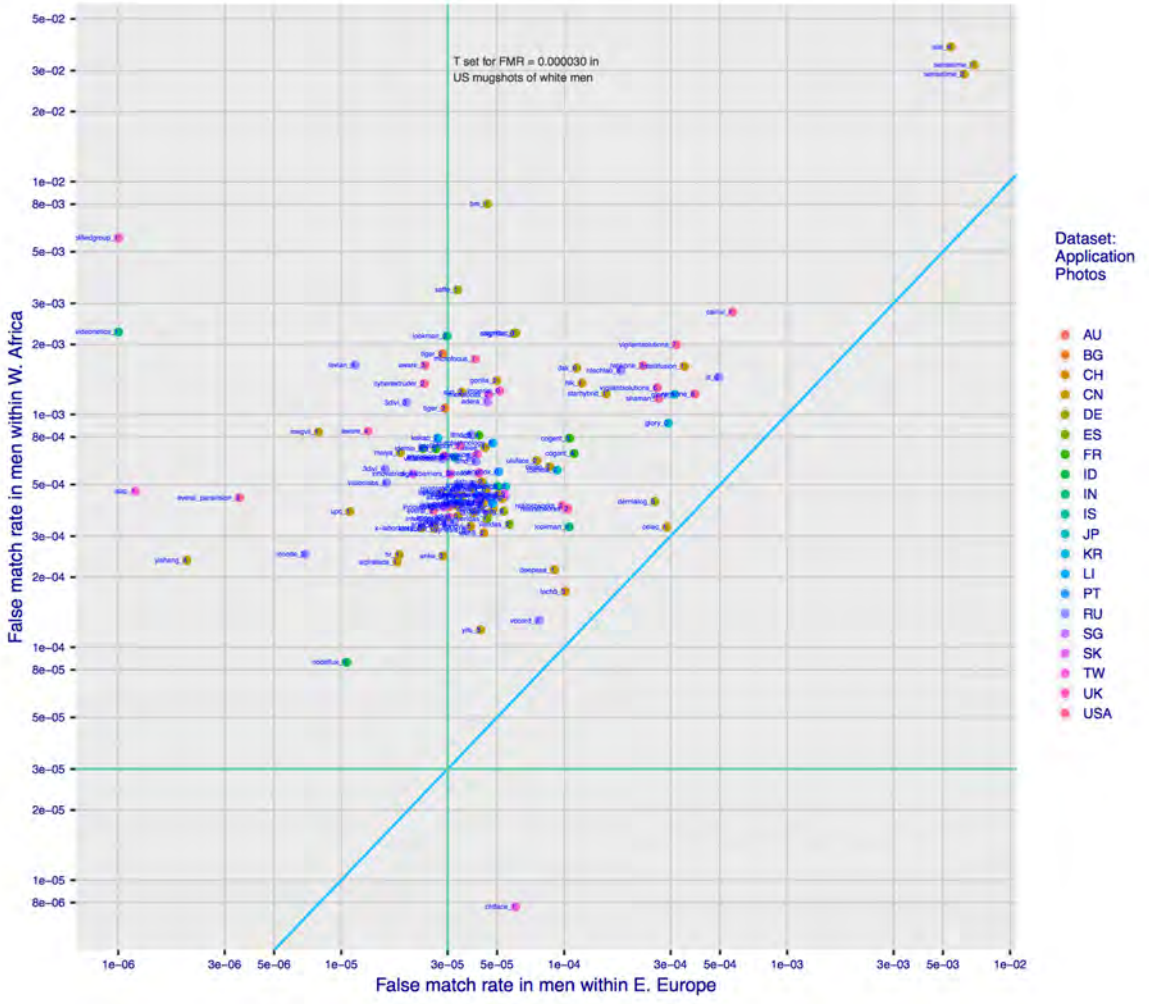
**Face Recognition Vendor Test (FRVT)**
**Part 3: Demographic Effects**

Patrick Grother
Mei Ngan
Kayee Hanaoka

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8280

NIST
National Institute of Standards and Technology
U.S. Department of Commerce

The **2019 NIST report** (third in a series) analyzed 189 algorithms from 99 (mostly commercial) developers. They assessed performance both for **1:1 matching** (for verification of a known individual) and for **1:*n* matching** (for attempted matches within a whole dataset). [Main report: 75pp, plus 1200pp technical appendices.]



# Facial Recognition Technologies

NISTIR 8280

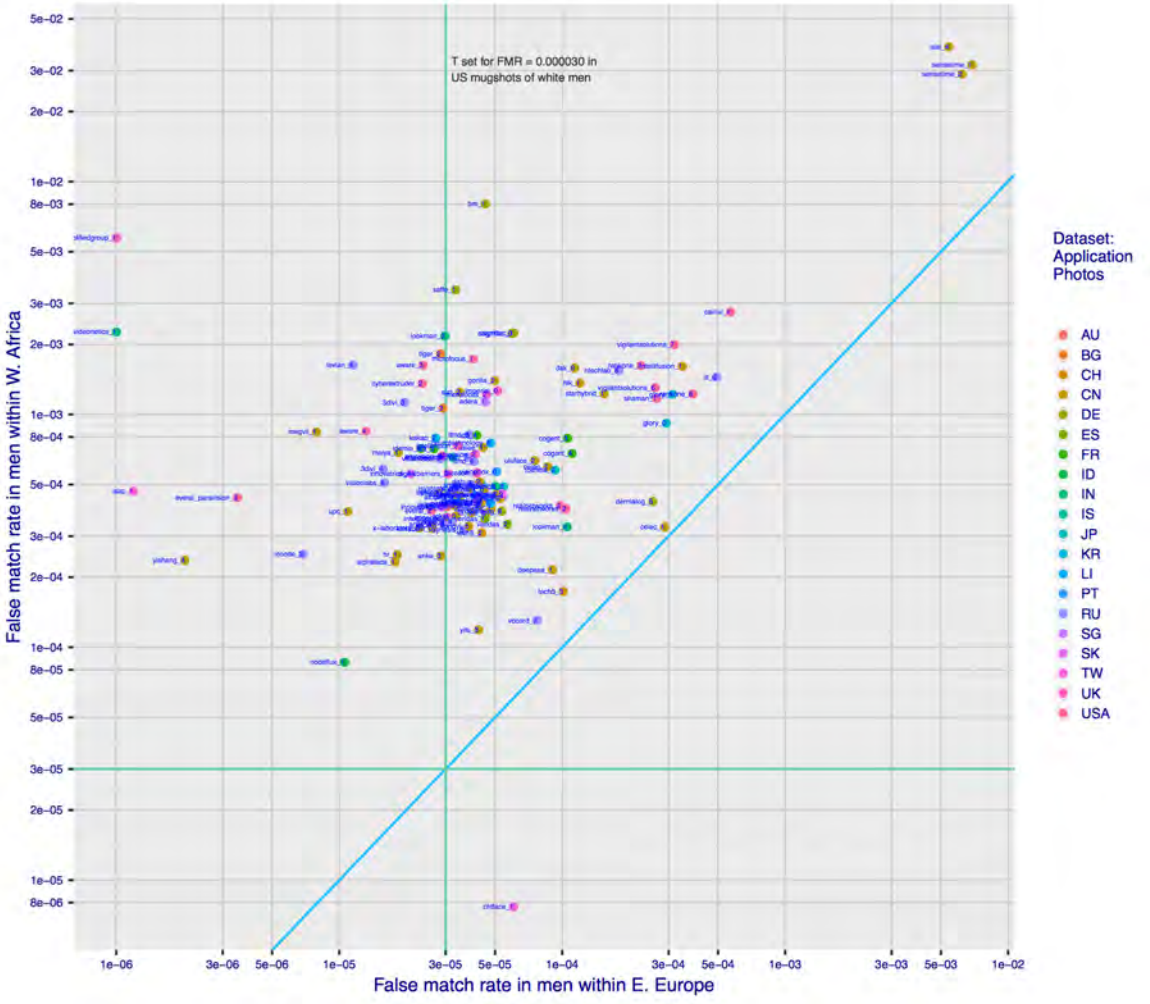
**Face Recognition Vendor Test (FRVT)**
**Part 3: Demographic Effects**

Patrick Grother
Mei Ngan
Kayee Hanaoka

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8280

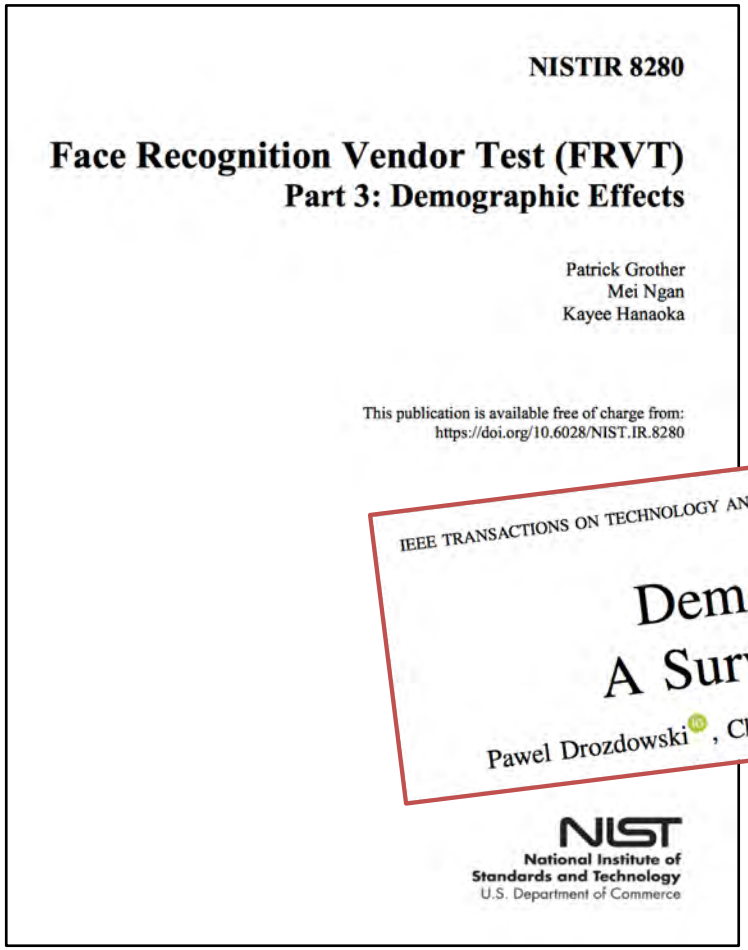
NIST
National Institute of Standards and Technology
U.S. Department of Commerce

The **2019 NIST report** (third in a series) analyzed 189 algorithms from 99 (mostly commercial) developers. They assessed performance both for **1:1 matching** (for verification of a known individual) and for **1:*n* matching** (for attempted matches within a whole dataset). [Main report: 75pp, plus 1200pp technical appendices.]




***Summary***: False positive rates highest for people from Africa and East Asia, lowest for people from Eastern Europe (large effect, ~$10^2x$). Across geographical sets, false positives higher for women than men (~ 2 – 5$x$).

