33RD USENIX SECURITY SYMPOSIUM

"I Don't Know If We're Doing Good. I Don't Know If We're Doing Bad" **Investigating How Practitioners Scope, Motivate, and Conduct Privacy Work When Developing AI Products**

Hao-Ping (Hank) Lee

Lan Gao

Stephanie Yang

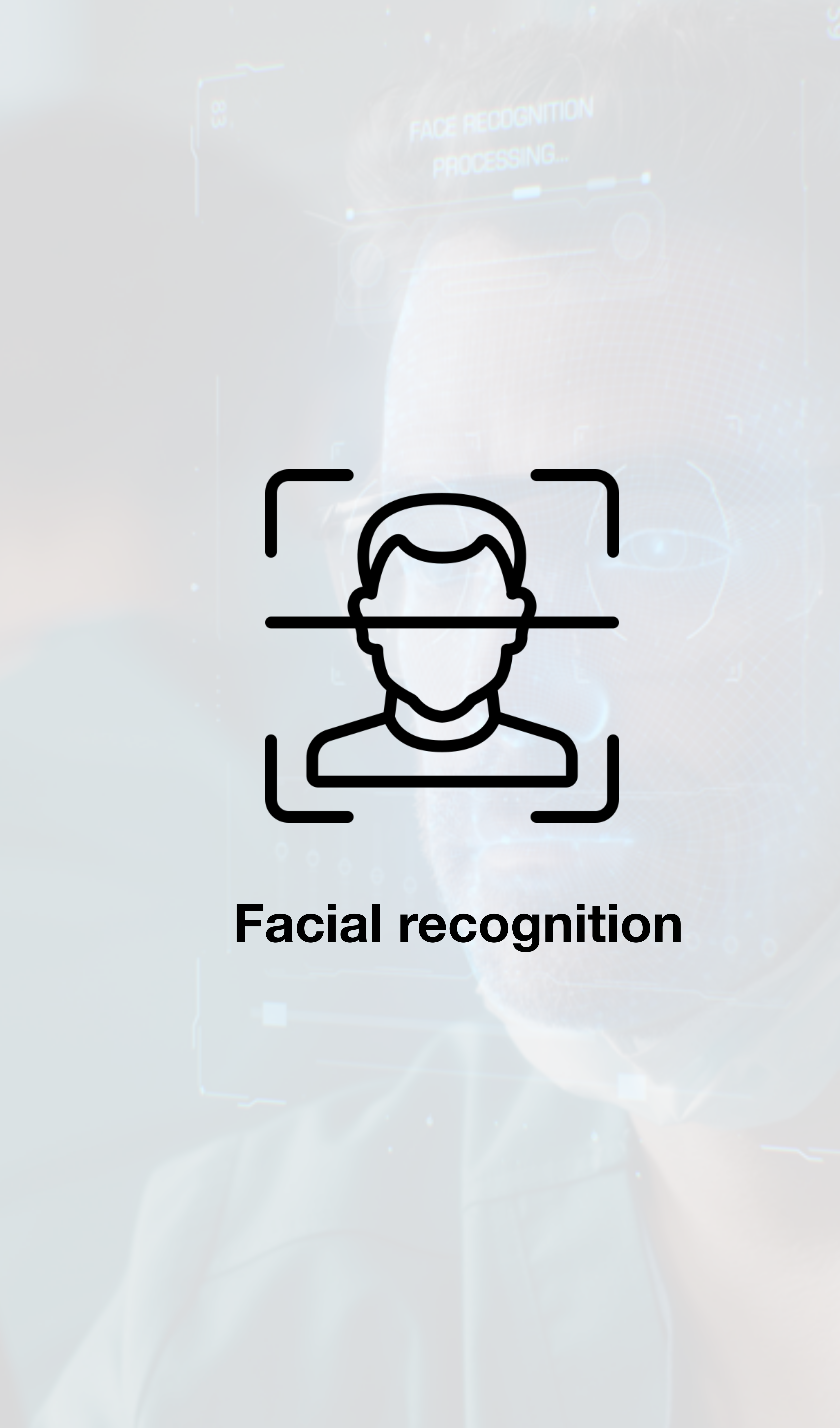
Jodi Forlizzi

Sauvik Das

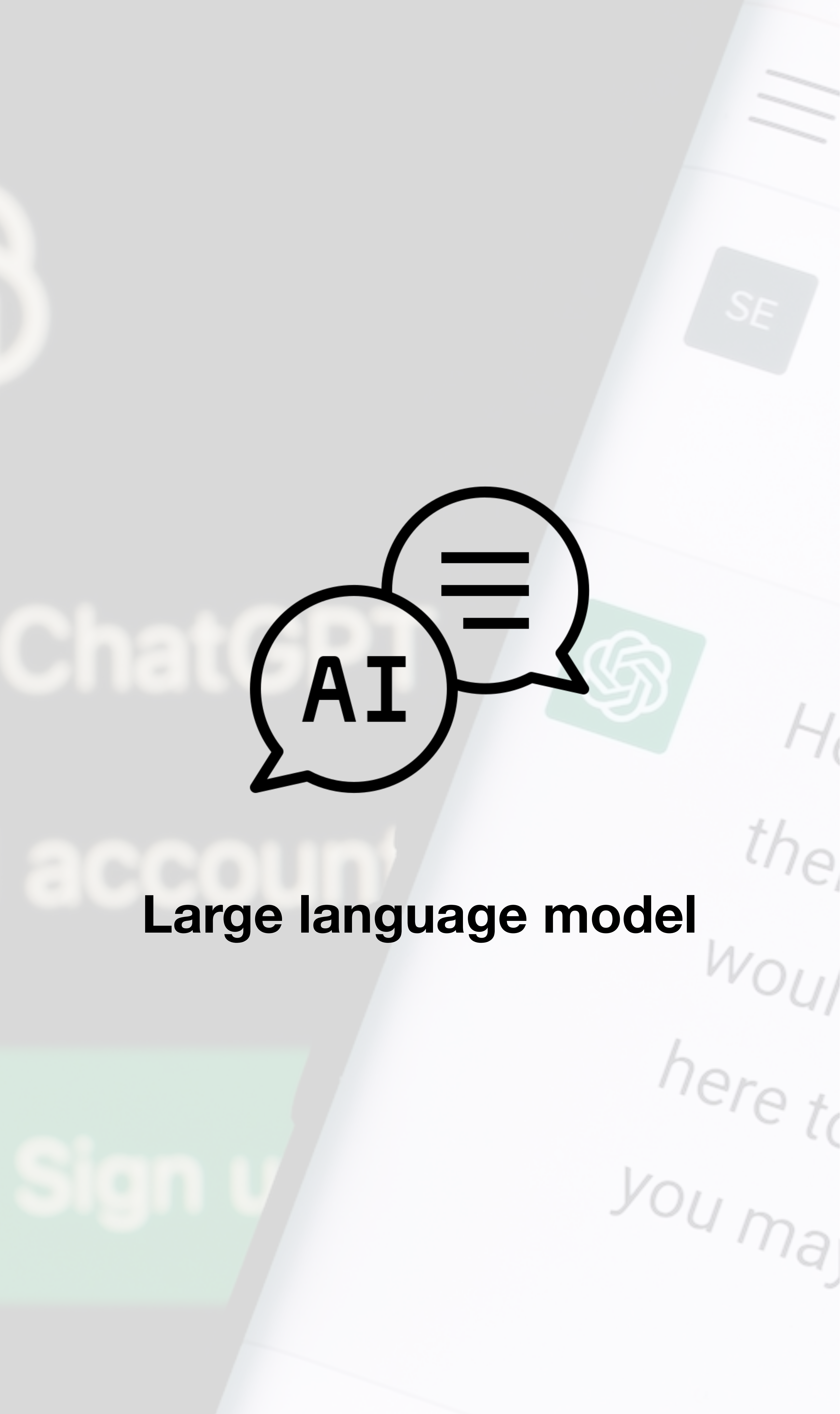
Human-Computer Interaction Institute

Carnegie Mellon University

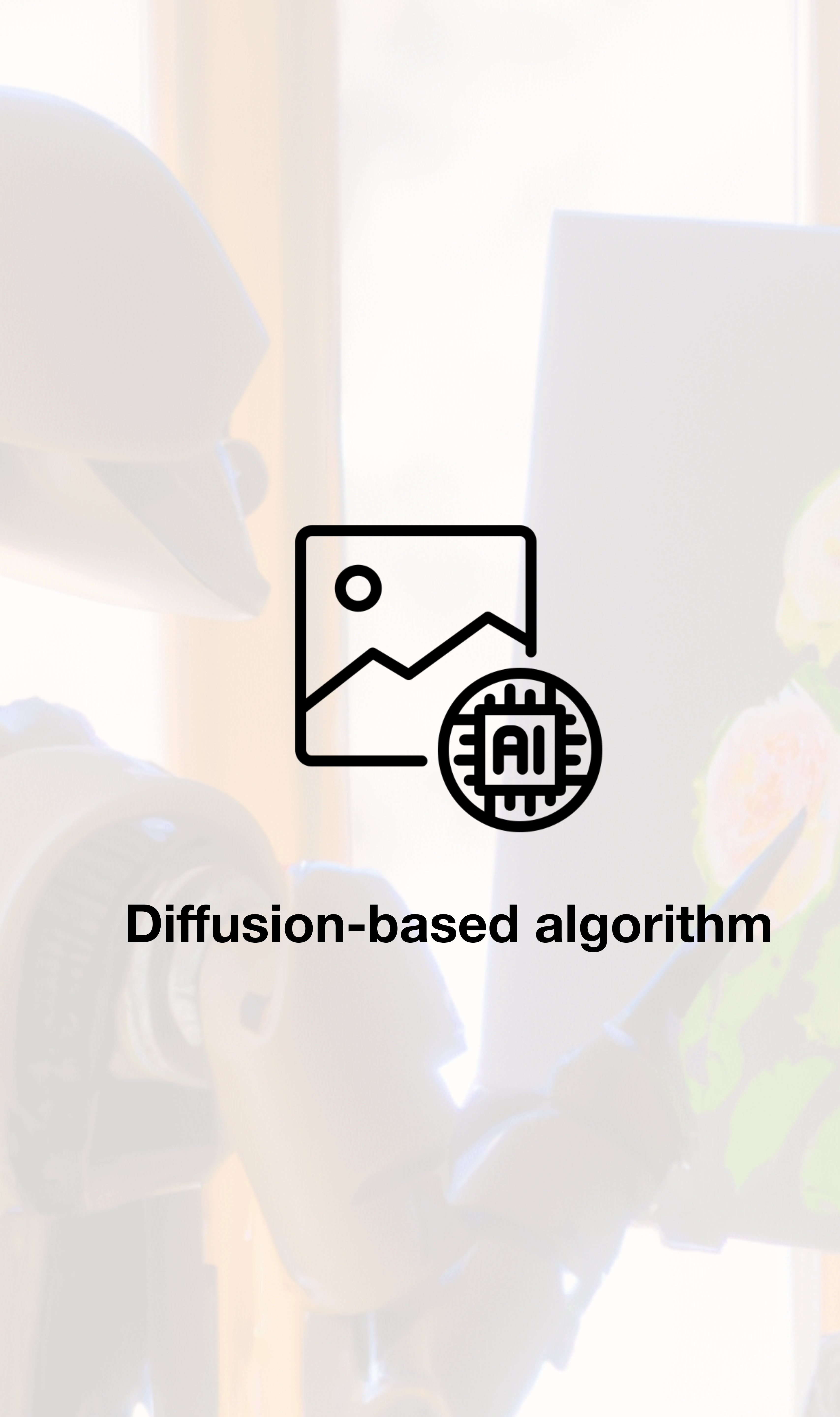
Georgia Tech



Facial recognition

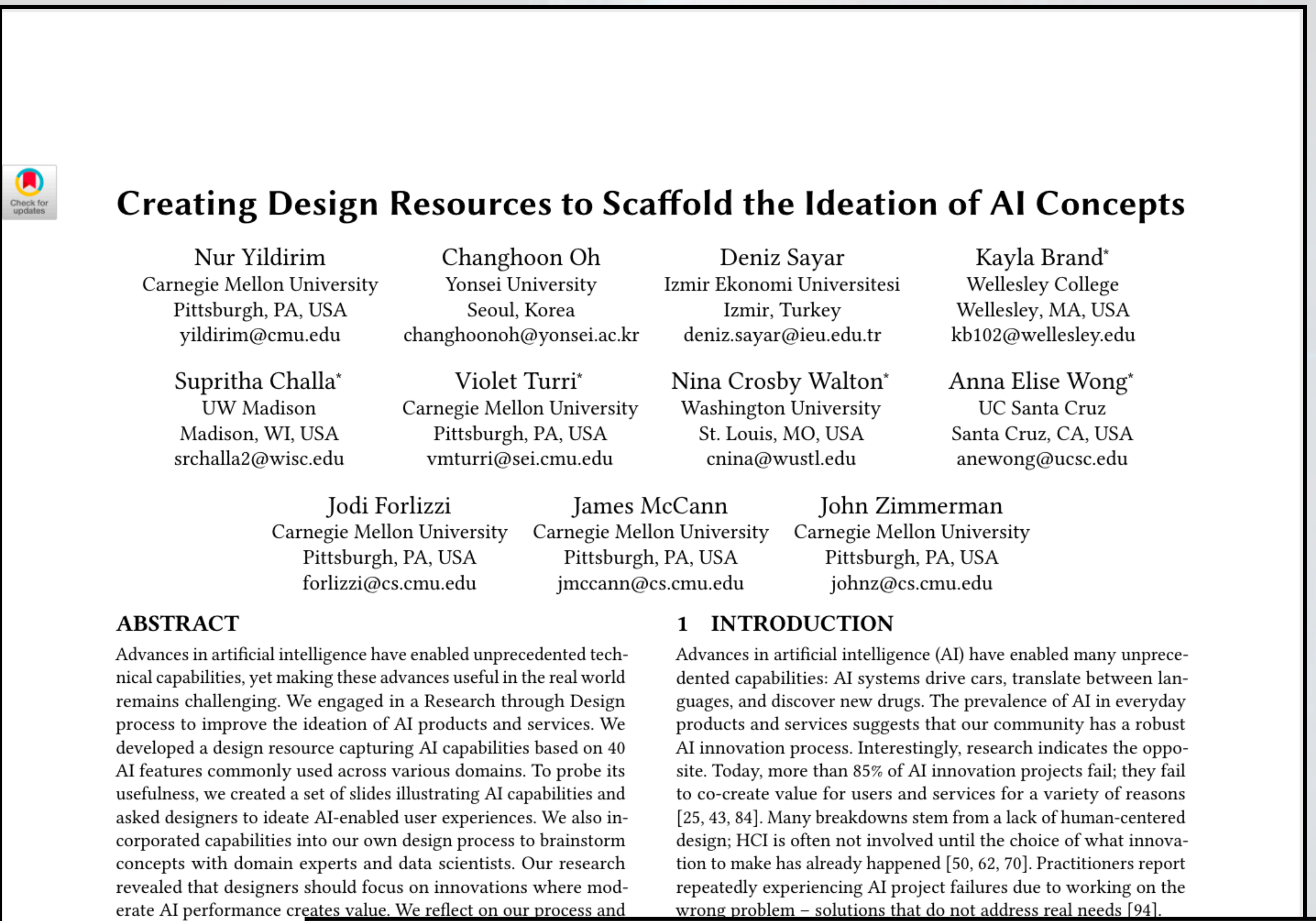


Large language model



Diffusion-based algorithm

AI as a design material



Creating Design Resources to Scaffold the Ideation of AI Concepts

Nur Yildirim
Carnegie Mellon University
Pittsburgh, PA, USA
yildirim@cmu.edu

Changhoon Oh
Yonsei University
Seoul, Korea
changhoonoh@yonsei.ac.kr

Deniz Sayar
Izmir Ekonomi Universitesi
Izmir, Turkey
deniz.sayar@ieu.edu.tr

Kayla Brand*
Wellesley College
Wellesley, MA, USA
kb102@wellesley.edu

Supritha Challa*
UW Madison
Madison, WI, USA
srchalla2@wisc.edu

Violet Turri*
Carnegie Mellon University
Pittsburgh, PA, USA
vmturri@sei.cmu.edu

Nina Crosby Walton*
Washington University
St. Louis, MO, USA
cnina@wustl.edu

Anna Elise Wong*
UC Santa Cruz
Santa Cruz, CA, USA
anewong@ucsc.edu

Jodi Forlizzi
Carnegie Mellon University
Pittsburgh, PA, USA
forlizzi@cs.cmu.edu

James McCann
Carnegie Mellon University
Pittsburgh, PA, USA
jmccann@cs.cmu.edu

John Zimmerman
Carnegie Mellon University
Pittsburgh, PA, USA
johnz@cs.cmu.edu

ABSTRACT

Advances in artificial intelligence have enabled unprecedented technical capabilities, yet making these advances useful in the real world remains challenging. We engaged in a Research through Design process to improve the ideation of AI products and services. We developed a design resource capturing AI capabilities based on 40 AI features commonly used across various domains. To probe its usefulness, we created a set of slides illustrating AI capabilities and asked designers to ideate AI-enabled user experiences. We also incorporated capabilities into our own design process to brainstorm concepts with domain experts and data scientists. Our research revealed that designers should focus on innovations where moderate AI performance creates value. We reflect on our process and
discuss research implica
to systematically explore

CCS CONCEPTS

• Human-centered cor
and methods.

KEYWORDS

User experience, artificial

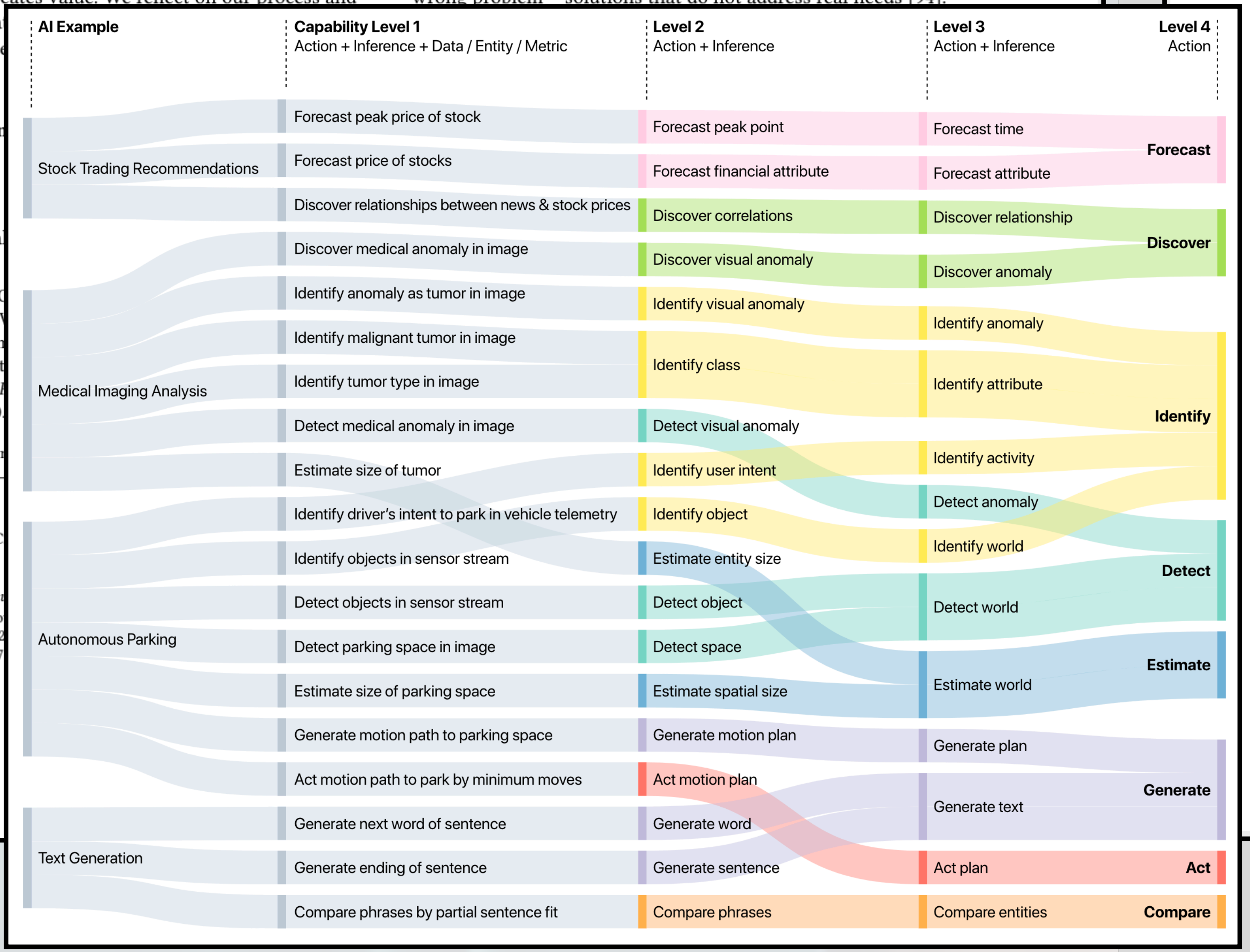
ACM Reference Format:
Nur Yildirim, Changhoon C
Violet Turri, Nina Crosby W