# Facial Recognition Technologies



NISTIR 8280

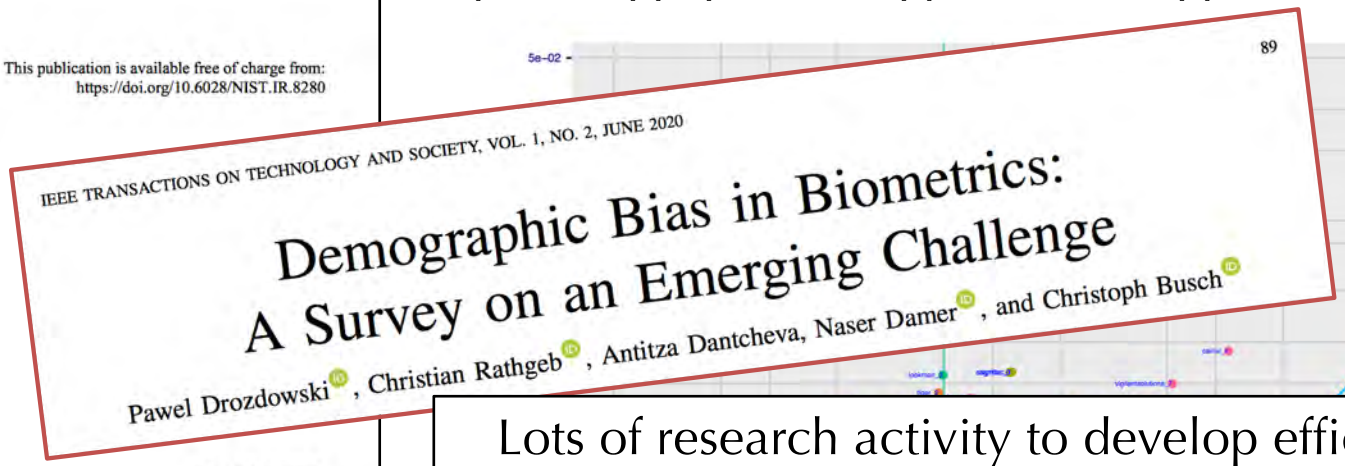
**Face Recognition Vendor Test (FRVT)**
**Part 3: Demographic Effects**

Patrick Grother
Mei Ngan
Kayee Hanaoka

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8280

NIST
National Institute of Standards and Technology
U.S. Department of Commerce

The **2019 NIST report** (third in a series) analyzed 189 algorithms from 99 (mostly commercial) developers. They assessed performance both for **1:1 matching** (for verification of a known individual) and for **1:*n* matching** (for attempted matches within a whole dataset). [Main report: 75pp, plus 1200pp technical appendices.]





IEEE TRANSACTIONS ON TECHNOLOGY AND SOCIETY, VOL. 1, NO. 2, JUNE 2020

Demographic Bias in Biometrics:
A Survey on an Emerging Challenge

Pawel Drozdowski, Christian Rathgeb, Antitza Dantcheva, Naser Damer, and Christoph Busch

Lots of research activity to develop efficient mitigation strategies *in laboratory settings*...




***Summary***: False positive rates highest for people from Africa and East Asia, lowest for people from Eastern Europe (large effect, ~$10^2x$). Across geographical sets, false positives higher for women than men (~ 2 – 5$x$).

# Facial Recognition Technologies

The New York Times

## *Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match*

A New Jersey man was accused of shoplifting and trying to hit an officer with a car. He is the third known Black man to be wrongfully arrested based on face recognition.

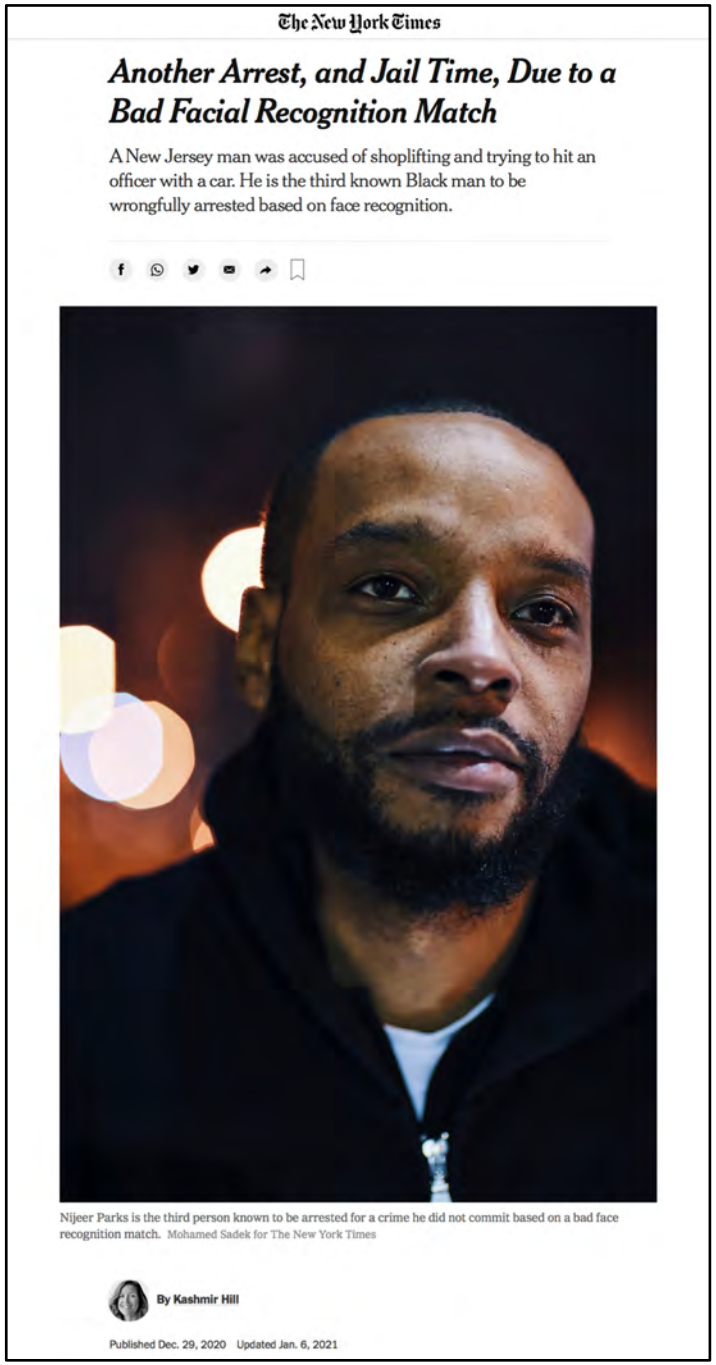
Nijeer Parks is the third person known to be arrested for a crime he did not commit based on a bad face recognition match. Mohamed Sadek for The New York Times

By Kashmir Hill

Published Dec. 29, 2020 Updated Jan. 6, 2021

The US has *thousands* of distinct law-enforcement jurisdictions. Commercial facial-recognition technologies are ***already being used*** across the country, subject to ***no regulation, standardization, or oversight***.

# Facial Recognition Technologies



The New York Times

**Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match**

A New Jersey man was accused of shoplifting and trying to hit an officer with a car. He is the third known Black man to be wrongfully arrested based on face recognition.

Nijeer Parks is the third person known to be arrested for a crime he did not commit based on a bad face recognition match. Mohamed Sadek for The New York Times

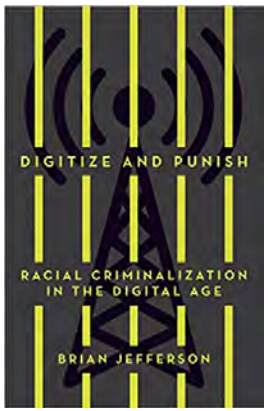
By Kashmir Hill

Published Dec. 29, 2020 Updated Jan. 6, 2021

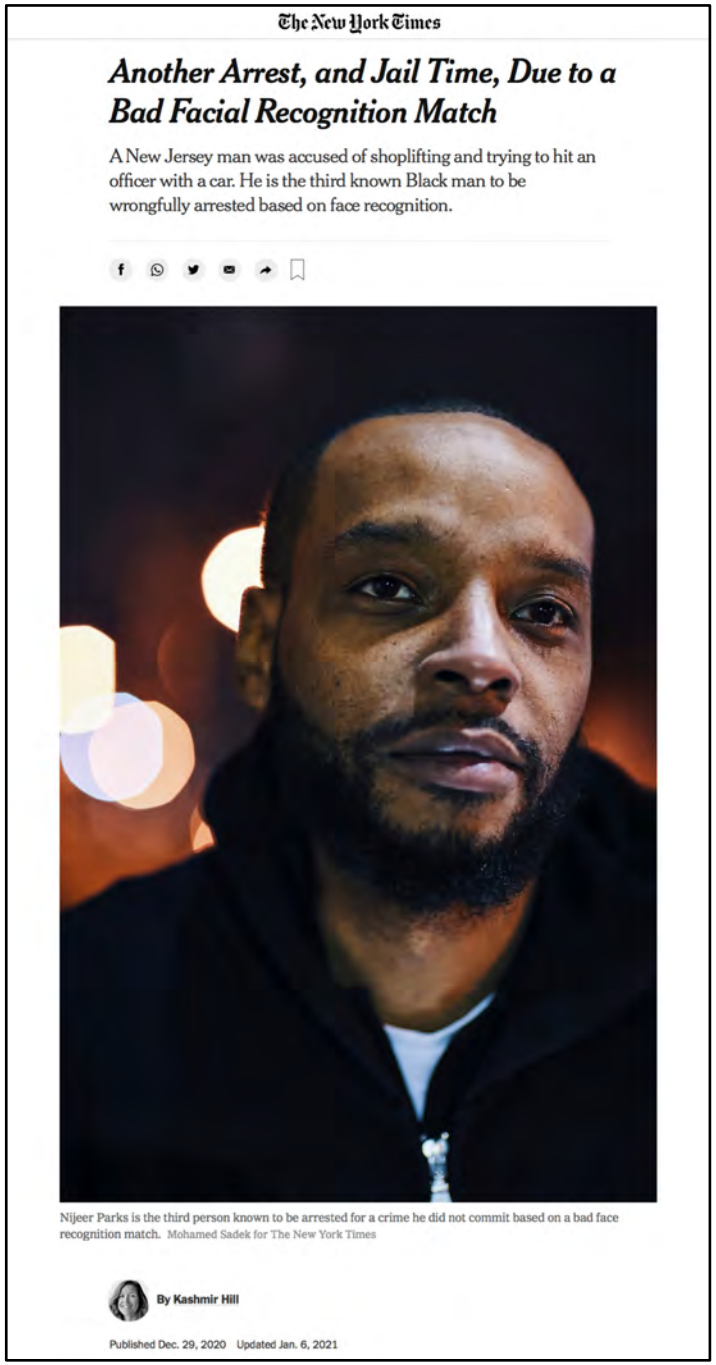
The US has *thousands* of distinct law-enforcement jurisdictions. Commercial facial-recognition technologies are ***already being used*** across the country, subject to ***no regulation, standardization, or oversight***.

Several reported cases of Black men being wrongfully arrested due to incorrect FRT matches indicate a ***combination of failures***: inadequate technical calibrations ***plus*** human failures to follow appropriate procedures.

More generally: new technologies are deployed within ***existing institutional frameworks***.


# Facial Recognition Technologies



The New York Times

***Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match***

A New Jersey man was accused of shoplifting and trying to hit an officer with a car. He is the third known Black man to be wrongfully arrested based on face recognition.

Nijeer Parks is the third person known to be arrested for a crime he did not commit based on a bad face recognition match. Mohamed Sadek for The New York Times

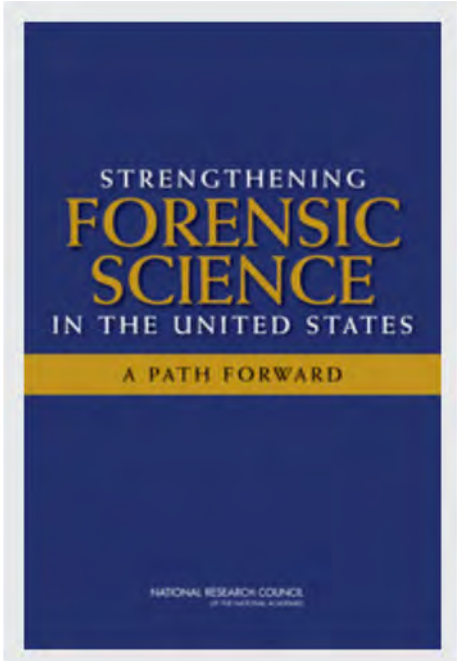
By Kashmir Hill

Published Dec. 29, 2020 Updated Jan. 6, 2021

The US has *thousands* of distinct law-enforcement jurisdictions. Commercial facial-recognition technologies are ***already being used*** across the country, subject to ***no regulation, standardization, or oversight***.

Several reported cases of Black men being wrongfully arrested due to incorrect FRT matches indicate a ***combination of failures***: inadequate technical calibrations ***plus*** human failures to follow appropriate procedures.

Contrast this with other examples of ***forensic science*** within the US, such as ***fingerprinting***: far from perfect, but subject to ***expert review, training,*** and ***standardization***.



# Facial Recognition Technologies

The New York Times

## *Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match*

A New Jersey man was accused of shoplifting and trying to hit an officer with a car. He is the third known Black man to be wrongfully arrested based on face recognition.