McCann, and John Zimmern
the Ideation of AI Concept
(DIS '23), July 10–14, 2023, P
21 pages. https://doi.org/10.

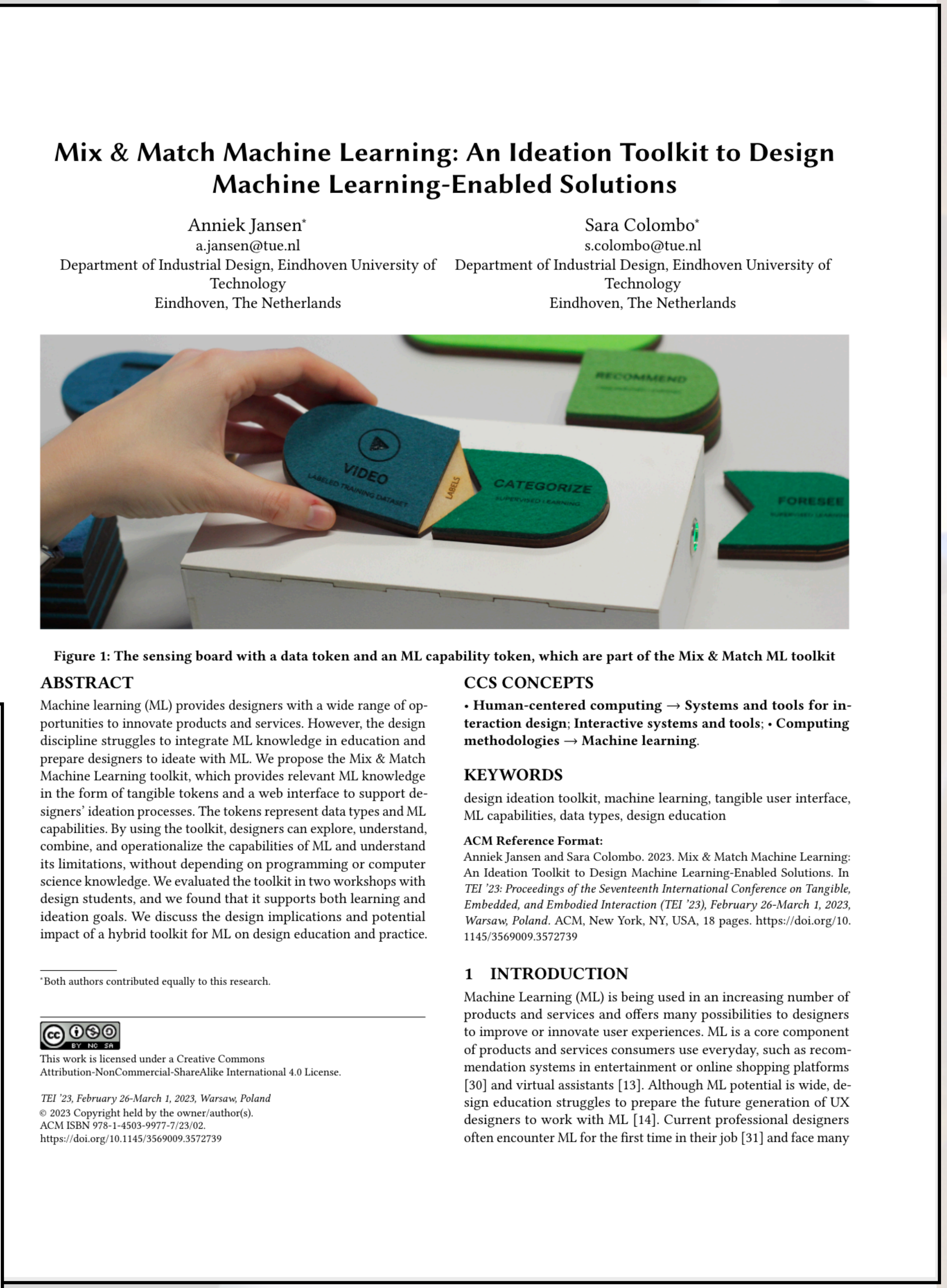
*Work done as a research intern

DIS '23, July 10–14, 2023, Pittsbu
© 2023 Copyright held by the o
ACM ISBN 978-1-4503-9893-0/2
https://doi.org/10.1145/3563657

1 INTRODUCTION

Advances in artificial intelligence (AI) have enabled many unprecedented capabilities: AI systems drive cars, translate between languages, and discover new drugs. The prevalence of AI in everyday products and services suggests that our community has a robust AI innovation process. Interestingly, research indicates the opposite. Today, more than 85% of AI innovation projects fail; they fail to co-create value for users and services for a variety of reasons [25, 43, 84]. Many breakdowns stem from a lack of human-centered design; HCI is often not involved until the choice of what innovation to make has already happened [50, 62, 70]. Practitioners report repeatedly experiencing AI project failures due to working on the wrong problem – solutions that do not address real needs [94].





Mix & Match Machine Learning: An Ideation Toolkit to Design Machine Learning-Enabled Solutions

Anniek Jansen*
a.jansen@tue.nl
Department of Industrial Design, Eindhoven University of Technology
Eindhoven, The Netherlands

Sara Colombo*
s.colombo@tue.nl
Department of Industrial Design, Eindhoven University of Technology
Eindhoven, The Netherlands

Figure 1: The sensing board with a data token and an ML capability token, which are part of the Mix & Match ML toolkit

ABSTRACT

Machine learning (ML) provides designers with a wide range of opportunities to innovate products and services. However, the design discipline struggles to integrate ML knowledge in education and prepare designers to ideate with ML. We propose the Mix & Match Machine Learning toolkit, which provides relevant ML knowledge in the form of tangible tokens and a web interface to support designers' ideation processes. The tokens represent data types and ML capabilities. By using the toolkit, designers can explore, understand, combine, and operationalize the capabilities of ML and understand its limitations, without depending on programming or computer science knowledge. We evaluated the toolkit in two workshops with design students, and we found that it supports both learning and ideation goals. We discuss the design implications and potential impact of a hybrid toolkit for ML on design education and practice.

*Both authors contributed equally to this research.

TEI '23, February 26-March 1, 2023, Warsaw, Poland
© 2023 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9977-7/23/02.
https://doi.org/10.1145/3569009.3572739

CCS CONCEPTS

• **Human-centered computing → Systems and tools for interaction design; Interactive systems and tools; • Computing methodologies → Machine learning.**

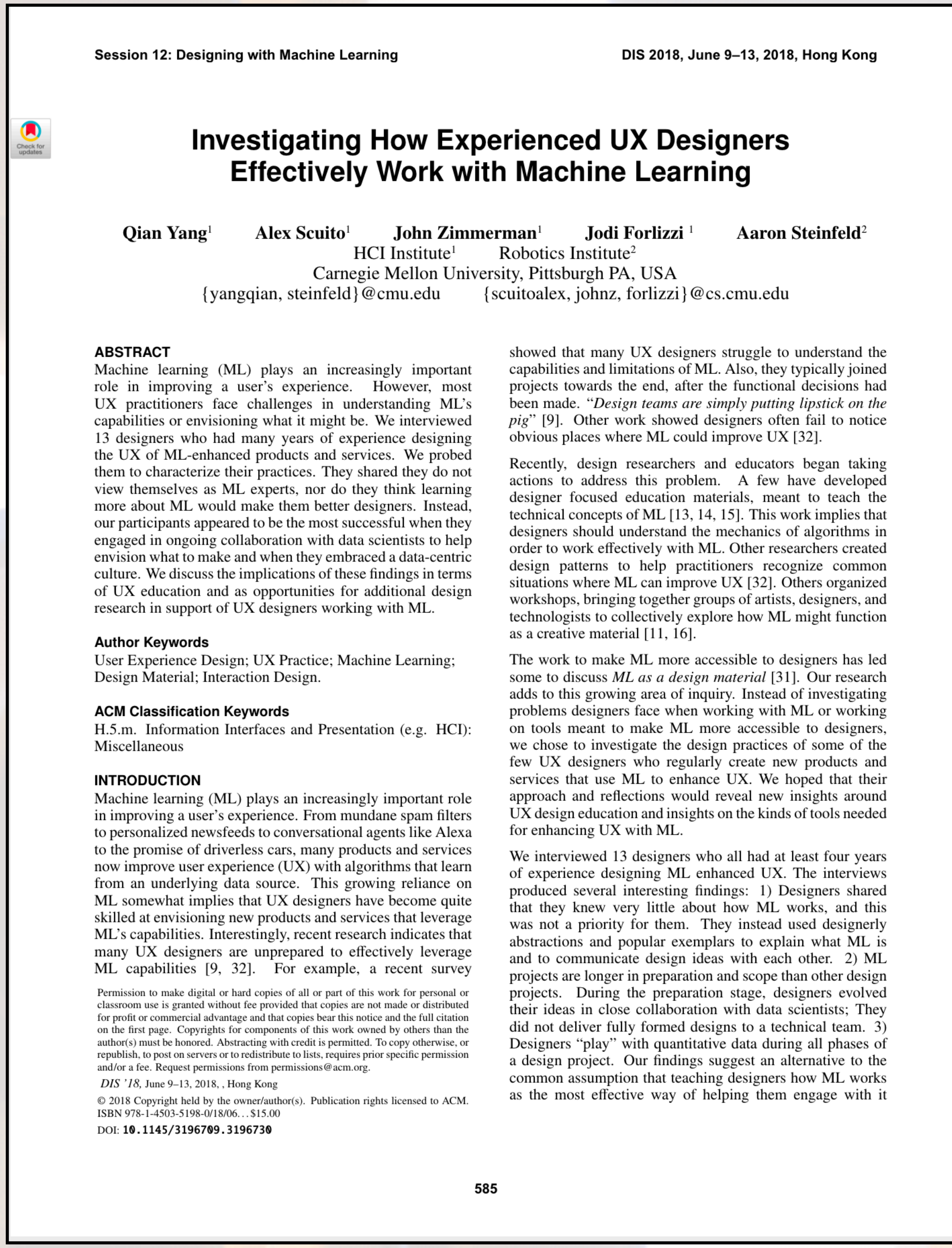
KEYWORDS

design ideation toolkit, machine learning, tangible user interface, ML capabilities, data types, design education

ACM Reference Format:
Anniek Jansen and Sara Colombo. 2023. Mix & Match Machine Learning: An Ideation Toolkit to Design Machine Learning-Enabled Solutions. In *TEI '23: Proceedings of the Seventeenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '23), February 26-March 1, 2023, Warsaw, Poland.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3569009.3572739

1 INTRODUCTION

Machine Learning (ML) is being used in an increasing number of products and services and offers many possibilities to designers to improve or innovate user experiences. ML is a core component of products and services consumers use everyday, such as recommendation systems in entertainment or online shopping platforms [30] and virtual assistants [13]. Although ML potential is wide, design education struggles to prepare the future generation of UX designers to work with ML [14]. Current professional designers often encounter ML for the first time in their job [31] and face many



Investigating How Experienced UX Designers Effectively Work with Machine Learning

Qian Yang[1] Alex Scuito[1] John Zimmerman[1] Jodi Forlizzi [1] Aaron Steinfeld[2]
HCI Institute[1] Robotics Institute[2]
Carnegie Mellon University, Pittsburgh PA, USA
{yangqian, steinfeld}@cmu.edu {scuitoalex, johnz, forlizzi}@cs.cmu.edu

ABSTRACT

Machine learning (ML) plays an increasingly important role in improving a user's experience. However, most UX practitioners face challenges in understanding ML's capabilities or envisioning what it might be. We interviewed 13 designers who had many years of experience designing the UX of ML-enhanced products and services. We probed them to characterize their practices. They shared they do not view themselves as ML experts, nor do they think learning more about ML would make them better designers. Instead, our participants appeared to be the most successful when they engaged in ongoing collaboration with data scientists to help envision what to make and when they embraced a data-centric culture. We discuss the implications of these findings in terms of UX education and as opportunities for additional design research in support of UX designers working with ML.

Author Keywords

User Experience Design; UX Practice; Machine Learning; Design Material; Interaction Design.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Machine learning (ML) plays an increasingly important role in improving a user's experience. From mundane spam filters to personalized newsfeeds to conversational agents like Alexa to the promise of driverless cars, many products and services now improve user experience (UX) with algorithms that learn from an underlying data source. This growing reliance on ML somewhat implies that UX designers have become quite skilled at envisioning new products and services that leverage ML's capabilities. Interestingly, recent research indicates that many UX designers are unprepared to effectively leverage ML capabilities [9, 32]. For example, a recent survey showed that many UX designers struggle to understand the capabilities and limitations of ML. Also, they typically joined projects towards the end, after the functional decisions had been made. "*Design teams are simply putting lipstick on the pig*" [9]. Other work showed designers often fail to notice obvious places where ML could improve UX [32].

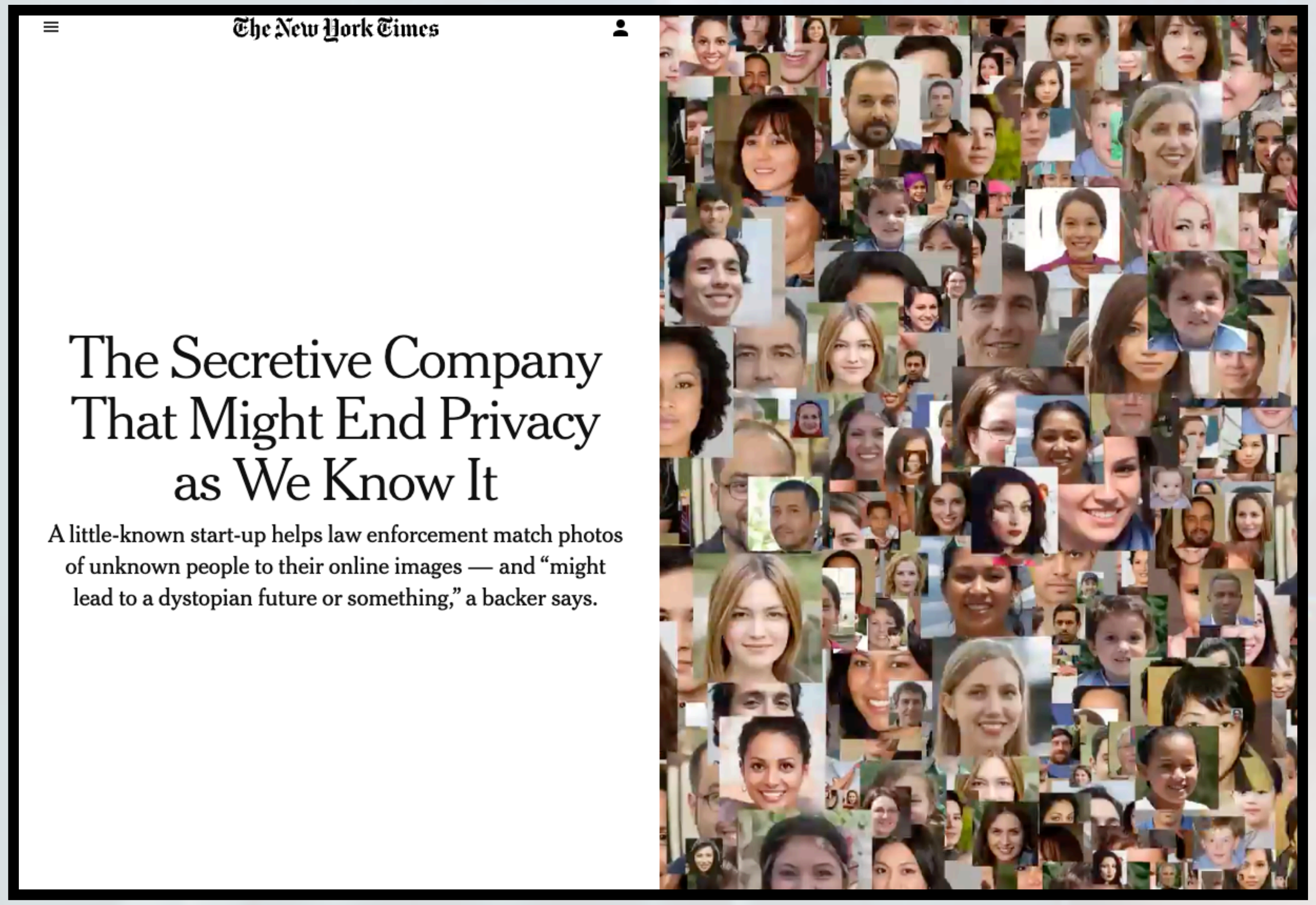
Recently, design researchers and educators began taking actions to address this problem. A few have developed designer focused education materials, meant to teach the technical concepts of ML [13, 14, 15]. This work implies that designers should understand the mechanics of algorithms in order to work effectively with ML. Other researchers created design patterns to help practitioners recognize common situations where ML can improve UX [32]. Others organized workshops, bringing together groups of artists, designers, and technologists to collectively explore how ML might function as a creative material [11, 16].

The work to make ML more accessible to designers has led some to discuss *ML as a design material* [31]. Our research adds to this growing area of inquiry. Instead of investigating problems designers face when working with ML or working on tools meant to make ML more accessible to designers, we chose to investigate the design practices of some of the few UX designers who regularly create new products and services that use ML to enhance UX. We hoped that their approach and reflections would reveal new insights around UX design education and insights on the kinds of tools needed for enhancing UX with ML.

We interviewed 13 designers who all had at least four years of experience designing ML enhanced UX. The interviews produced several interesting findings: 1) Designers shared that they knew very little about how ML works, and this was not a priority for them. They instead used designerly abstractions and popular exemplars to explain what ML is and to communicate design ideas with each other. 2) ML projects are longer in preparation and scope than other design projects. During the preparation stage, designers evolved their ideas in close collaboration with data scientists; They did not deliver fully formed designs to a technical team. 3) Designers "play" with quantitative data during all phases of a design project. Our findings suggest an alternative to the common assumption that teaching designers how ML works as the most effective way of helping them engage with it

DIS '18, June 9–13, 2018, , Hong Kong
© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. ISBN 978-1-4503-5198-0/18/06...$15.00
DOI: 10.1145/3196709.3196730