By Kashmir Hill

Published Dec. 29, 2020 Updated Jan. 6, 2021

***One billion surveillance*** cameras are now in place within 50 countries. Within the US alone, facial images of ***half the adult US population*** are already included in databases accessible to law enforement.

The US has *thousands* of distinct law-enforcement jurisdictions. Commercial facial-recognition technologies are ***already being used*** across the country, subject to ***no regulation, standardization, or oversight***.

Several reported cases of Black men being wrongfully arrested due to incorrect FRT matches indicate a ***combination of failures***: inadequate technical calibrations ***plus*** human failures to follow appropriate procedures.

Forbes

Jun 4, 2019, 06:26pm EDT | 2,146 views

## Government Watchdog Questions FBI On Its 640-Million-Photo Facial Recognition Database

Monica Melton Former Staff
Innovation

Follow

Officials testify a House Oversight and Reform committee hearing on facial recognition technology in ... [+] JACQUELYN MARTIN/ASSOCIATED PRESS

## *As China Tracked Muslims, Alibaba Showed Customers How They Could, Too*

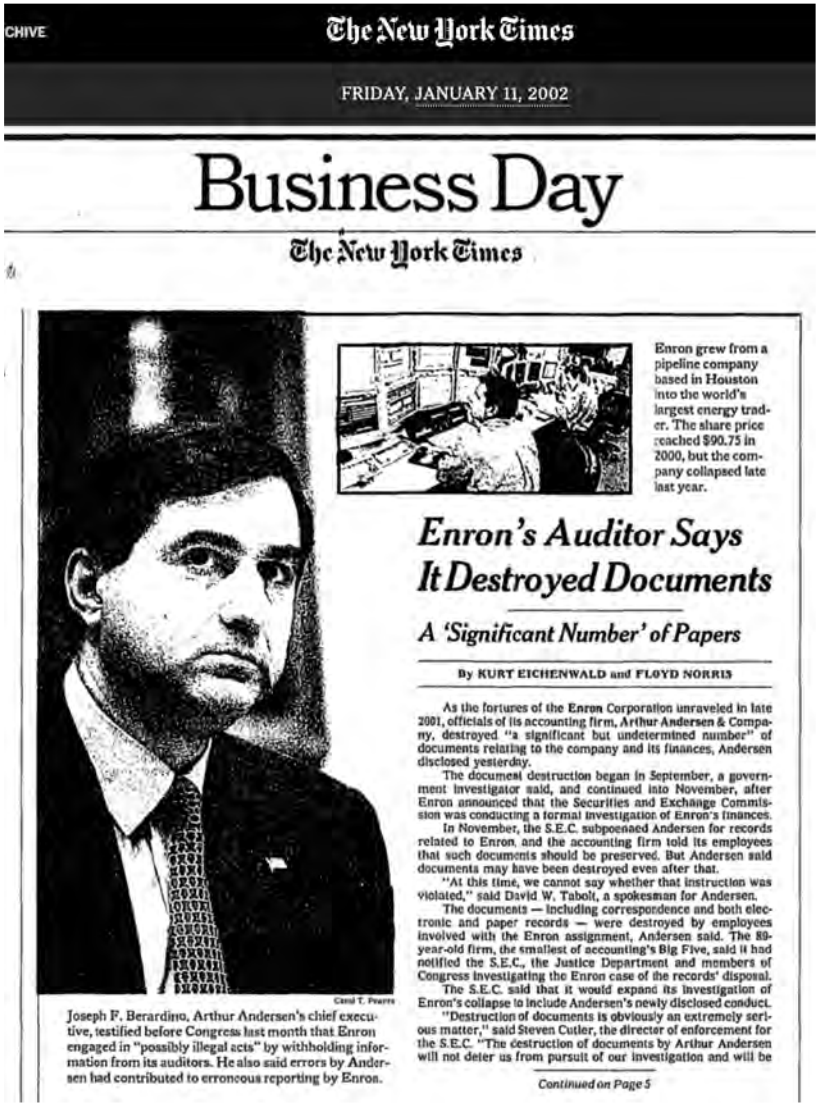
The website for the tech titan's cloud business described facial recognition software that could detect members of a minority group whose persecution has drawn international condemnation.


Alibaba, the leading provider of cloud computing services in China, could be drawn into the global controversy over China's treatment of Uighurs. Aly Song/Reuters

By Raymond Zhong

Published Dec. 16, 2020 Updated Jan. 20, 2021

# Natural-Language Processing



CHIVE

The New York Times

FRIDAY, JANUARY 11, 2002

Business Day

The New York Times

Enron grew from a pipeline company based in Houston into the world's largest energy trader. The share price reached $90.75 in 2000, but the company collapsed late last year.

## Enron's Auditor Says It Destroyed Documents

### A 'Significant Number' of Papers

By KURT EICHENWALD and FLOYD NORRIS

As the fortunes of the Enron Corporation unraveled in late 2001, officials of its accounting firm, Arthur Andersen & Company, destroyed "a significant but undetermined number" of documents relating to the company and its finances, Andersen disclosed yesterday.

The document destruction began in September, a government investigator said, and continued into November, after Enron announced that the Securities and Exchange Commission was conducting a formal investigation of Enron's finances.

In November, the S.E.C. subpoenaed Andersen for records related to Enron, and the accounting firm told its employees that such documents should be preserved. But Andersen said documents may have been destroyed even after that.

"At this time, we cannot say whether that instruction was violated," said David W. Tabolt, a spokesman for Andersen.

The documents — including correspondence and both electronic and paper records — were destroyed by employees involved with the Enron assignment, Andersen said. The 89-year-old firm, the smallest of accounting's Big Five, said it had notified the S.E.C., the Justice Department and members of Congress investigating the Enron case of the records' disposal.

The S.E.C. said that it would expand its investigation of Enron's collapse to include Andersen's newly disclosed conduct.

"Destruction of documents is obviously an extremely serious matter," said Steven Cutler, the director of enforcement for the S.E.C. "The destruction of documents by Arthur Andersen will not deter us from pursuit of our investigation and will be

Continued on Page 5

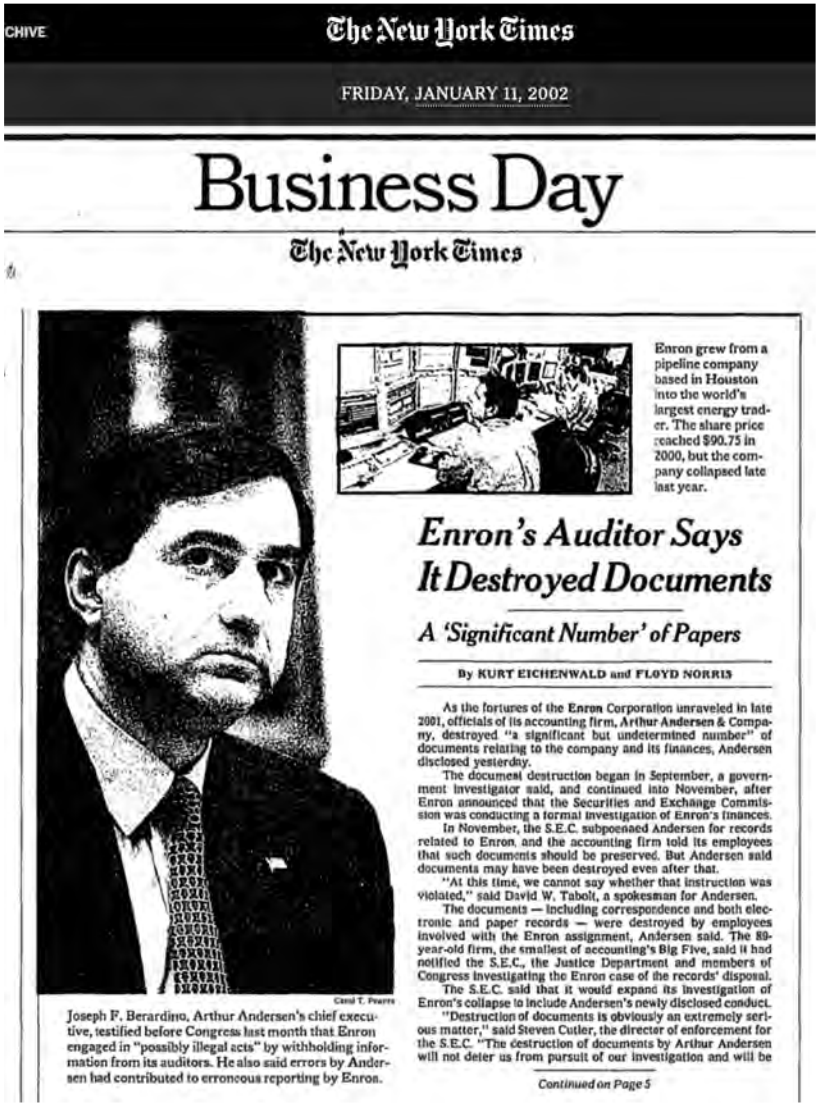
Carol T. Powers

Joseph F. Berardino, Arthur Andersen's chief executive, testified before Congress last month that Enron engaged in "possibly illegal acts" by withholding information from its auditors. He also said errors by Andersen had contributed to erroneous reporting by Enron.

To date, one of the largest and most-used ***corpora of email*** — critical to algorithmic natural-language processing — comes from employees of ***Enron***, the large Texas-based energy company that declared bankruptcy in December 2001.

Ultimately more than ***20 executives pleaded guilty*** or were ***convicted***, including on multiple felony fraud charges.

# Natural-Language Processing



CHIVE

The New York Times

FRIDAY, JANUARY 11, 2002

Business Day

The New York Times

Enron grew from a pipeline company based in Houston into the world's largest energy trader. The share price reached $90.75 in 2000, but the company collapsed late last year.

## Enron's Auditor Says It Destroyed Documents

### A 'Significant Number' of Papers

By KURT EICHENWALD and FLOYD NORRIS

As the fortunes of the Enron Corporation unraveled in late 2001, officials of its accounting firm, Arthur Andersen & Company, destroyed "a significant but undetermined number" of documents relating to the company and its finances, Andersen disclosed yesterday.

The document destruction began in September, a government investigator said, and continued into November, after Enron announced that the Securities and Exchange Commission was conducting a formal investigation of Enron's finances.

In November, the S.E.C. subpoenaed Andersen for records related to Enron, and the accounting firm told its employees that such documents should be preserved. But Andersen said documents may have been destroyed even after that.

"At this time, we cannot say whether that instruction was violated," said David W. Tabolt, a spokesman for Andersen.

The documents — including correspondence and both electronic and paper records — were destroyed by employees involved with the Enron assignment, Andersen said. The 89-year-old firm, the smallest of accounting's Big Five, said it had notified the S.E.C., the Justice Department and members of Congress investigating the Enron case of the records' disposal.

The S.E.C. said that it would expand its investigation of Enron's collapse to include Andersen's newly disclosed conduct.

"Destruction of documents is obviously an extremely serious matter," said Steven Cutler, the director of enforcement for the S.E.C. "The destruction of documents by Arthur Andersen will not deter us from pursuit of our investigation and will be

Continued on Page 5

Carol T. Powers

Joseph F. Berardino, Arthur Andersen's chief executive, testified before Congress last month that Enron engaged in "possibly illegal acts" by withholding information from its auditors. He also said errors by Andersen had contributed to erroneous reporting by Enron.

To date, one of the largest and most-used ***corpora of email*** — critical to algorithmic natural-language processing — comes from employees of ***Enron***, the large Texas-based energy company that declared bankruptcy in December 2001.

Ultimately more than ***20 executives pleaded guilty*** or were ***convicted***, including on multiple felony fraud charges.

In 2003, the U.S. Federal Energy Regulatory Commission released ***1.6 million emails*** sent to or from ***158 Enron senior executives*** between 2000 – 2002. After minimal processing, the emails were simply made publicly available on a website. Employees were given a ***limited opt-out*** period.

(Recall 2003: Mark Zuckerberg was still an undergraduate ...)

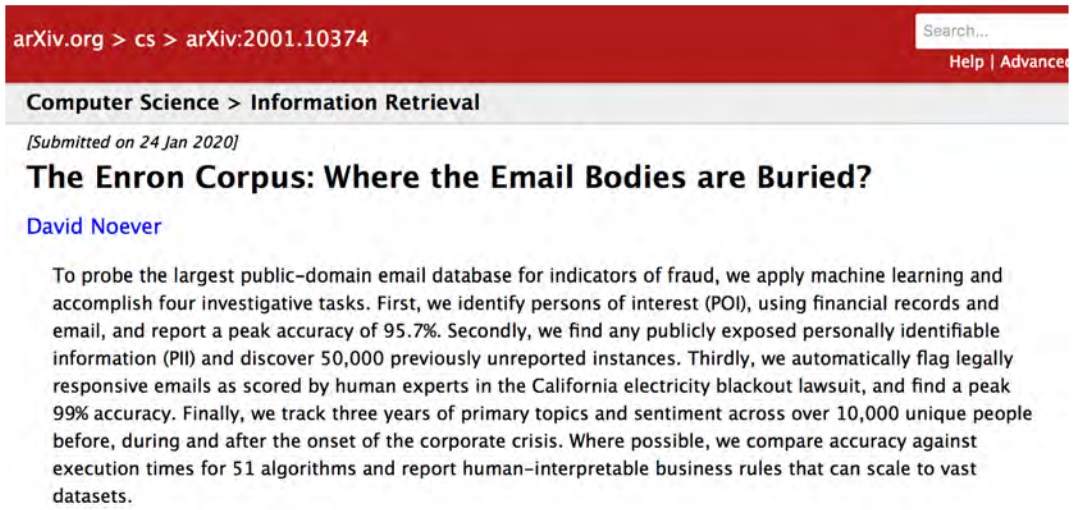
# Natural-Language Processing


More than ***20k academic studies*** have been published that make use of the Enron email corpus (~1k per year).

# Natural-Language Processing


More than ***20k academic studies*** have been published that make use of the Enron email corpus (~1k per year).

Although no legal actions were ever brought against 99.83% of Enron email users, researchers are ***still*** finding $O(10^4)$ examples of ***sensitive personally identifiable information*** within the publicly available corpus, including Social Security Numbers, credit card numbers, birthdates, bank account numbers... Not to mention multiple extramarital affairs and other embarrassing episodes.



# Natural-Language Processing


More than ***20k academic studies*** have been published that make use of the Enron email corpus (~1k per year).

Beyond the privacy concerns, the ubiquitous corpus continues to be used as ***training data*** for various ***natural-language-processing algorithms***.

"If you think there might be **significant biases** embedded in emails sent among employees of a Texas oil-and-gas company that collapsed under federal investigation for **fraud** stemming from **systemic, institutionalized unethical culture**, you would be right. The Enron emails are simply **not representative**—not geographically, not socioeconomically, not even in terms of race or gender. Indeed, researchers have used the Enron emails *specifically* to analyze gender bias and power dynamics. ***And yet the Enron emails remain a go-to dataset for training AI systems***."

The Enron email corpus is a canonical example of "***biased, low-friction data***."

***Amanda Levendowski***, "How copyright law can fix artificial intelligence's implicit bias problem," ***Washington Law Review*** 93 (2018): 579-630.

# Beyond "Algorithmic Bias"

Summer 2021 +

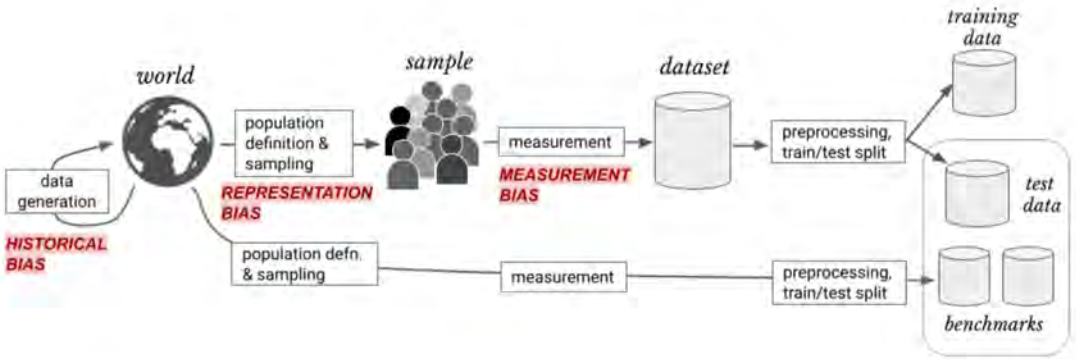
Published on Aug 10, 2021

## Understanding Potential Sources of Harm throughout the Machine Learning Life Cycle

As machine learning (ML) increasingly affects people and society, awareness of its potential unwanted consequences has also grown. To anticipate, prevent, and mitigate undesirable downstream consequences, it is critical that we understand when and how harm might be introduced...


**Harini Suresh**
Department of Electrical Engineering and Computer Science, and Computer Science and Artificial Intelligence Laboratory, MIT

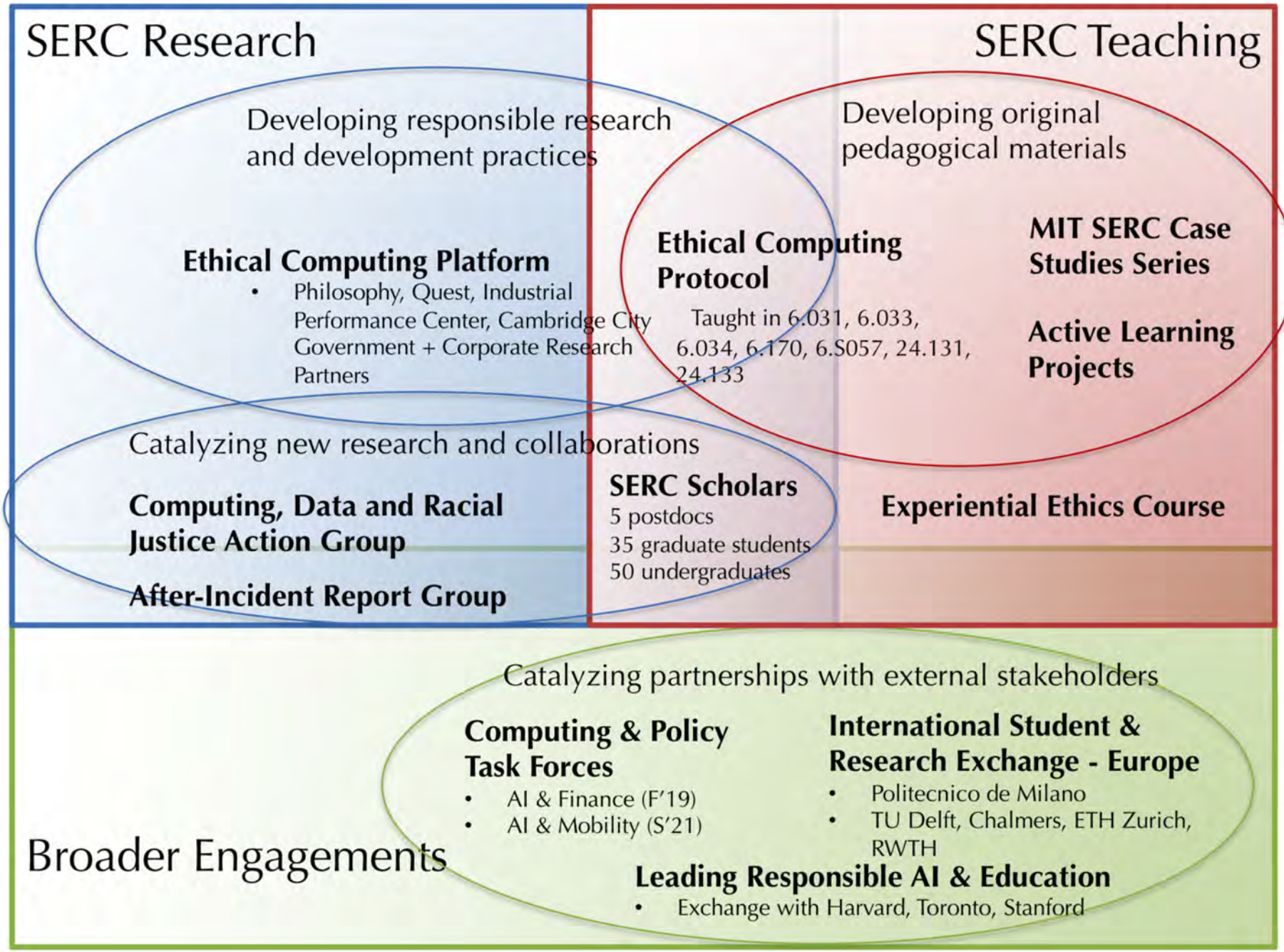
**John Guttag**
Department of Electrical Engineering and Computer Science, and Computer Science and Artificial Intelligence Laboratory, MIT


policing and sentencing; healthcare; real estate and finance; hiring …

1. **Historical bias** arises when there is a misalignment between world as it is and the values or objectives to be encoded and propagated in a model. It is a normative concern with the state of the world, and exists even given perfect sampling and feature selection.
2. **Representation bias** arises while defining and sampling a development population. It occurs when the development population under-represents, and subsequently fails to generalize well, for some part of the use population.
3. **Measurement Bias** arises when choosing and measuring features and labels to use; these are often proxies for the desired quantities. The chosen set of features and labels may leave out important factors or introduce group- or input-dependent noise that leads to differential performance.
4. **Aggregation bias** arises during model construction, when distinct populations are inappropriately combined. In many applications, the population of interest is heterogeneous and a single model is unlikely to suit all subgroups.
5. **Evaluation bias** occurs during model iteration and evaluation. It can arise when the testing or external benchmark populations do not equally represent the various parts of the use population. Evaluation bias can also arise from the use of performance metrics that are not appropriate for the way in which the model will be used.
6. **Deployment Bias** occurs after model deployment, when a system is used or interpreted in inappppropriate ways.

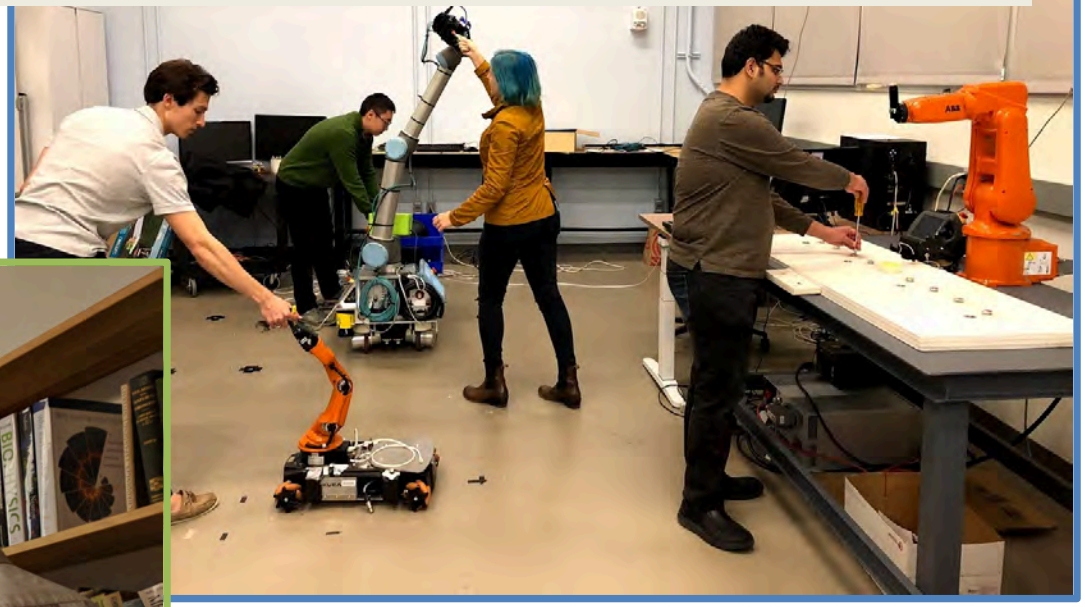
# Social and Ethical Responsibilities of Computing at MIT