Facial recognition

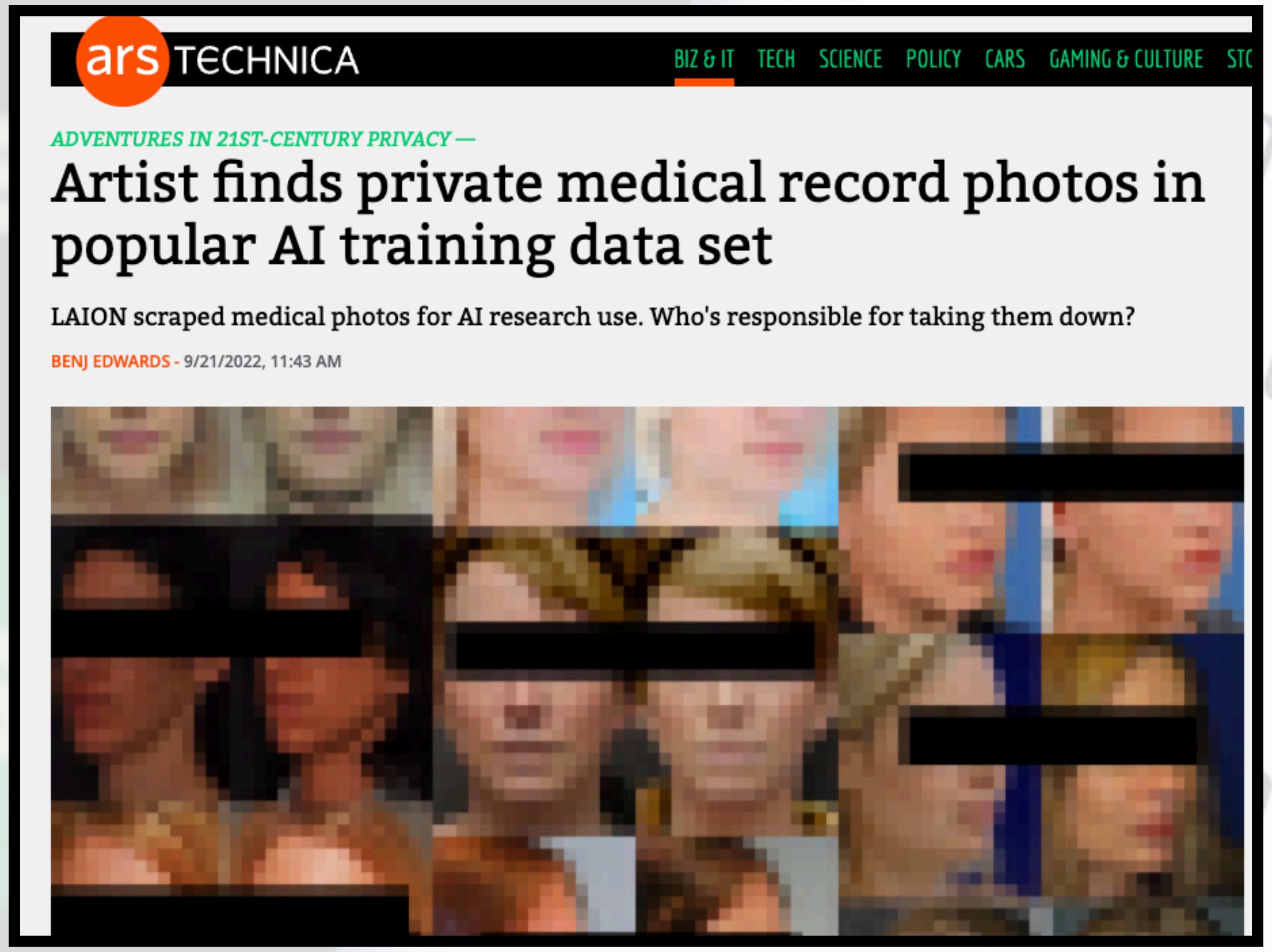


The New York Times

The Secretive Company That Might End Privacy as We Know It

A little-known start-up helps law enforcement match photos of unknown people to their online images — and "might lead to a dystopian future or something," a backer says.

Large language model



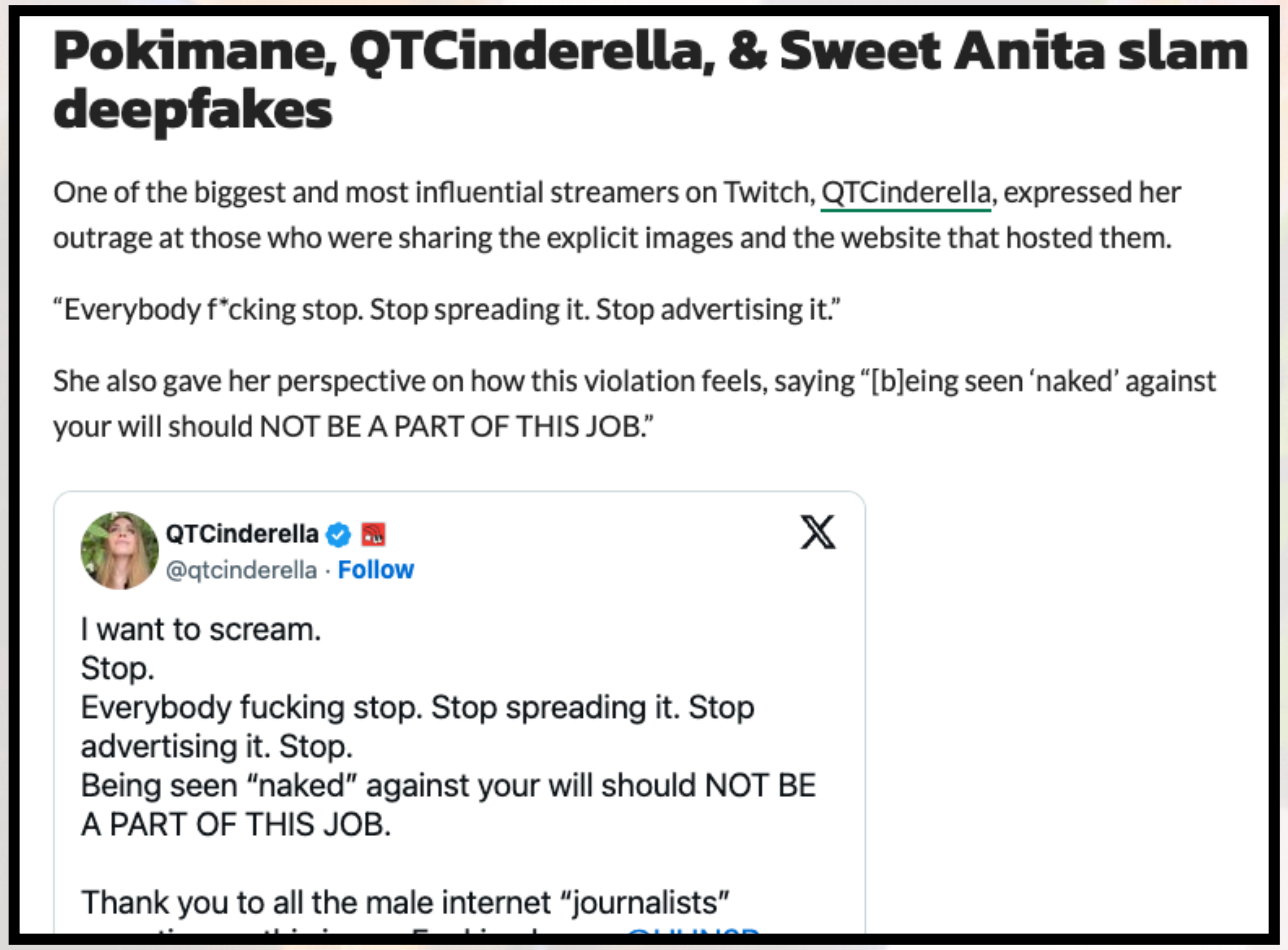
ADVENTURES IN 21ST-CENTURY PRIVACY —

Artist finds private medical record photos in popular AI training data set

LAION scraped medical photos for AI research use. Who's responsible for taking them down?

BENJ EDWARDS - 9/21/2022, 11:43 AM

Diffusion-based algorithm



Pokimane, QTCinderella, & Sweet Anita slam deepfakes

One of the biggest and most influential streamers on Twitch, QTCinderella, expressed her outrage at those who were sharing the explicit images and the website that hosted them.

"Everybody f*cking stop. Stop spreading it. Stop advertising it."

She also gave her perspective on how this violation feels, saying "[b]eing seen 'naked' against your will should NOT BE A PART OF THIS JOB."

QTCinderella @qtcinderella · Follow

I want to scream.
Stop.
Everybody fucking stop. Stop spreading it. Stop advertising it. Stop.
Being seen "naked" against your will should NOT BE A PART OF THIS JOB.

Thank you to all the male internet "journalists"

Motivation

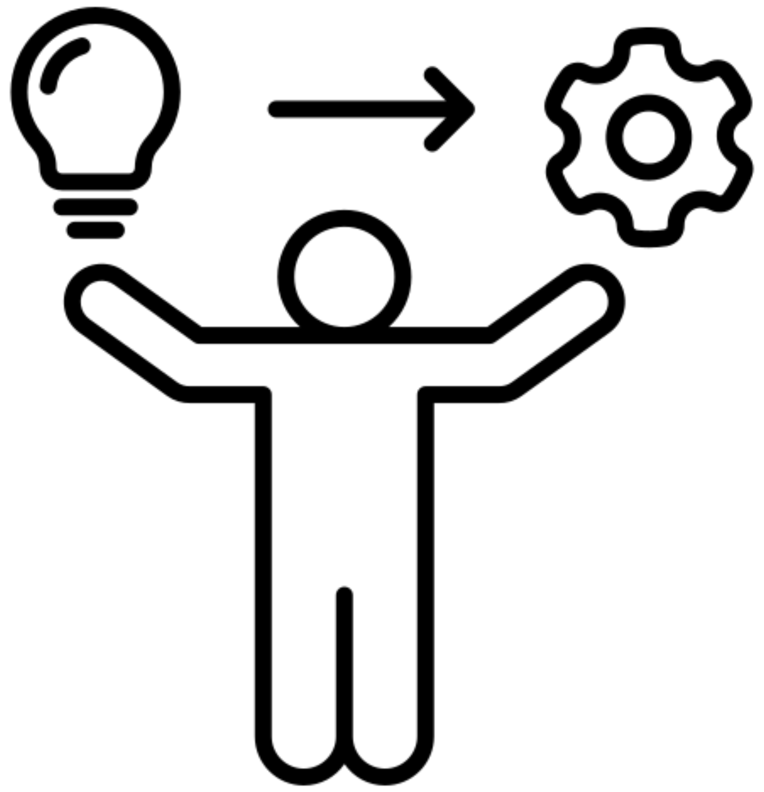
Are practitioners, who are building AI technologies, equipped to **recognize** and **mitigate** the privacy risks introduced by AI?

Potential barriers for AI practitioners to design for privacy

gap between principle and practice

unique privacy harms due to capabilities

Potential barriers for AI practitioners to design for privacy



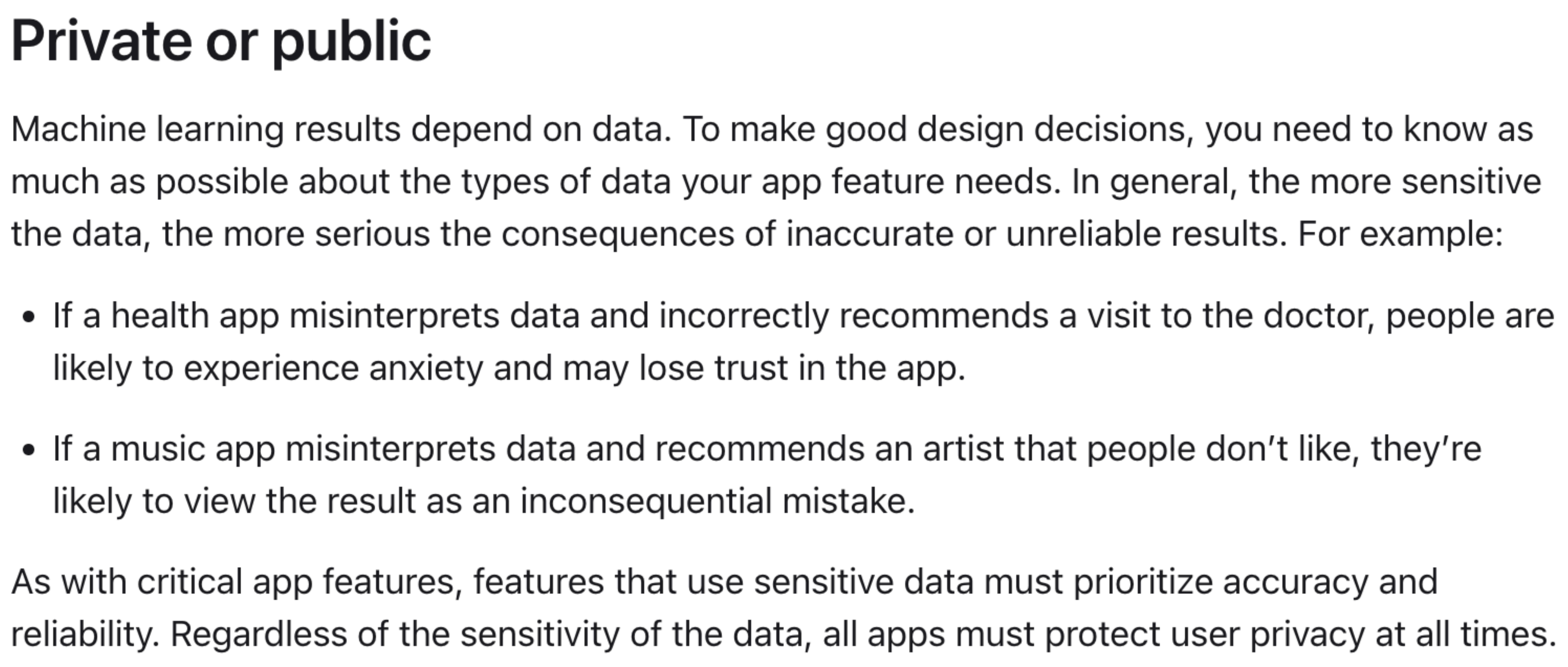
gap between principle and practice

There remains a substantial **“gap between principle and practice”** in human-centered AI [1].

[1] Shneiderman. Bridging the gap between ethics and practice: guidelines for reliable, safe, and trustwor- thy human-centered ai systems (TiiS’20)

Potential barriers for AI practitioners to design for privacy

Apple's Human Interface Guidelines for Machine Learning



Private or public

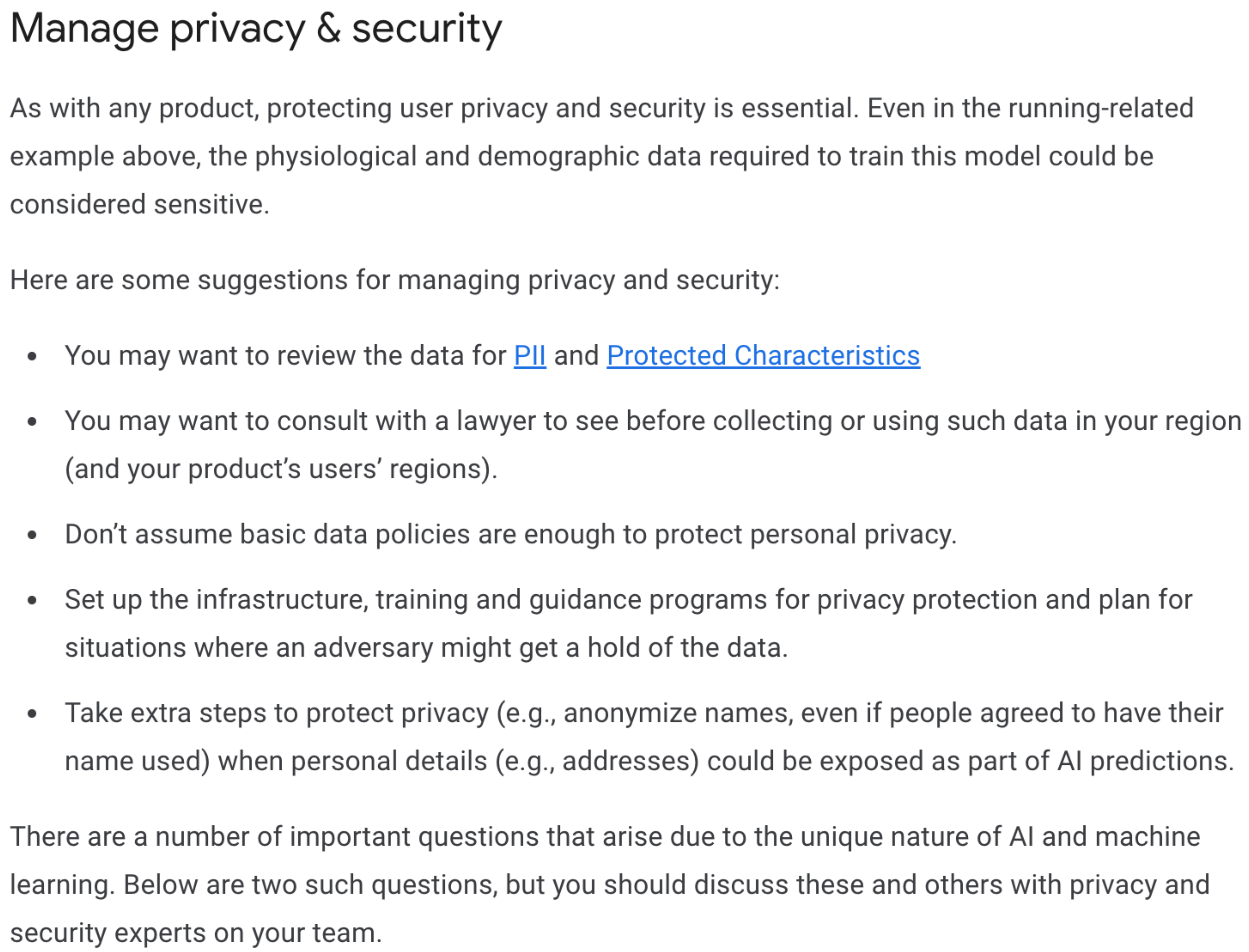
Machine learning results depend on data. To make good design decisions, you need to know as much as possible about the types of data your app feature needs. In general, the more sensitive the data, the more serious the consequences of inaccurate or unreliable results. For example:

- If a health app misinterprets data and incorrectly recommends a visit to the doctor, people are likely to experience anxiety and may lose trust in the app.
- If a music app misinterprets data and recommends an artist that people don't like, they're likely to view the result as an inconsequential mistake.

As with critical app features, features that use sensitive data must prioritize accuracy and reliability. Regardless of the sensitivity of the data, all apps must protect user privacy at all times.

"all apps must protect user privacy at all times"

Google's PAIR Guidebook



Manage privacy & security

As with any product, protecting user privacy and security is essential. Even in the running-related example above, the physiological and demographic data required to train this model could be considered sensitive.

Here are some suggestions for managing privacy and security:

- You may want to review the data for PII and Protected Characteristics
- You may want to consult with a lawyer to see before collecting or using such data in your region (and your product's users' regions).
- Don't assume basic data policies are enough to protect personal privacy.
- Set up the infrastructure, training and guidance programs for privacy protection and plan for situations where an adversary might get a hold of the data.
- Take extra steps to protect privacy (e.g., anonymize names, even if people agreed to have their name used) when personal details (e.g., addresses) could be exposed as part of AI predictions.

There are a number of important questions that arise due to the unique nature of AI and machine learning. Below are two such questions, but you should discuss these and others with privacy and security experts on your team.

"…you should discuss these and others with privacy and security experts on your team"

Potential barriers for AI practitioners to design for privacy

unique privacy harms due to capabilities

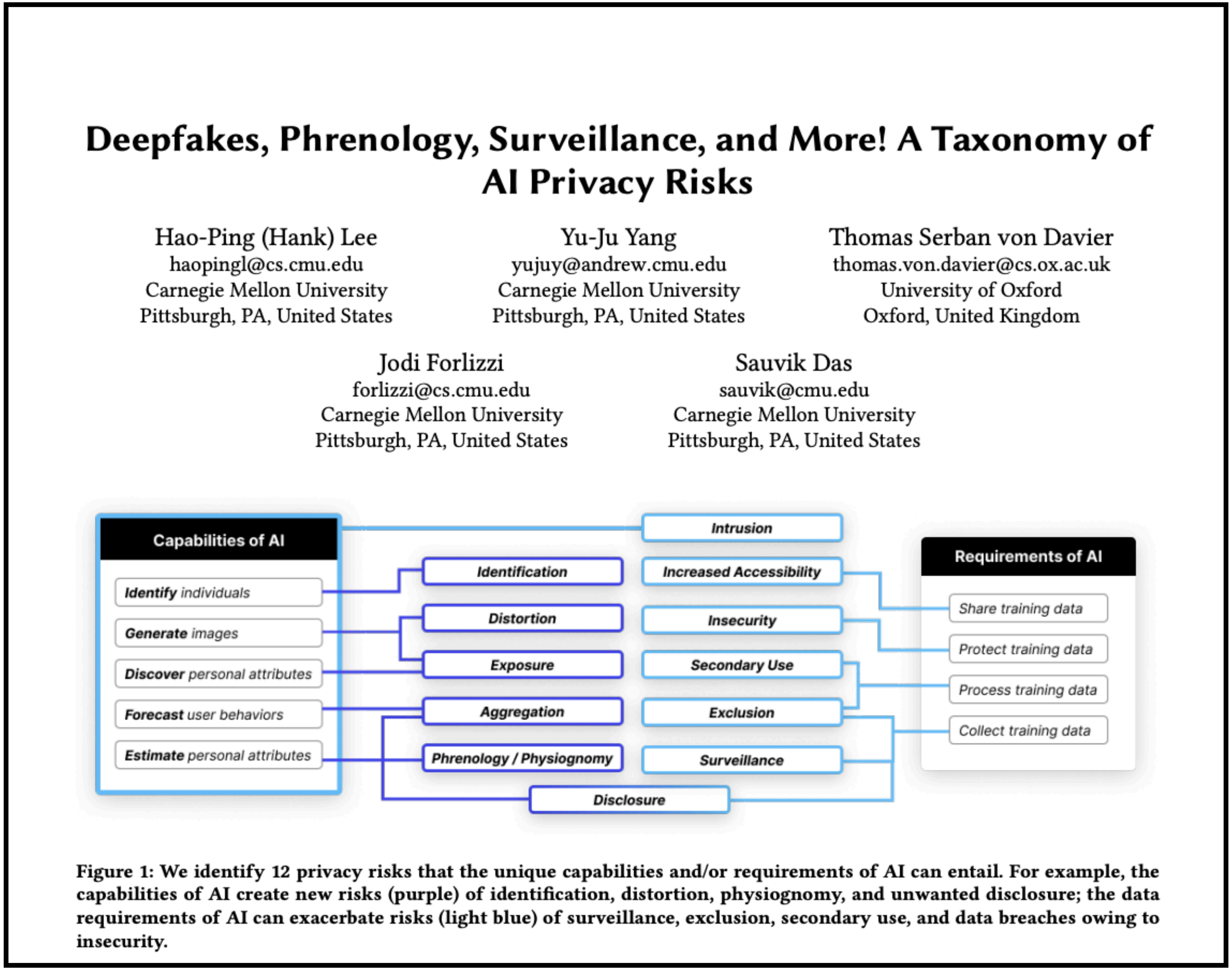
AI technologies have the potential to pose **unique privacy harms** due to its unique capabilities:

- Face recognition (e.g., [2])
- Deep fake (e.g., [3])
- Reconstructing training data (e.g., [4])

[2] *Hill. The secretive company that might end privacy as we know it (2020)*

[3] *Burgess. The Biggest Deepfake Abuse Site Is Grow- ing in Disturbing Ways (2021)*

[4] *Webster et al. This person (probably) exists. identity membership attacks against gan generated faces (2021)*



Deepfakes, Phrenology, Surveillance, and More! A Taxonomy of AI Privacy Risks

Hao-Ping (Hank) Lee
haopingl@cs.cmu.edu
Carnegie Mellon University
Pittsburgh, PA, United States

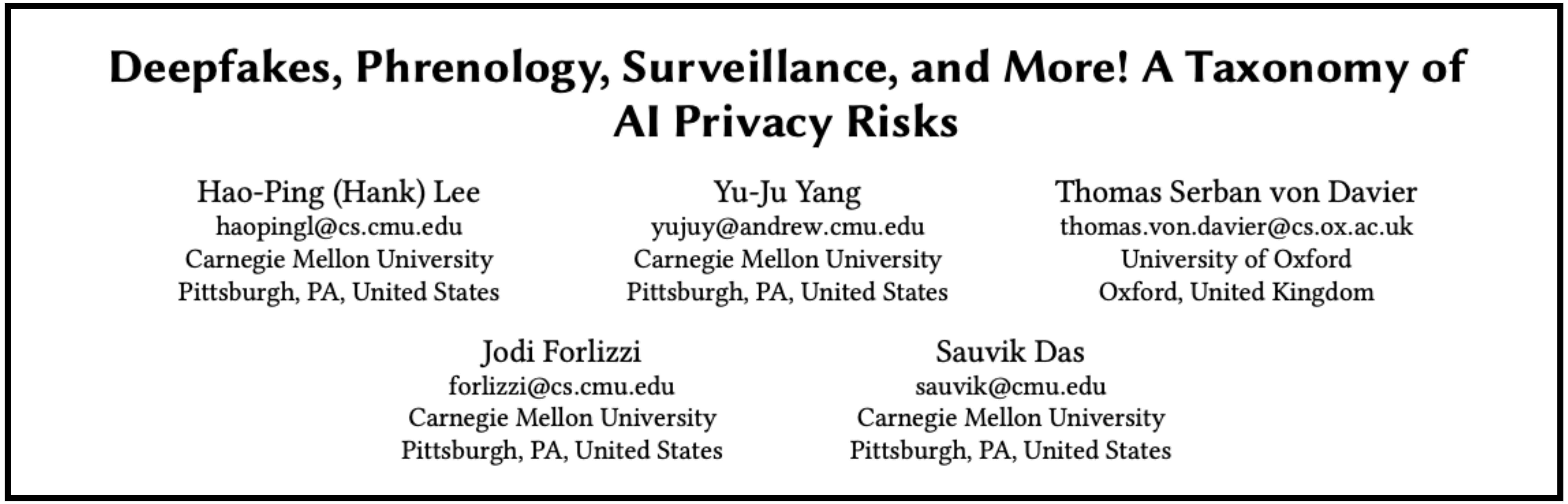
Yu-Ju Yang
yujuy@andrew.cmu.edu
Carnegie Mellon University
Pittsburgh, PA, United States

Thomas Serban von Davier
thomas.von.davier@cs.ox.ac.uk
University of Oxford
Oxford, United Kingdom

Jodi Forlizzi
forlizzi@cs.cmu.edu
Carnegie Mellon University
Pittsburgh, PA, United States

Sauvik Das
sauvik@cmu.edu
Carnegie Mellon University
Pittsburgh, PA, United States

Figure 1: We identify 12 privacy risks that the unique capabilities and/or requirements of AI can entail. For example, the capabilities of AI create new risks (purple) of identification, distortion, physiognomy, and unwanted disclosure; the data requirements of AI can exacerbate risks (light blue) of surveillance, exclusion, secondary use, and data breaches owing to insecurity.