# SERC Leadership Team

**Julie Shah**

Professor, Department of Aeronautics and Astronautics
Director, Interactive Robotics Group, CSAIL
Associate Dean, SERC



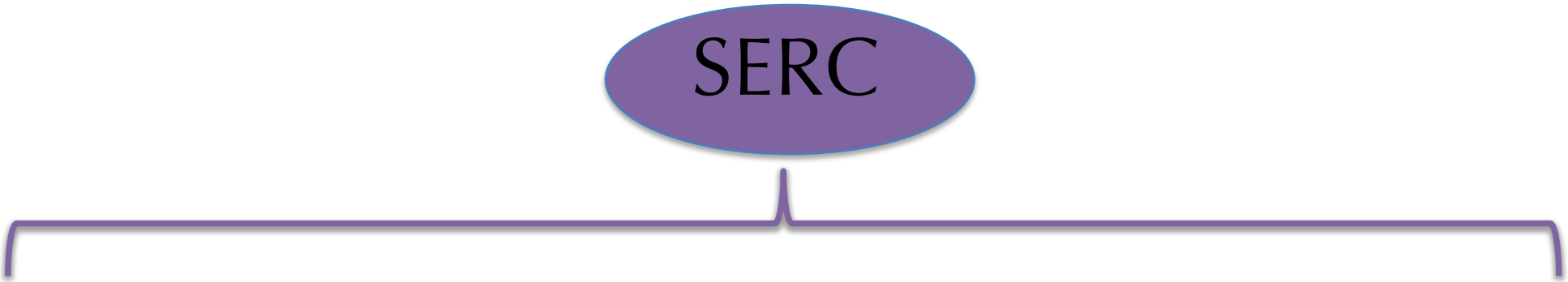
**David Kaiser**

Germeshausen Professor of the History of Science
Professor of Physics
Associate Dean, SERC


***Since Spring 2020:***

70+ faculty and PIs
50+ undergraduates
35+ graduate students
7+ postdocs


Participants in **SERC Action Groups** represent **all 5 Schools** at MIT plus the new MIT Schwarzman College of Computing. They have engaged in a sustained fashion within **multidisciplinary groups** for one or more semesters during the past two years.




## SERC Teaching

### *Vision:*

**Fractal model** for embedding SERC material throughout the curriculum, making it ***inescapable***.

Develop original pedagogical materials by **multidisciplinary teams** with members from across computing, data sciences, humanities, arts, and social sciences—for use in ***each*** of these types of classes.

# SERC

## SERC Teaching

### *Vision:*

**Fractal model** for embedding SERC material throughout the curriculum, making it ***inescapable***.

Develop original pedagogical materials by **multidisciplinary teams** with members from across computing, data sciences, humanities, arts, and social sciences—for use in ***each*** of these types of classes.

Commission and publish a new series of **peer-reviewed case studies**. The cases are **brief** (4k words), based on **original research**, and appropriate for use in **undergraduate instruction** across a range of existing courses and fields of study.

OPEN ACCESS


https://mit-serc.pubpub.org

# SERC

## SERC Teaching

### *Vision:*

**Fractal model** for embedding SERC material throughout the curriculum, making it ***inescapable***.

Develop original pedagogical materials by **multidisciplinary teams** with members from across computing, data sciences, humanities, arts, and social sciences—for use in ***each*** of these types of classes.

Commission and publish a new series of **peer-reviewed case studies**. The cases are **brief** (4k words), based on **original research**, and appropriate for use in **undergraduate instruction** across a range of existing courses and fields of study.

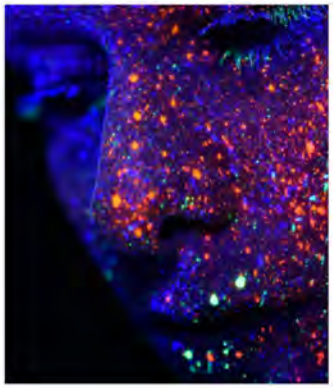
Each submission is **reviewed by 4-6 senior researchers at MIT**, drawn equally from computing and data sciences and from arts, humanities, and social sciences. Each submission is **also reviewed** by **MIT undergraduate volunteers** for balance and accessibility.

OPEN ACCESS

Cases are written by **subject-area experts**, not limited to MIT.


**https://mit-serc.pubpub.org**

# Winter 2021



## The Case of the Nosy Neighbors

*by Johanna Gunawan and Woodrow Hartzog*

Inspired by companies like Clearview AI, Nextdoor, and Amazon, this case study asks students to assume the role of a high-ranking ethics-focused employee at a (fictional) neighborhood-focused social media company. It

## Who Collects the Data? A Tale of Three Maps

*by Catherine D'Ignazio and Lauren Klein*

Who makes maps and who gets mapped? Using a comparative reading of three maps, this case study introduces the idea that data may be useful, but they are not neutral. Rather, they

## The Bias in the Machine: Facial Recognition Technology and Racial Disparities

*by Sidney Perkowitz*

Facial recognition technology (FRT) appears in uses from providing secure access to smartphones, to identifying criminal suspects

## The Dangers of Risk Prediction in the Criminal Justice System

*by Julia Dressel and Hany Farid*

Courts across the United States are using computer software to predict whether a person will commit a crime, the results of which are

https://mit-serc.pubpub.org

MIT Schwarzman College of Computing

Create Pub Search Dashboard DK

HOME AIMS AND SCOPE EDITORS ABOUT THE CASES AUTHOR RESOURCES

# Summer 2021

## Hacking Technology, Hacking Communities: Codes of Conduct and Community Standards in Open Source

*by Christina Dunbar-Hester*

*Published: Aug 10, 2021*

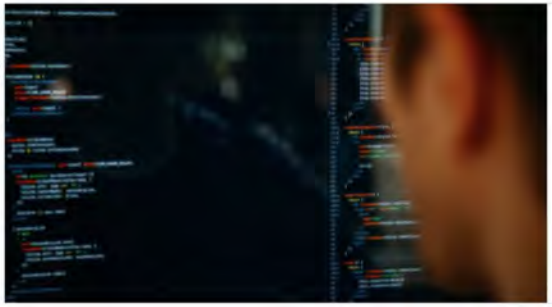
In recent years, the free/libre and open source software (FLOSS) and hacking communities have engaged in lively debates about diversity and inclusion in their ranks...

## Identity, Advertising, and Algorithmic Targeting: Or How (Not) to Target Your "Ideal User"

*by Tanya Kant*

*Published: Aug 10, 2021*

Targeted or "personalized" marketing is an everyday part of most web users' experience. But how do companies "personalize" commercial web content in the context of mass data aggregation?



## Understanding Potential Sources of Harm throughout the Machine Learning Life Cycle

*by Harini Suresh and John Guttag*

*Published: Aug 10, 2021*

As machine learning (ML) increasingly affects people and society, awareness of its potential unwanted consequences has also grown. To anticipate, prevent, and mitigate undesirable downstream consequences, it is critical that we understand when and how harm might be introduced...

## Wrestling with Killer Robots: The Benefits and Challenges of Artificial Intelligence for National Security

*by Erik Lin-Greenberg*

*Published: Aug 10, 2021*

Countries around the world are increasingly turning to artificial intelligence (AI) for national security tasks ranging from intelligence analysis to identifying and attacking targets without human input. This proliferation of AI-enabled military technologies has significant...

## Public Debate on Facial Recognition Technologies in China

*by Tristan G. Brown, Alexander Statman, and Celine Sui*

*Published: Aug 10, 2021*

China's ascent on the global stage in the fields of artificial intelligence and facial recognition has been widely noted in Western-language scholarship and media. Much of the attention has focused on the applications of these technologies in government security systems and...

https://mit-serc.pubpub.org

MIT Schwarzman College of Computing

Create Pub Search Dashboard DK

HOME AIMS AND SCOPE EDITORS ABOUT THE CASES AUTHOR RESOURCES

# Summer 2021

**Hacking Technology, Hacking Communities: Codes of Conduct and Community Standards in Open Source**

*by Christina Dunbar-Hester*

*Published: Aug 10, 2021*

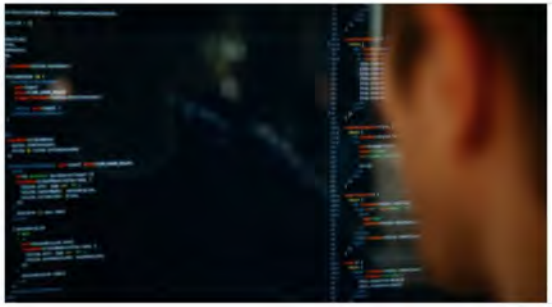
In recent years, the free/libre and open source software (FLOSS) and hacking communities have engaged in lively debates about diversity and inclusion in their ranks...

**Identity, Advertising, and Algorithmic Targeting: Or How (Not) to Target Your "Ideal User"**

*by Tanya Kant*

*Published: Aug 10, 2021*

Targeted or "personalized" marketing is an everyday part of most web users' experience. But how do companies "personalize" commercial web content in the context of mass data aggregation?



**Understanding Potential Sources of Harm throughout the Machine Learning Life Cycle**

*by Harini Suresh and John Guttag*

*Published: Aug 10, 2021*

As machine learning (ML) increasingly affects people and society, awareness of its potential unwanted consequences has also grown. To anticipate, prevent, and mitigate undesirable downstream consequences, it is critical that we understand when and how harm might be introduced...

**Wrestling with Killer Robots: The Benefits and Challenges of Artificial Intelligence for National Security**

*by Erik Lin-Greenberg*

*Published: Aug 10, 2021*

Countries around the world are increasingly turning to artificial intelligence (AI) for national security tasks ranging from intelligence analysis to identifying and attacking targets without human input. This proliferation of AI-enabled military technologies has significant...

computer science, law, urban studies, literature, anthropology, political science, media studies, history

**Public Debate on Facial Recognition Technologies in China**

*by Tristan G. Brown, Alexander Statman, and Celine Sui*

*Published: Aug 10, 2021*

China's ascent on the global stage in the fields of artificial intelligence and facial recognition has been widely noted in Western-language scholarship and media. Much of the attention has focused on the applications of these technologies in government security systems and...

https://mit-serc.pubpub.org

MIT Schwarzman College of Computing

Create Pub Search Dashboard DK

HOME AIMS AND SCOPE EDITORS ABOUT THE CASES AUTHOR RESOURCES

# Summer 2021

**Hacking Technology, Hacking Communities: Codes of Conduct and Community Standards in Open Source**

*by Christina Dunbar-Hester*

*Published: Aug 10, 2021*

In recent years, the free/libre and open source software (FLOSS) and hacking communities have engaged in lively debates about diversity and inclusion in their ranks...

**Identity, Advertising, and Algorithmic Targeting: Or How (Not) to Target Your "Ideal User"**

*by Tanya Kant*

*Published: Aug 10, 2021*

Targeted or "personalized" marketing is an everyday part of most web users' experience. But how do comp... "personalize" commercial web... mass data aggregation...

**Understanding Potential Sources of Harm throughout the Machine... Cycle**

**...ficial Intelligence ...National Security**

*by Erik Lin-Greenberg*

*Published: Aug 10, 2021*

Countries around the world are increasingly turning to artificial intelligence (AI) for national security tasks ranging from intelligence analysis to identifying and attacking targets without human input. This proliferation of AI-enabled military technologies has significant...

Companion site coming soon: original homework problems, in-class demos, and active learning projects.

MIT OCW

computer science, law, urban studies, literature, anthropology, political science, media studies, history

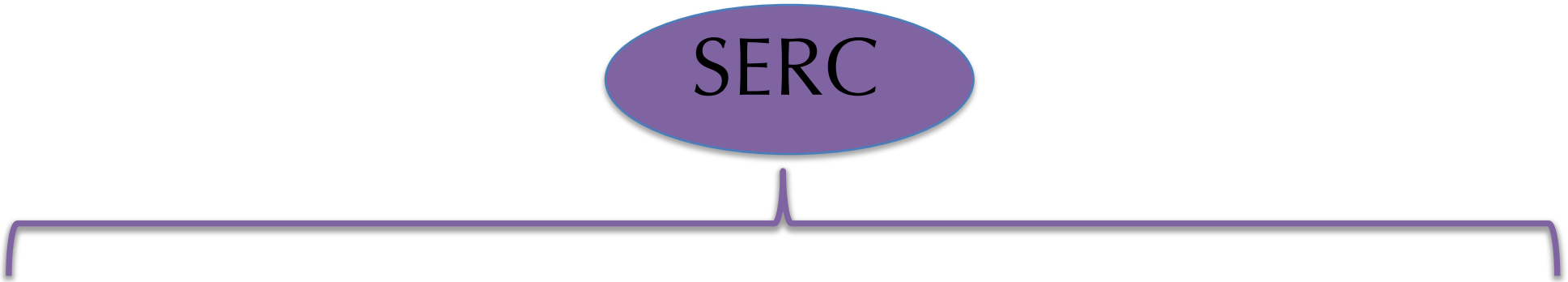
**Public Debate on Facial Recognition Technologies in China**

*by Tristan G. Brown, Alexander Statman, and Celine Sui*

*Published: Aug 10, 2021*

China's ascent on the global stage in the fields of artificial intelligence and facial recognition has been widely noted in Western-language scholarship and media. Much of the attention has focused on the applications of these technologies in government security systems and...

https://mit-serc.pubpub.org




## SERC Research

### *Vision:*

**Resources** and **tools** used by researchers who are engaged in large-scale applied research partnerships.

Research programs conducted with **external partners** that inform the design and use of the tools for **better outcomes**.

Faculty stewards and champions trained to work in **multidisciplinary teams** to assess potential harms as well as benefits from computing research.

# SERC

## SERC Research

### *Vision:*

**Resources** and **tools** used by researchers who are engaged in large-scale applied research partnerships.

Research programs conducted with **external partners** that inform the design and use of the tools for **better outcomes**.

Faculty stewards and champions trained to work in **multidisciplinary teams** to assess potential harms as well as benefits from computing research.

Informed by safety-engineering practices in other fields, develop **after-incident reports** on **unintended consequences** originating from computing research, development, or implementation.

# SERC

## SERC Research

### *Vision:*

**Resources** and **tools** used by researchers who are engaged in large-scale applied research partnerships.

Research programs conducted with **external partners** that inform the design and use of the tools for **better outcomes**.

Faculty stewards and champions trained to work in **multidisciplinary teams** to assess potential harms as well as benefits from computing research.

Informed by safety-engineering practices in other fields, develop **after-incident reports** on **unintended consequences** originating from computing research, development, or implementation.

Train crash at Montparnasse station, Paris, 1895

# SERC

## SERC Research

### *Vision:*

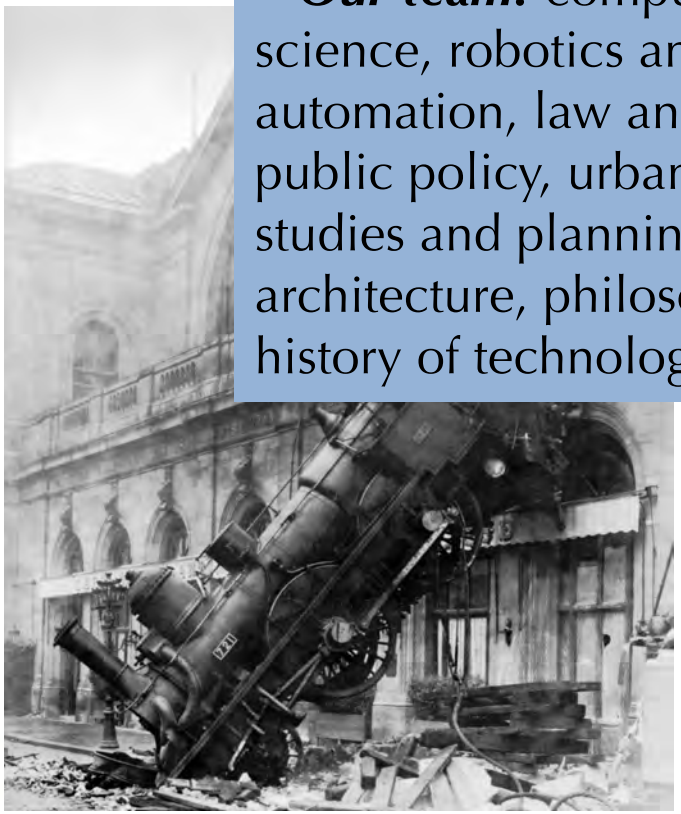
**Resources** and **tools** used by researchers who are engaged in large-scale applied research partnerships.

Research programs conducted with **external partners** that inform the design and use of the tools for **better outcomes**.

Faculty stewards and champions trained to work in **multidisciplinary teams** to assess potential harms as well as benefits from computing research.

Informed by safety-engineering practices in other fields, develop **after-incident reports** on **unintended consequences** originating from computing research, development, or implementation.

***Our team:*** computer science, robotics and automation, law and public policy, urban studies and planning, architecture, philosophy, history of technology.



Train crash at Montparnasse station, Paris, 1895

# SERC

## SERC Research

### *Vision:*

**Resources** and **tools** used by researchers who are engaged in large-scale applied research partnerships.

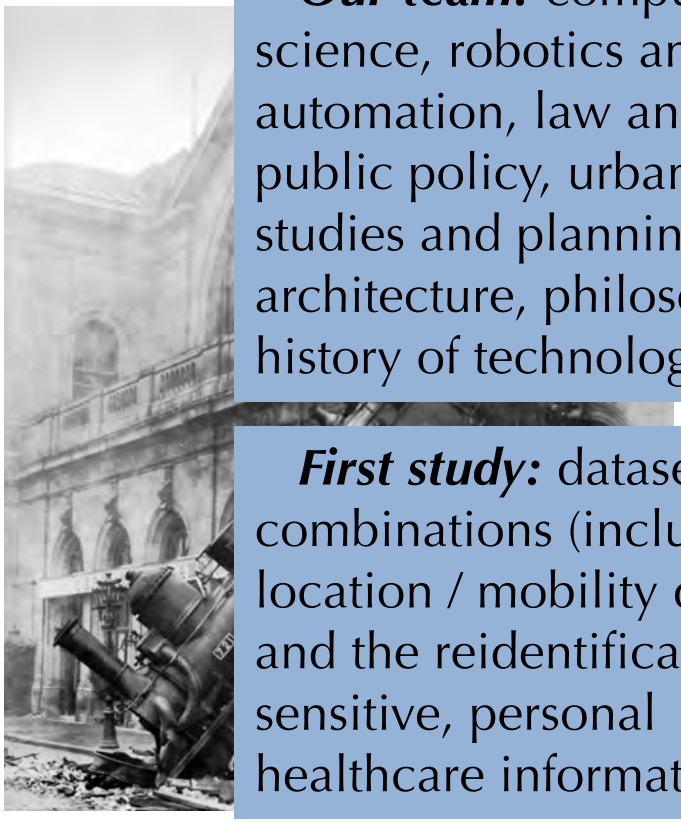
Research programs conducted with **external partners** that inform the design and use of the tools for **better outcomes**.

Faculty stewards and champions trained to work in **multidisciplinary teams** to assess potential harms as well as benefits from computing research.

Informed by safety-engineering practices in other fields, develop **after-incident reports** on **unintended consequences** originating from computing research, development, or implementation.

***Our team:*** computer science, robotics and automation, law and public policy, urban studies and planning, architecture, philosophy, history of technology.

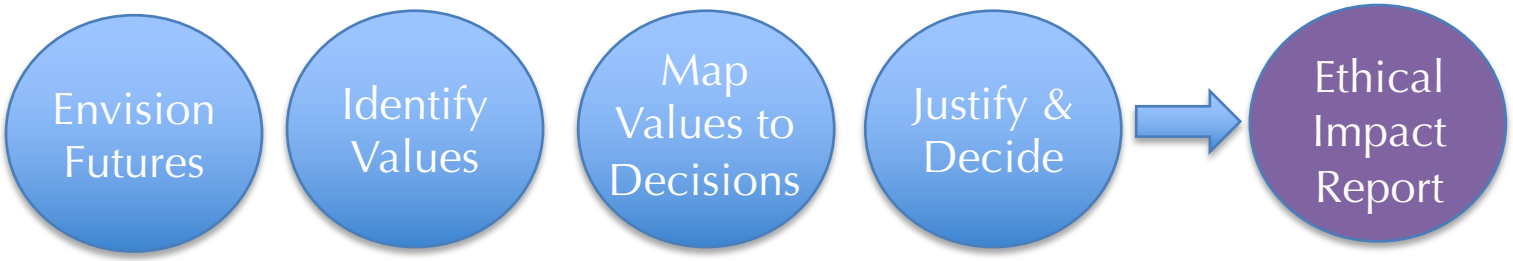
***First study:*** dataset combinations (including location / mobility data) and the reidentification of sensitive, personal healthcare information.



Train crash at Montparnasse station, Paris, 1895

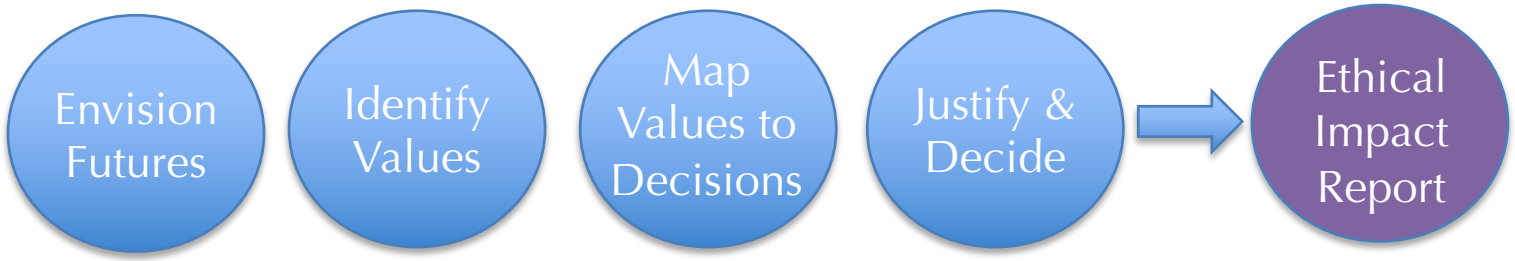
# New Tools

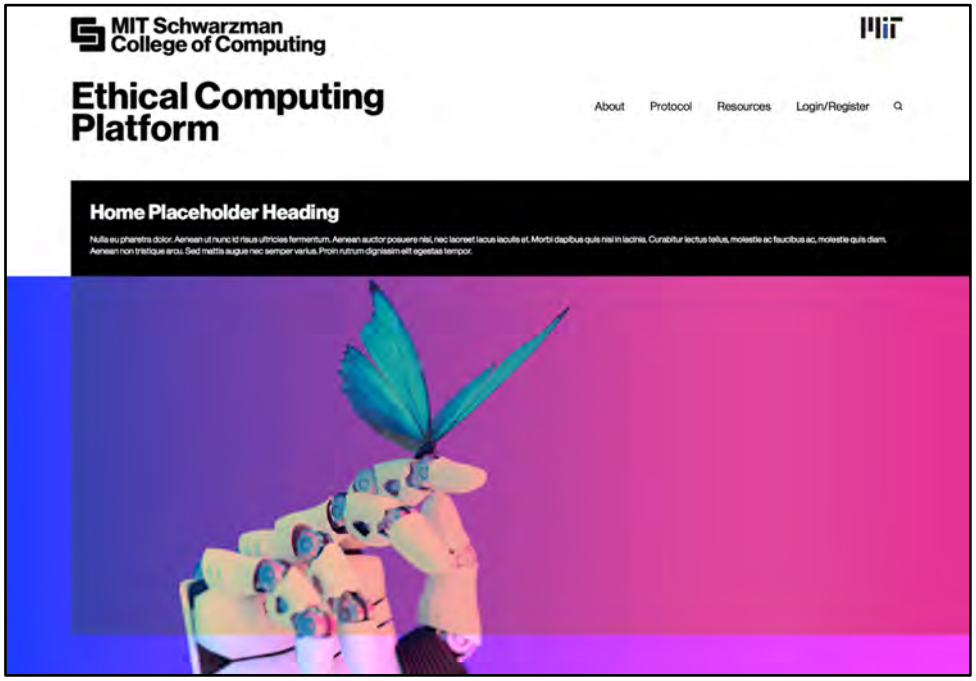
***Ethical Computing Platform:*** a free and open **online platform** that research groups can use ***iteratively*** to address a series of questions about **benefits and burdens** of their **proposed research design**, **data collection** and **protection protocols**, and **potential unintended consequences** of a given project.