Deepfakes, Phrenology, Surveillance, and More! A Taxonomy of AI Privacy Risks

Hao-Ping (Hank) Lee
haopingl@cs.cmu.edu
Carnegie Mellon University
Pittsburgh, PA, United States

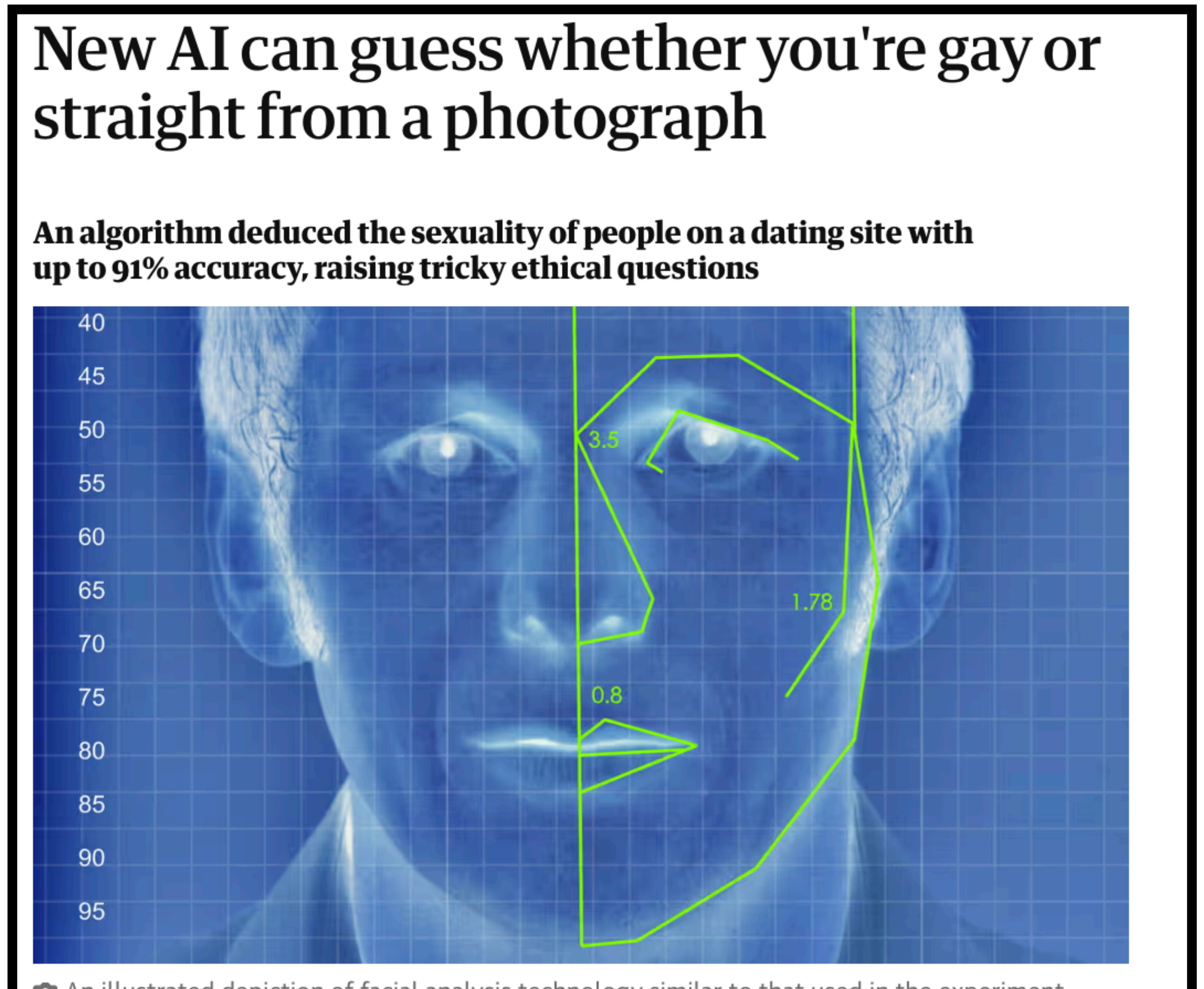
Yu-Ju Yang
yujuy@andrew.cmu.edu
Carnegie Mellon University
Pittsburgh, PA, United States

Thomas Serban von Davier
thomas.von.davier@cs.ox.ac.uk
University of Oxford
Oxford, United Kingdom

Jodi Forlizzi
forlizzi@cs.cmu.edu
Carnegie Mellon University
Pittsburgh, PA, United States

Sauvik Das
sauvik@cmu.edu
Carnegie Mellon University
Pittsburgh, PA, United States

Phrenology / Physiognomy

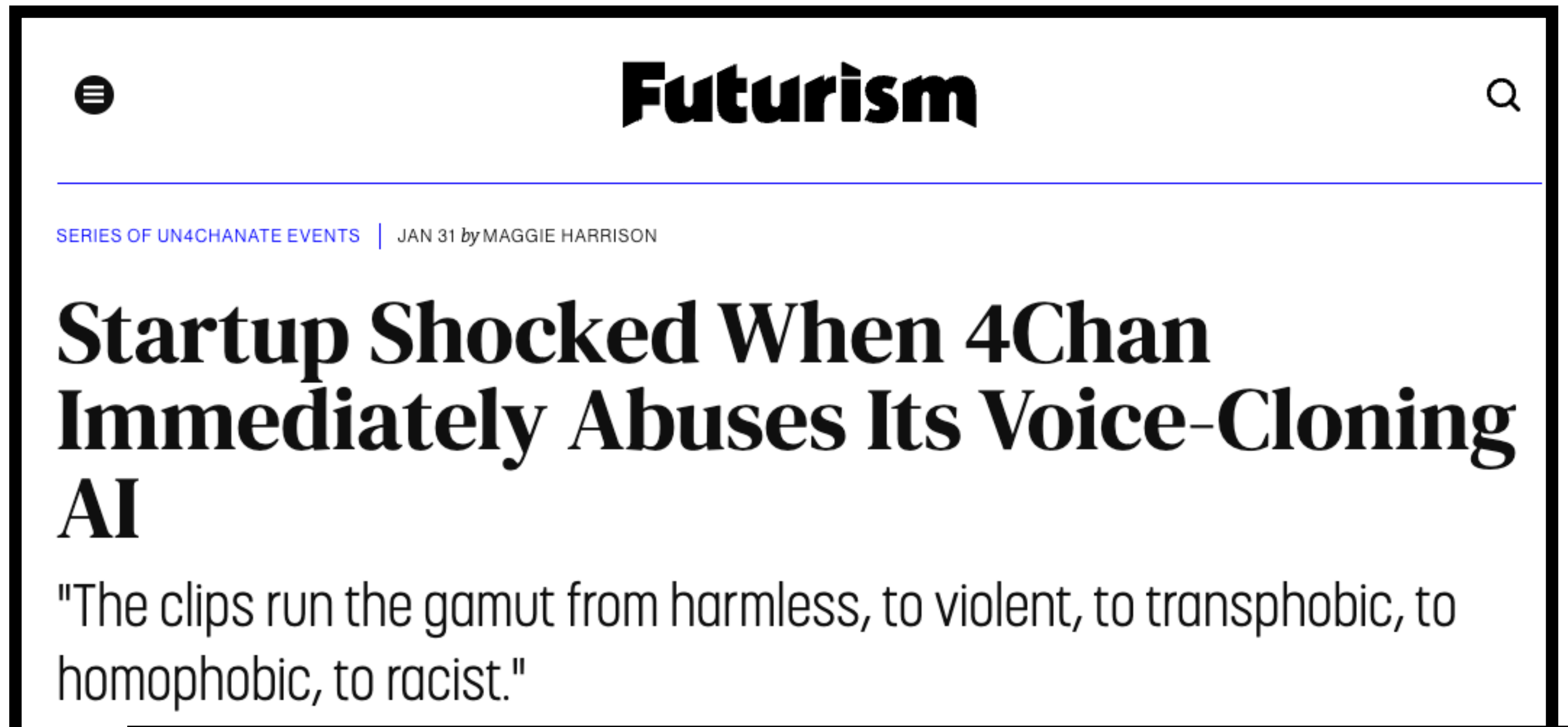


New AI can guess whether you're gay or straight from a photograph

An algorithm deduced the sexuality of people on a dating site with up to 91% accuracy, raising tricky ethical questions

An illustrated depiction of facial analysis technology similar to that used in the experiment

Distortion

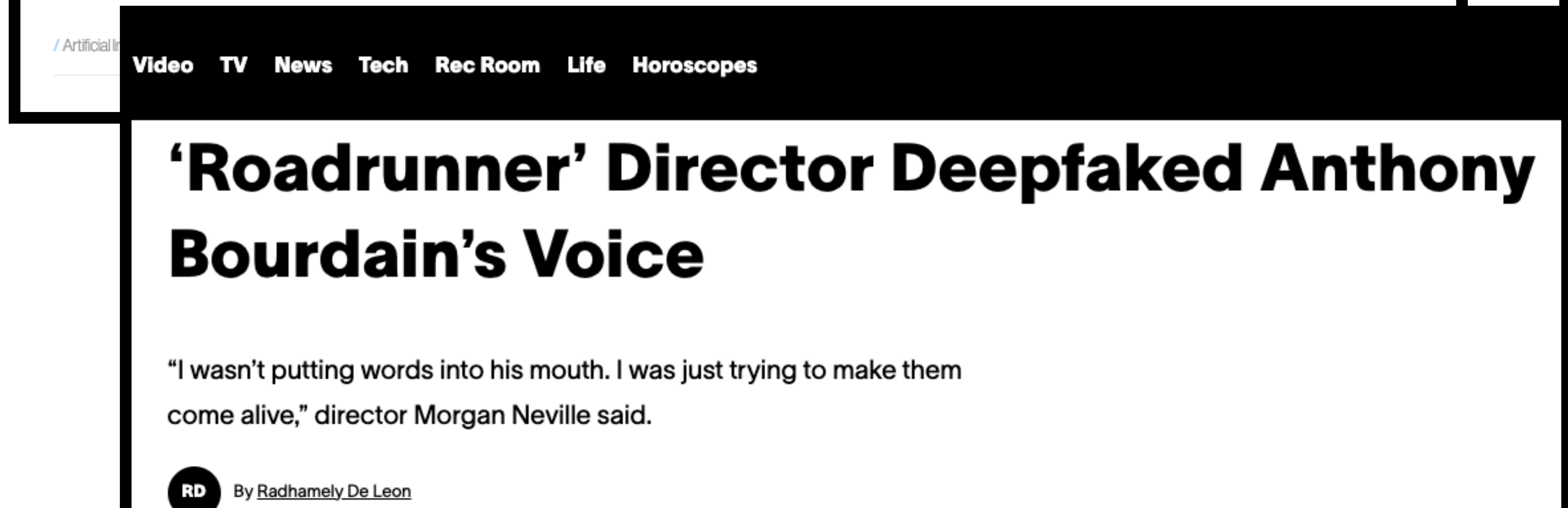


SERIES OF UN4CHANATE EVENTS | JAN 31 *by* MAGGIE HARRISON

Startup Shocked When 4Chan Immediately Abuses Its Voice-Cloning AI

"The clips run the gamut from harmless, to violent, to transphobic, to homophobic, to racist."

/ Artificial I

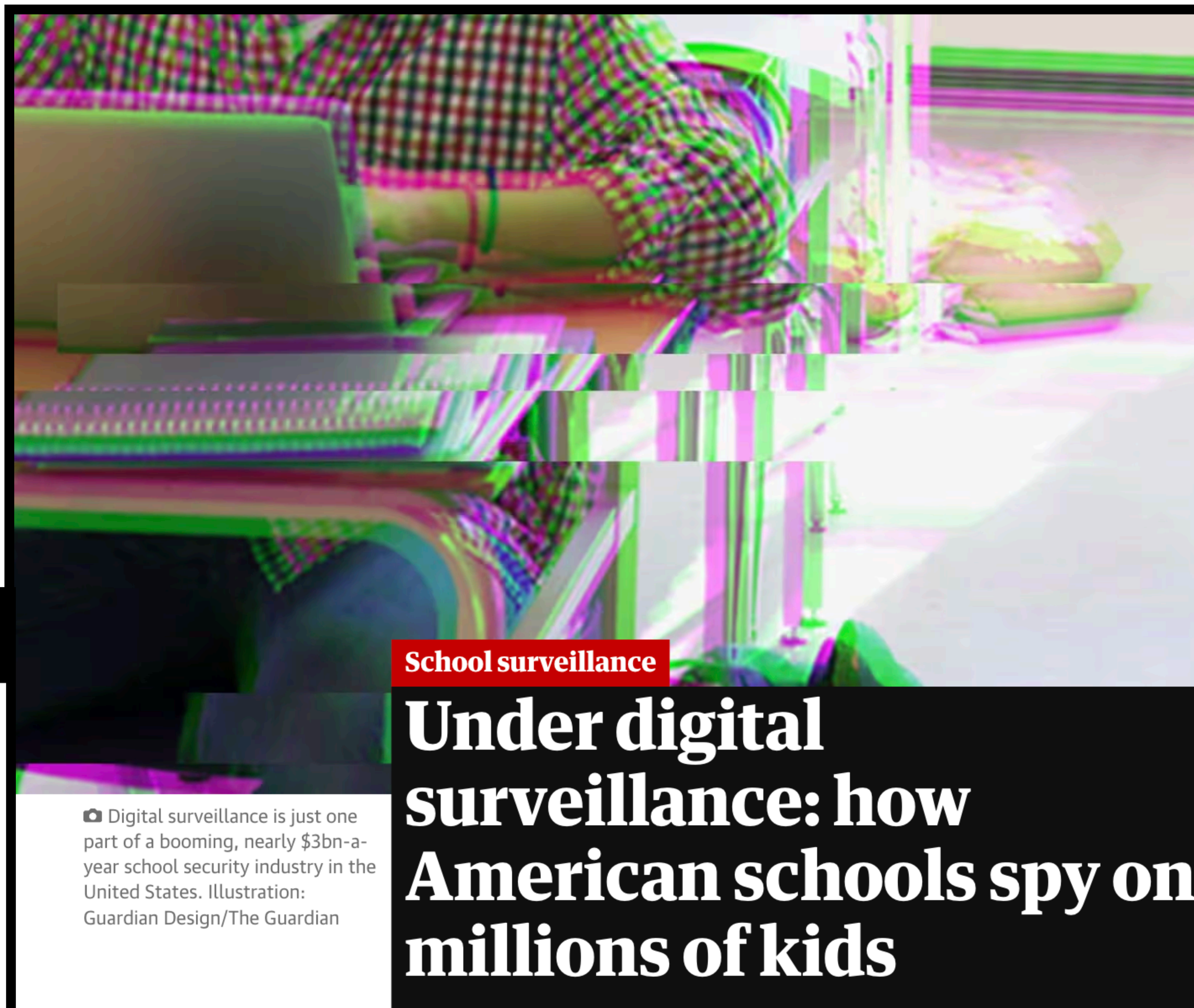


‘Roadrunner’ Director Deepfaked Anthony Bourdain’s Voice

“I wasn’t putting words into his mouth. I was just trying to make them come alive,” director Morgan Neville said.

RD By Radhamely De Leon

Surveillance



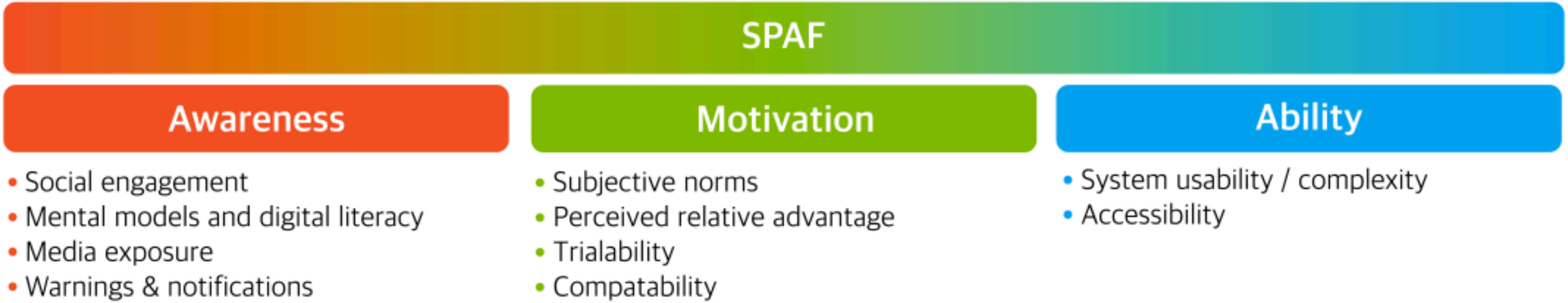
Digital surveillance is just one part of a booming, nearly $3bn-a-year school security industry in the United States. Illustration: Guardian Design/The Guardian

School surveillance

Under digital surveillance: how American schools spy on millions of kids

Analyze barriers for AI practitioners to design for privacy

Security and Privacy Acceptance Framework (SPAF)



SPAF

Awareness

- Social engagement
- Mental models and digital literacy
- Media exposure
- Warnings & notifications

Motivation

- Subjective norms
- Perceived relative advantage
- Trialability
- Compatability

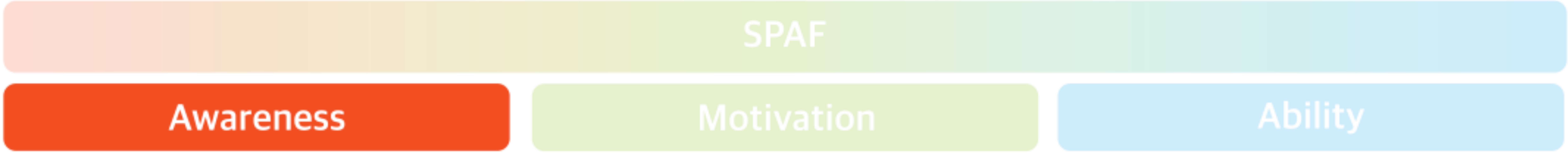
Ability

- System usability / complexity
- Accessibility

Das et al. The Security and Privacy Acceptance Framework (SPAF): A review of why users accept or reject cybersecurity and privacy best practices (2022)

Analyze barriers for AI practitioners to design for privacy

Security and Privacy Acceptance Framework (SPAF)

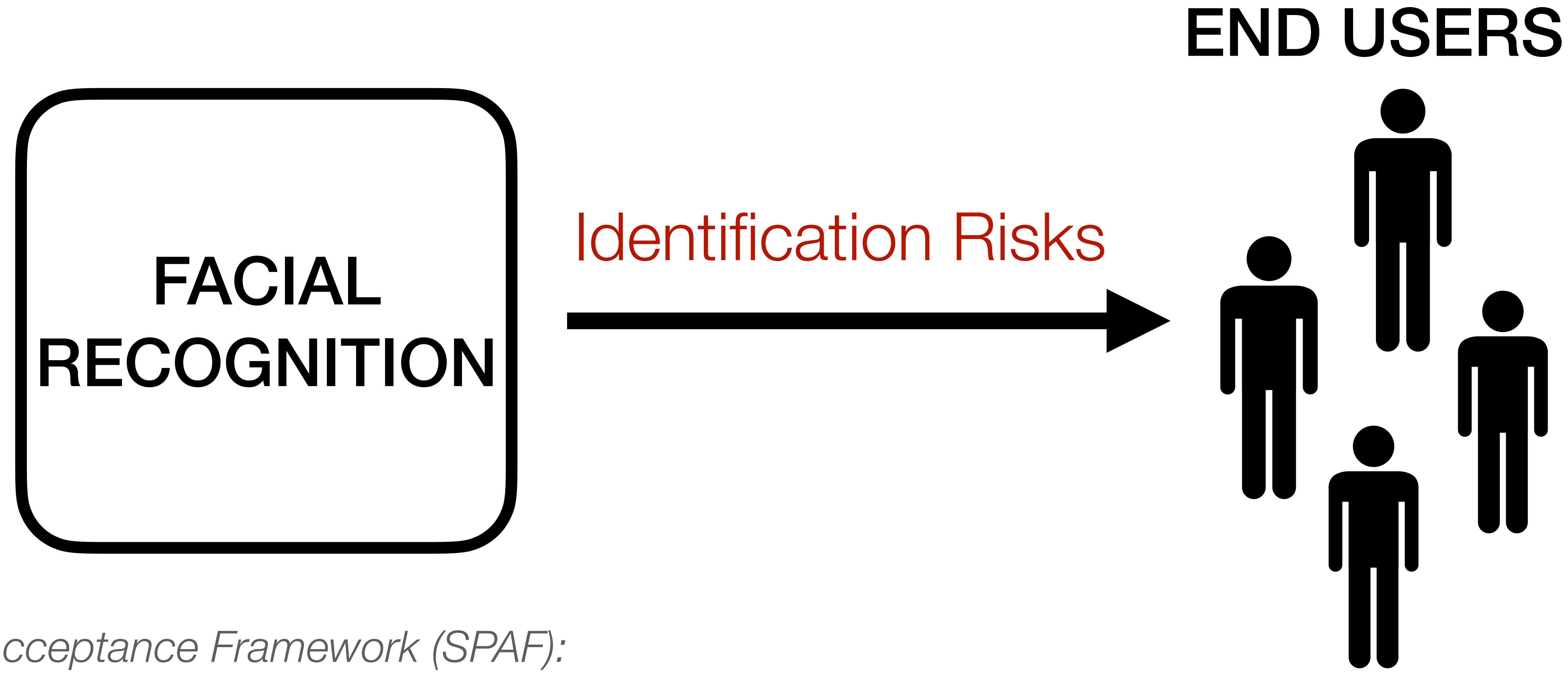


SPAF

Awareness

Motivation

Ability



FACIAL RECOGNITION

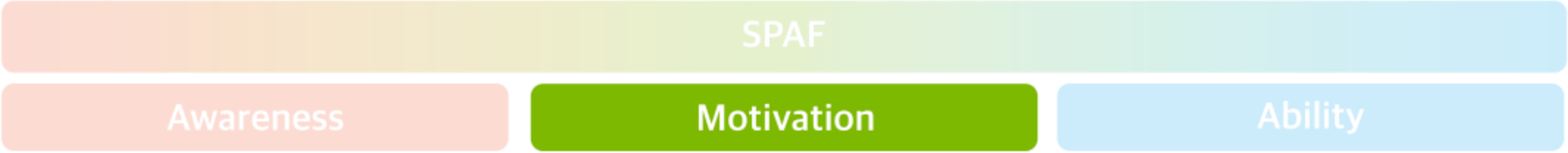
Identification Risks

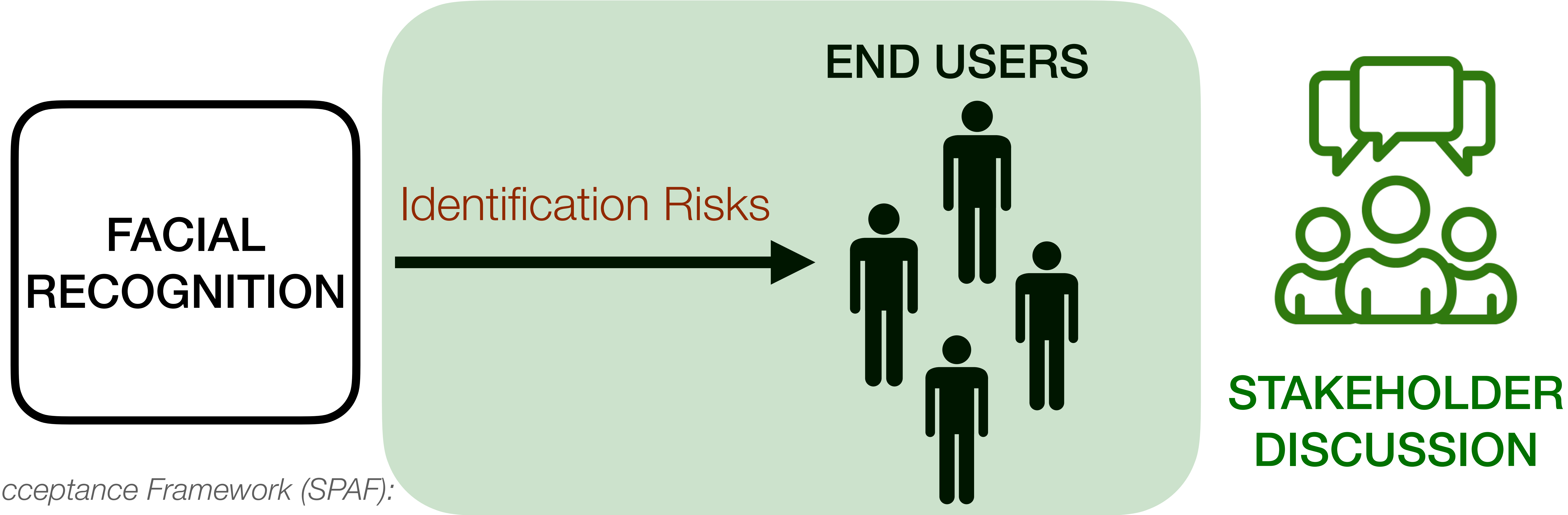
END USERS

Das et al. The Security and Privacy Acceptance Framework (SPAF): A review of why users accept or reject cybersecurity and privacy best practices (2022)

Analyze barriers for AI practitioners to design for privacy

Security and Privacy Acceptance Framework (SPAF)

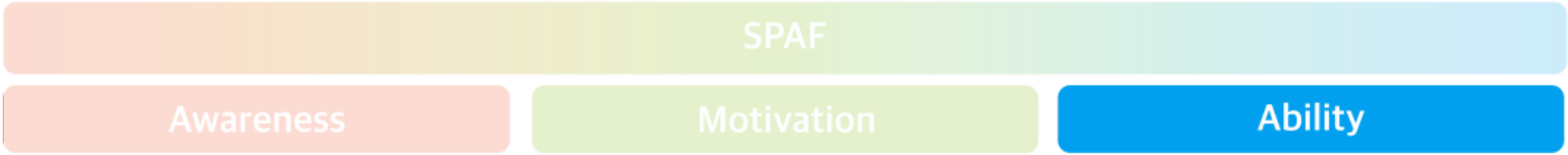




Das et al. The Security and Privacy Acceptance Framework (SPAF): A review of why users accept or reject cybersecurity and privacy best practices (2022)

Analyze barriers for AI practitioners to design for privacy

Security and Privacy Acceptance Framework (SPAF)