# New Tools

***Ethical Computing Platform:*** a free and open **online platform** that research groups can use ***iteratively*** to address a series of questions about **benefits and burdens** of their **proposed research design**, **data collection** and **protection protocols**, and **potential unintended consequences** of a given project.




*coming soon!*

**Abby Everett Jaques**
MIT Philosophy PhD, IPC Research Affiliate

**Milo Phillips-Brown**
SERC postdoc, now faculty at Oxford

**Aude Oliva**
MIT Quest for Intelligence

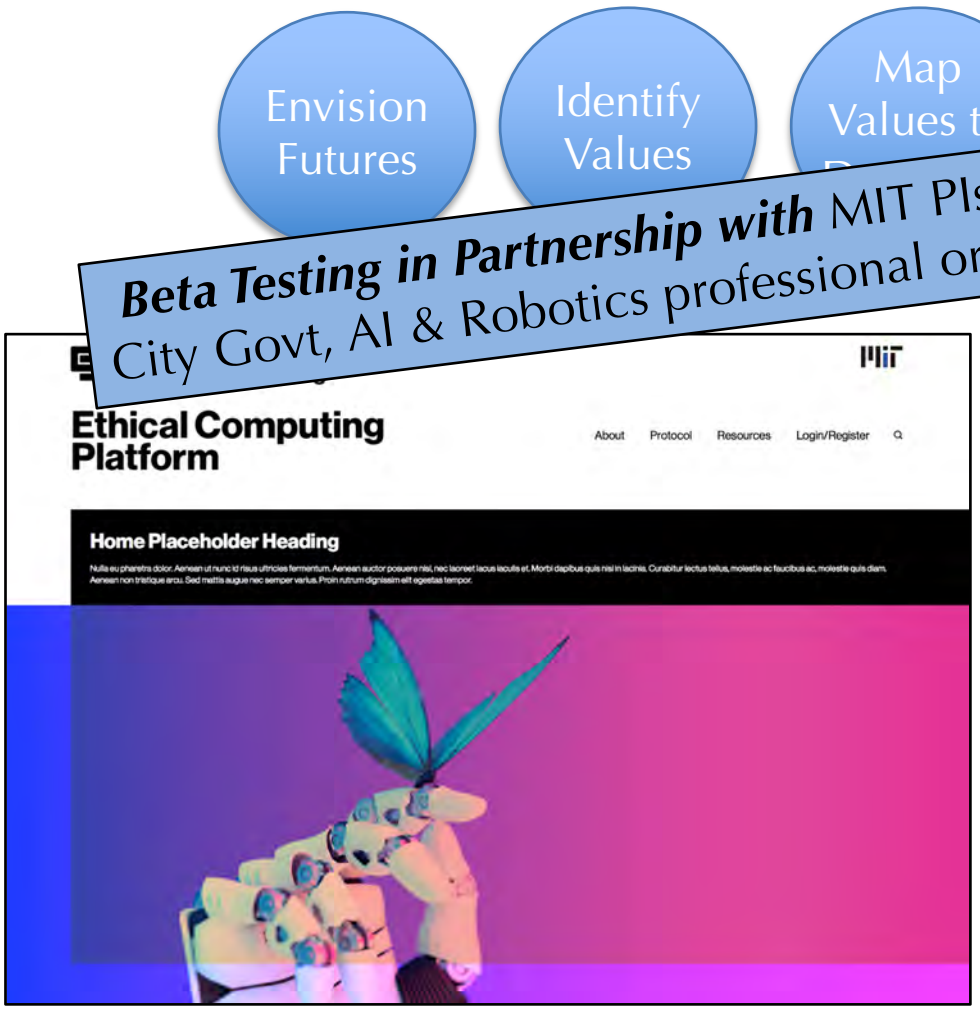
**Ben Armstrong**
Executive Director, MIT IPC

**David Kaiser**
STS/Physics, SERC

**Julie Shah**
AeroAstro/CSAIL, SERC

# New Tools

***Ethical Computing Platform:*** a free and open **online platform** that research groups can use ***iteratively*** to address a series of questions about **benefits and burdens** of their **proposed research design**, **data collection** and **protection protocols**, and **potential unintended consequences** of a given project.




***Beta Testing in Partnership with*** MIT PIs and graduate researchers, Cambridge City Govt, AI & Robotics professional organizations, industry research partners

*coming soon!*

**Abby Everett Jaques**
MIT Philosophy PhD,
IPC Research Affiliate

**Milo Phillips-Brown**
SERC postdoc, now
faculty at Oxford

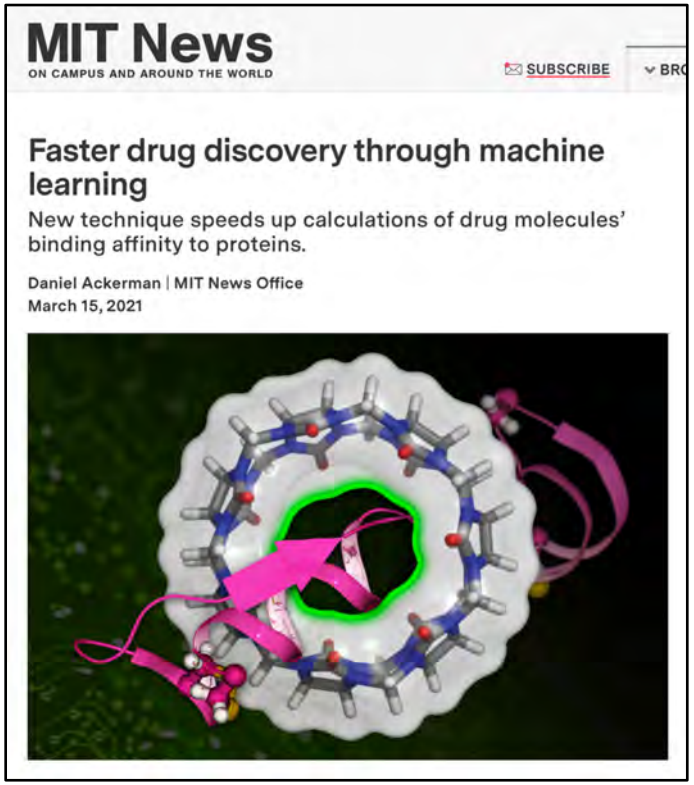
**Aude Oliva**
MIT Quest for
Intelligence

**Ben Armstrong**
Executive Director,
MIT IPC

**David Kaiser**
STS/Physics,
SERC

**Julie Shah**
AeroAstro/CSAIL,
SERC

Computing and the data sciences are enabling **fantastic progress** and **benefits for society**…



MIT News
ON CAMPUS AND AROUND THE WORLD

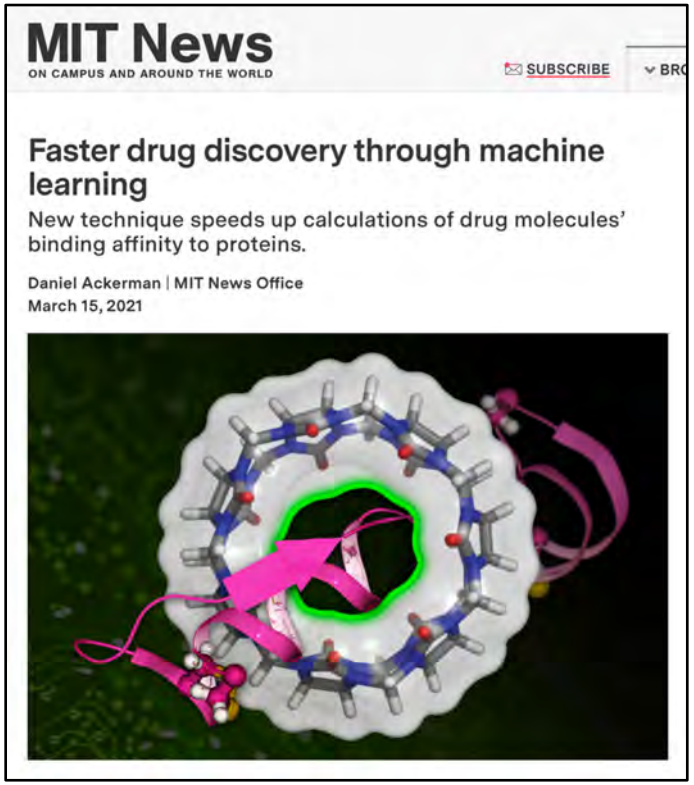
SUBSCRIBE

## Faster drug discovery through machine learning

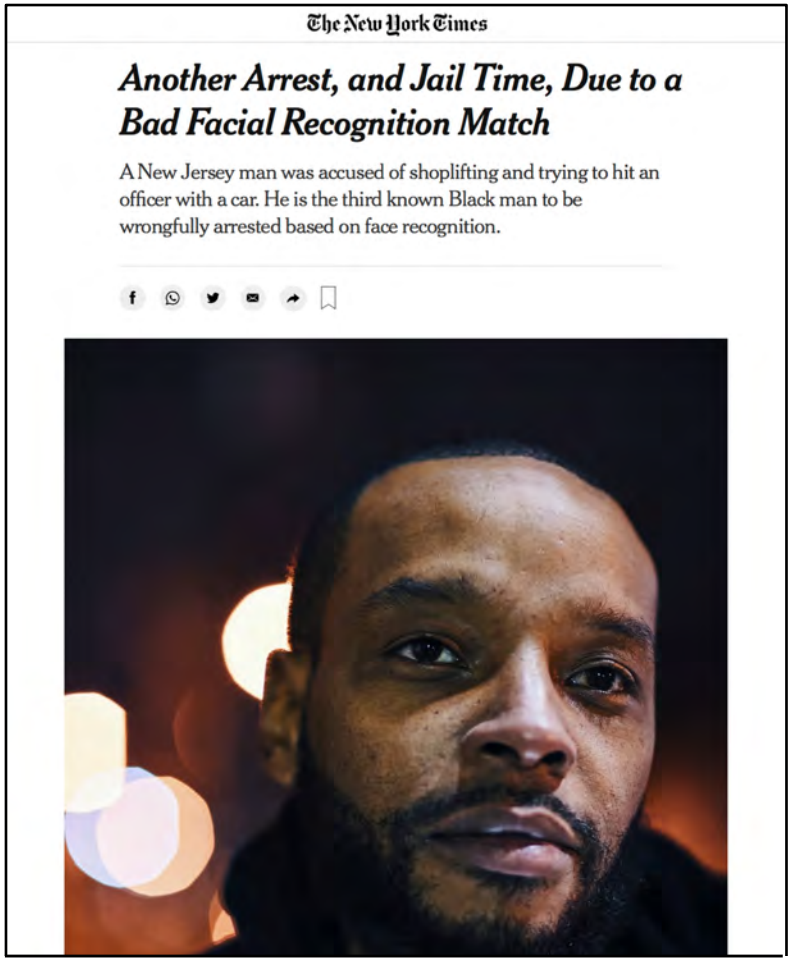
New technique speeds up calculations of drug molecules' binding affinity to proteins.

Daniel Ackerman | MIT News Office
March 15, 2021

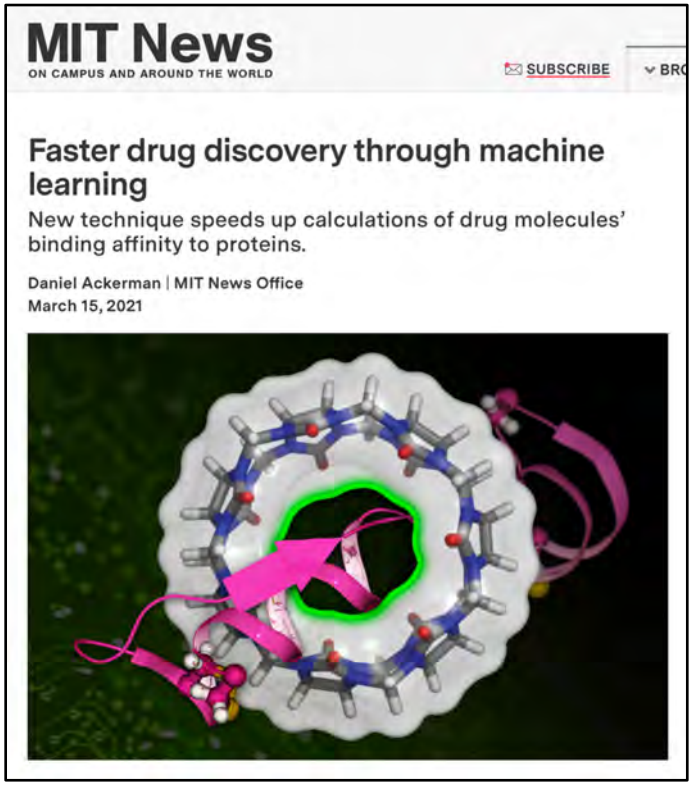
Computing and the data sciences are enabling **fantastic progress** and **benefits for society**…



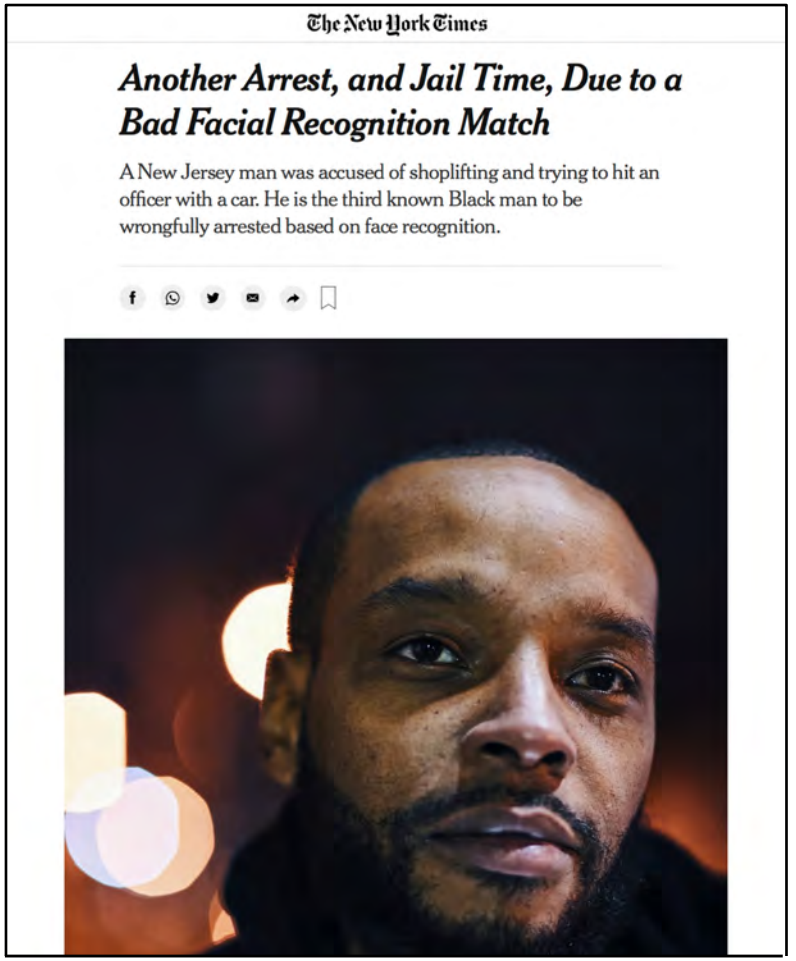

… but like **any new technology**, they carry risks of **unintended consequences** and **real harms** for individuals and groups.



Computing and the data sciences are enabling **fantastic progress** and **benefits for society**…

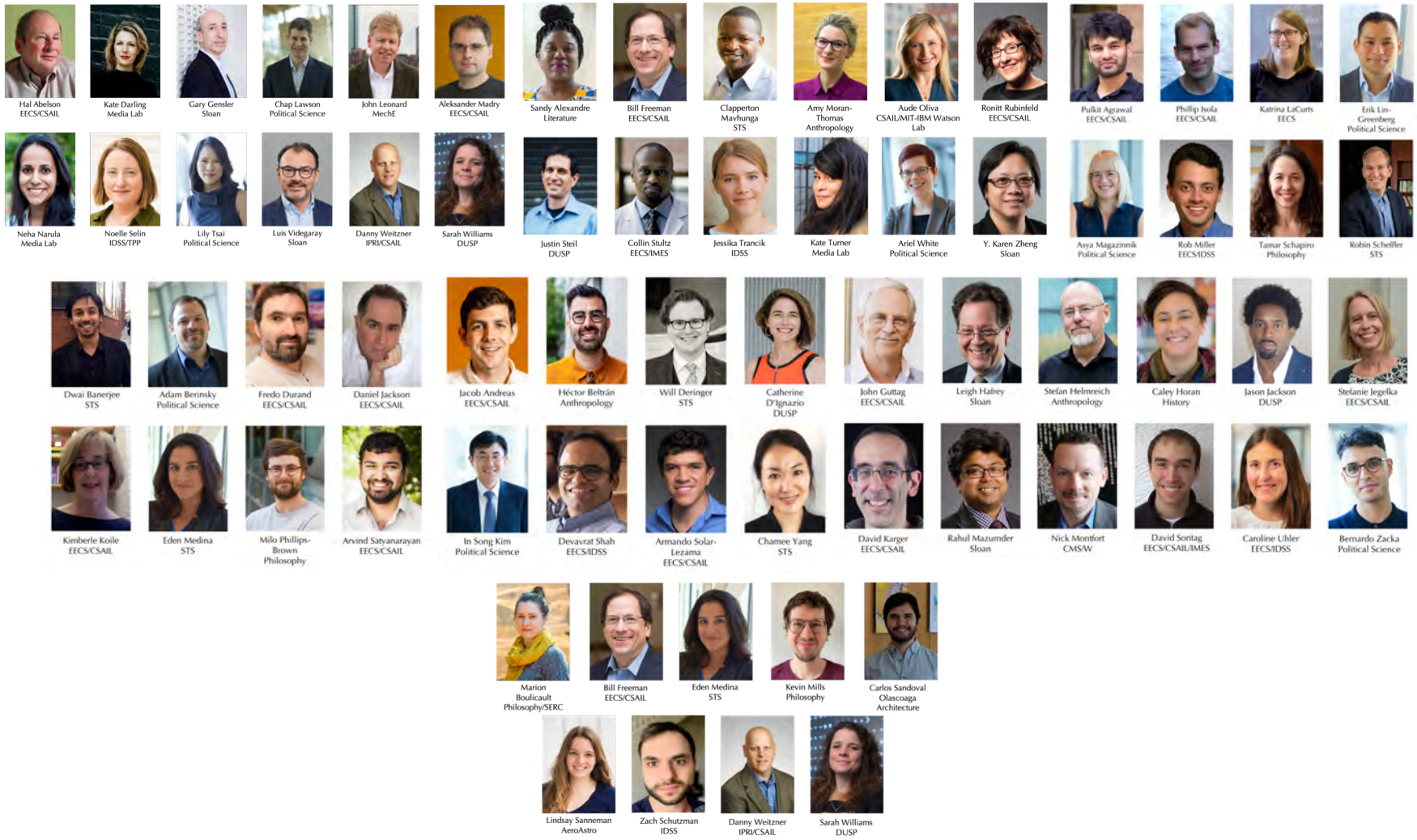



… but like **any new technology**, they carry risks of **unintended consequences** and **real harms** for individuals and groups.



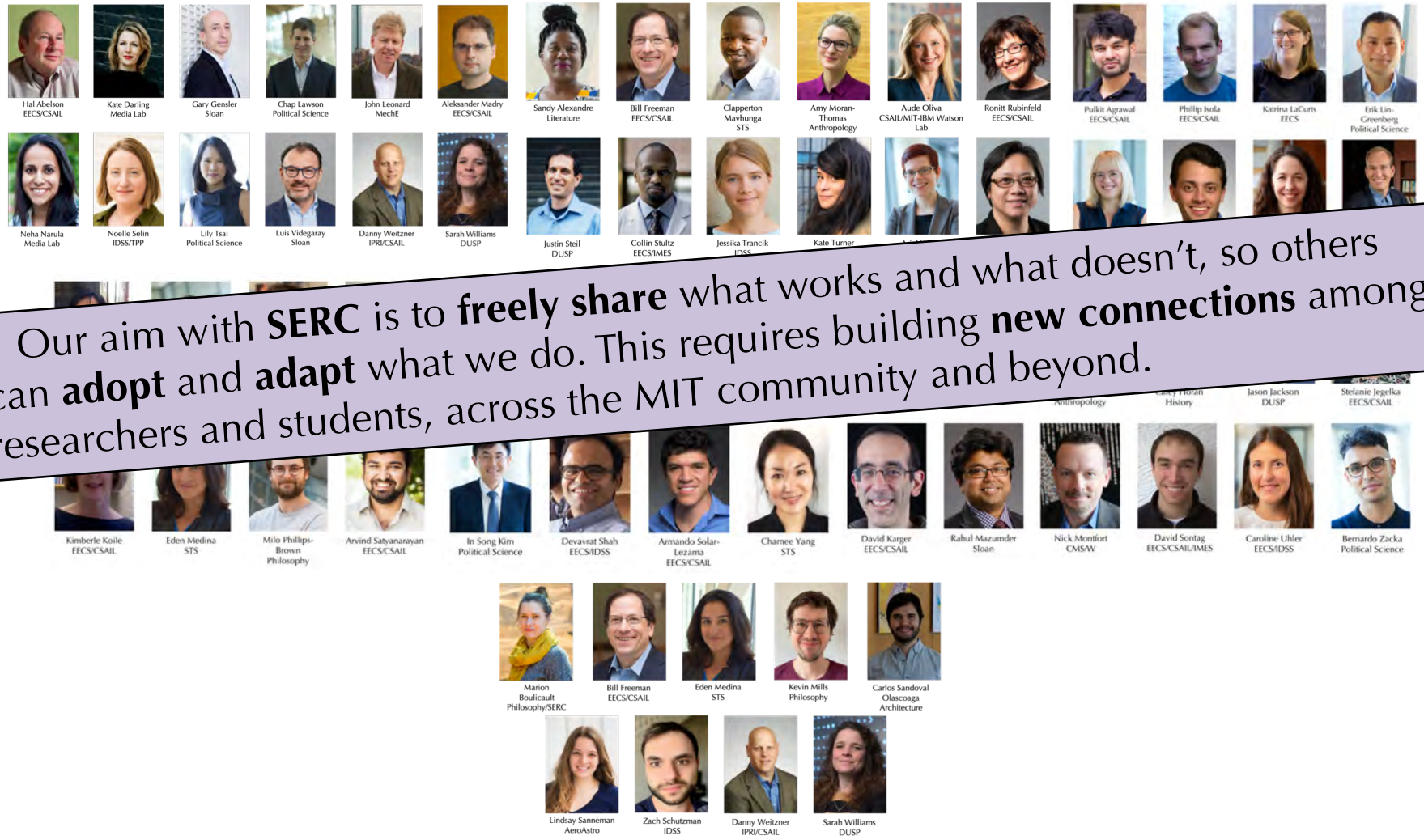

The most **significant challenges** arise when techniques **leave the lab** and impact **real-world situations**.

Identifying potential consequences and rectifying harms requires **input** and **expertise** from **across the fields** of science, engineering, humanities, and the social sciences …




… as well as working with **partners beyond academia** — including people who are **impacted** by the latest technologies as well as the people who **implement** them.

Identifying potential consequences and rectifying harms requires **input** and **expertise** from **across the fields** of science, engineering, humanities, and the social sciences …




Our aim with **SERC** is to **freely share** what works and what doesn't, so others can **adopt** and **adapt** what we do. This requires building **new connections** among researchers and students, across the MIT community and beyond.

… as well as working with **partners beyond academia** — including people who are **impacted** by the latest technologies as well as the people who **implement** them.