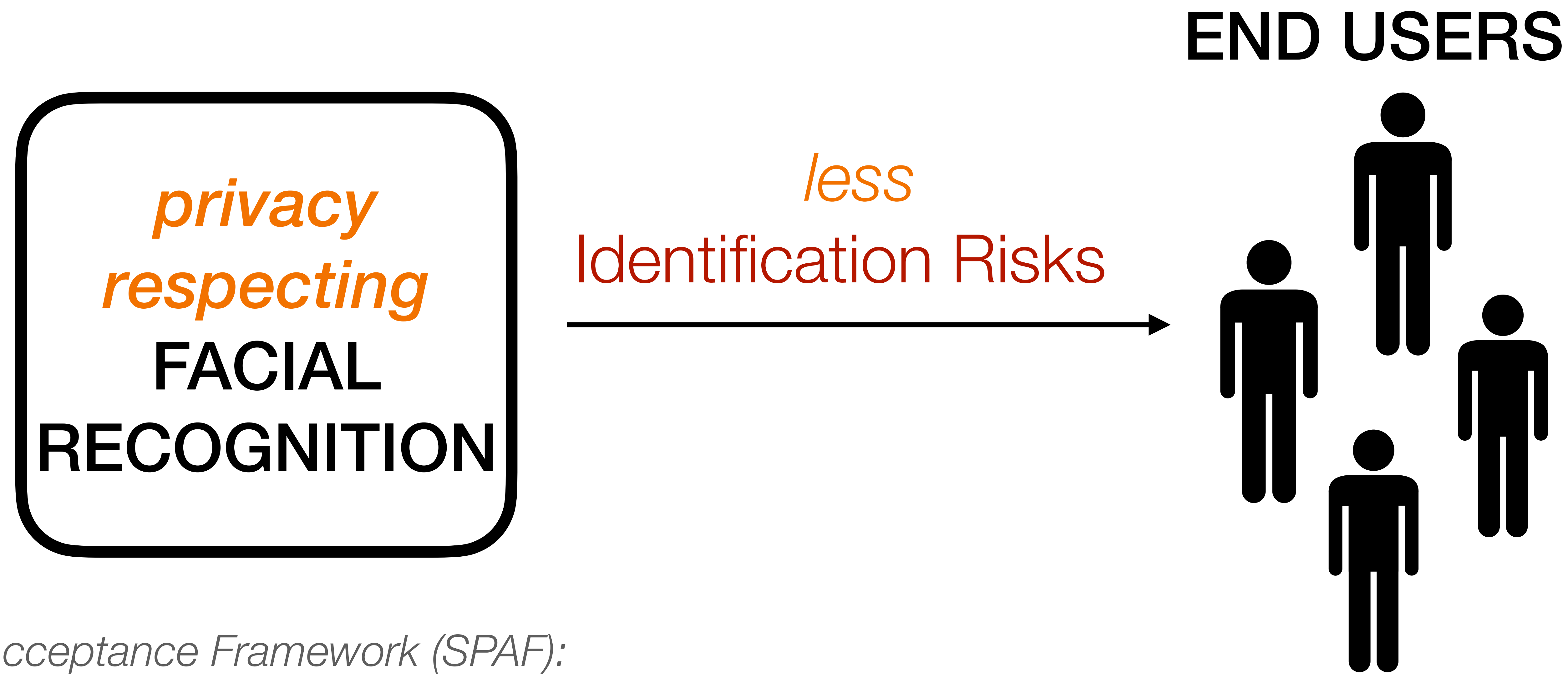


SPAF

Awareness

Motivation

Ability



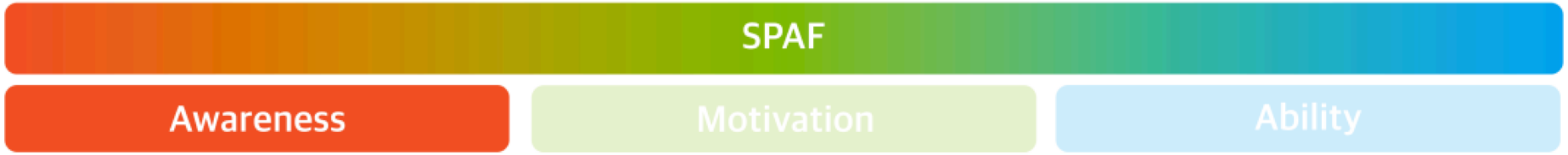
END USERS

privacy respecting FACIAL RECOGNITION

less Identification Risks

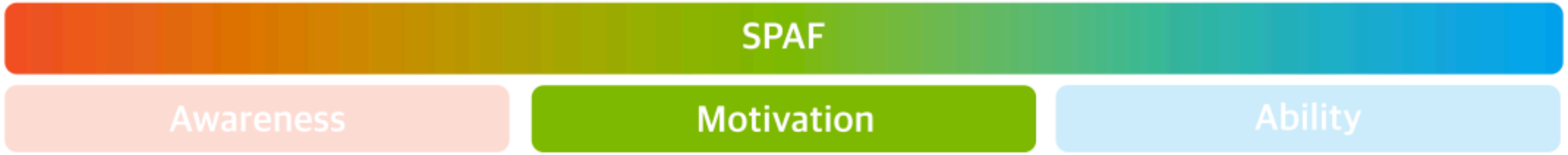
Das et al. The Security and Privacy Acceptance Framework (SPAF): A review of why users accept or reject cybersecurity and privacy best practices (2022)

Research Questions



RQ1: How well do AI practitioners' definitions of privacy work reflect ***awareness*** of AI-exacerbated privacy threats?

Research Questions



SPAF

Awareness

Motivation

Ability

RQ1: How well do AI practitioners' definitions of privacy work reflect awareness of AI-exacerbated privacy threats?

RQ2: What ***motivates*** and ***inhibits*** privacy work for consumer-facing AI products?

Research Questions

SPAF

Awareness

Motivation

Ability

RQ1: How well do AI practitioners' definitions of privacy work reflect awareness of AI-exacerbated privacy threats?

RQ2: What motivates and inhibits privacy work for consumer-facing AI products?

RQ3: What affects practitioners’ ***ability*** to do AI privacy work for consumer-facing AI products?

How well are practitioners equipped?

Semi-structured interviews with **35 AI industry practitioners** from 25 different companies with diverse roles (e.g., product manager, engineers, designers, and researchers), and work on different consumer-facing AI products (e.g., chatbots, recommenders, computer vision) and domains (e.g., healthcare, marketing, media & entertainment)

All of them have participated in **discussions about end-user privacy** related to the consumer-facing AI products/services that they have helped build.

Awareness

How do AI practitioners define privacy work?

Identification:

[P8, Tech Lead, ML Dev Tool]: *“you should only be able to analyze things in aggregate manners, and not be able to do that root cause to a single point that's potentially causing a behavior.”*

Secondary use:

[P3, Software Engineer, Recommendation System]: *“we want people’s information that they give us to be safe and not used for anything else other than actually recommending them clothes.”*

Insecurity:

[P28, Tech Lead, Document Co-pilot Tool]: *“any data that they [users] contribute to the product, that the lifecycle of that data is protected in some way”*

Awareness

👉 Key upshot

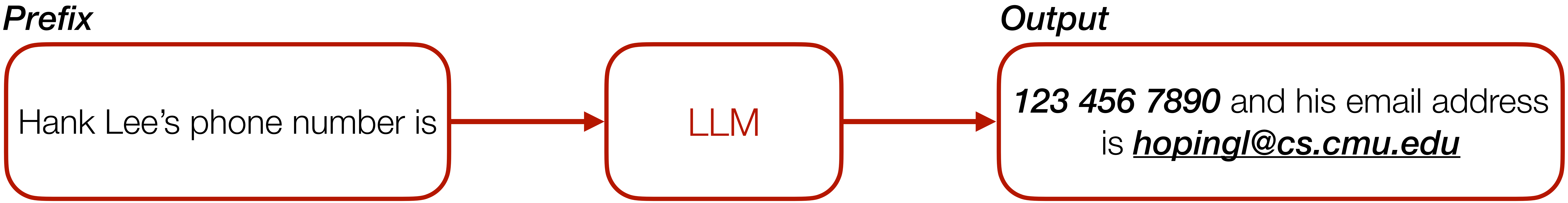
Our participants **exhibited limited awareness** of how the capabilities and requirements of AI might affect the privacy threats entailed by a product.

The structures in place for practitioners to **think about privacy for products remain generic and non-specific to AI**.

Awareness

👉 Key upshot

Our participants **exhibited limited awareness** of how the capabilities and requirements of AI might affect the privacy threats entailed by a product.



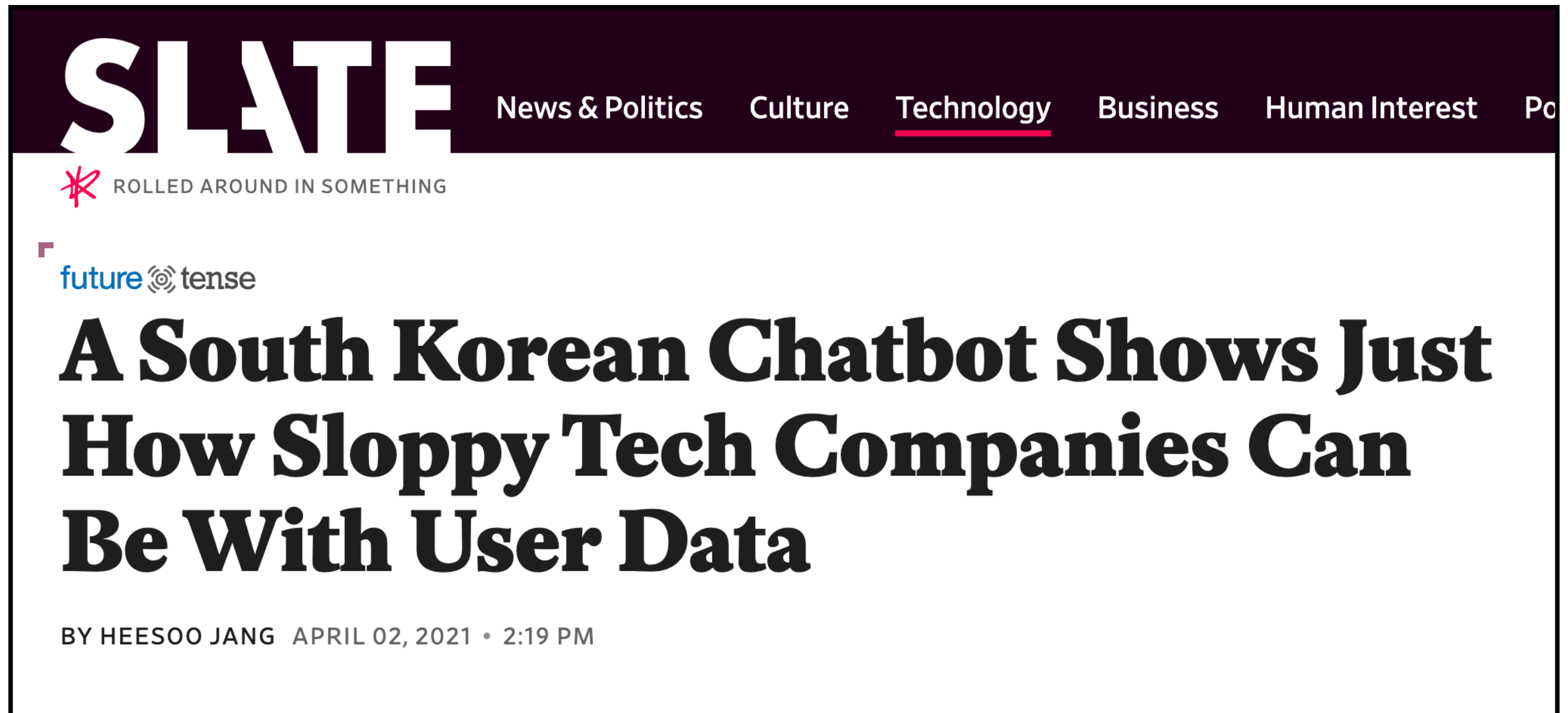
Prefix

Hank Lee's phone number is

LLM

Output

123 456 7890 and his email address is ***hopingl@cs.cmu.edu***



ROLLED AROUND IN SOMETHING

future tense

A South Korean Chatbot Shows Just How Sloppy Tech Companies Can Be With User Data

BY HEESOO JANG APRIL 02, 2021 • 2:19 PM

Peris et al., Privacy in the Time of Language Models. (WSDM '23)

Jang. Slate (April 2021)

Motivation

Motivators and inhibitors for AI privacy work

Motivators for AI privacy work

- Alignment with business interests
- Social responsibility
- Compliance

Inhibitors for AI privacy work

- Rigid compliance requirements
- Incentives
- Power in organizational structures
- Privacy education
- Ownership
- Opportunity costs

Motivators for AI privacy work

Alignment with business interests:

[P28, Tech Lead, Document Co-pilot Tool]: *“privacy can be a differentiator. And when you're doing a startup, especially if you're in a crowded space, you're looking for any way, any angle that you have to say that you're different from other things that are out there.”*

Social responsibility:

[P35, Researcher, ML building tools]: *“people that tend to come here that are building new ML features are aware of bad cases of ML being... inappropriately applied. And no one wants to have that happen.”*

Compliance:

[P31, Designer, AI App Dev Platform]: *“[privacy work] are considered compliance... we don't do it by choice, like it's always enforced.”*

Motivation

Motivators and inhibitors for AI privacy work

Motivators for AI privacy work

- Alignment with business interests
- Social responsibility
- Compliance

Inhibitors for AI privacy work

- Rigid compliance requirements
- Incentives
- Power in organizational structures
- Privacy education
- Ownership
- Opportunity costs

Inhibitors for AI privacy work

Rigid compliance requirements

the product was already “compliant”

[P16, Designer, ML Dev Tool]: *“in general product development, and what the engineers are doing, it's so standardized, that's not really a conversation, because there's nothing to be done about it. It just is the way that it is.”*

Incentives

[P11, Tech Lead, Recommendation System]: *“people are not really incentivized to do this correctly. And if they wanted to do things correctly, it becomes extra effort, and influences their completed work, fewer results, and as a result they get promoted slower than their peers.”*

Opportunity costs: model performance

[P5, Software Engineer, Recommendation System]: *“we are getting less and less idea about, for example, what an end-user is like, if having a higher standard of privacy.”*

Motivation

👉 Key upshot

AI privacy work is driven by meeting **non-AI specific compliance** standards.

Motivators for AI privacy work

- Alignment with business interests (9/35)
- Social responsibility (5/35)
- **Compliance (19/35)**

Motivation

👉 Key upshot

Practitioners currently **face many more inhibitors than motivators** for privacy work in developing consumer AI products

Motivators for AI privacy work

- Alignment with business interests
- Social responsibility
- Compliance

Inhibitors for AI privacy work

- Rigid compliance requirements
- Incentives
- Power in organizational structures
- Privacy education
- Ownership
- Opportunity costs

Ability

What constitutes AI privacy work?

Privacy value negotiations

Privacy training

Design references & compliance consultations

Developer tools & artifacts

Privacy training

[P2, Software Engineer, Speech Recognition]: *“[trainings have] nothing but a general concept… They don’t care if you work for [different types of products]. They just give everyone a very high-level idea”*

Design references & Compliance consultations

[P33, UX Researcher, Chatbot]: *“some of it is examples of what other teams have done... there's like learnings from other groups that we can take advantage of... like, how do other teams collect terms of service, or how do other teams do platform agreements?”*

[P31, ML Engineer, Chatbot]: *“we refer to our legal experts whenever we are confused or when we feel we don't know if we're doing the right thing.”*

Ability

👉 Key upshot

Practitioners often **lack a holistic view of the data pipeline**.

[P34, Product Director, Job Matching Tool]: *"<u>The technology is often really sophisticated</u>, and so sometimes the data is leaving your AWS account, sometimes it's not. All kinds of AI and policies control, like who can and can't see that data... And so <u>it becomes difficult [to] tease out the true risk</u>."*

Practitioners **lack guidance** but must rely on individual judgment.

[P16, Designer, ML Dev Tool]: *"<u>I'm more doing computer vision stuff. It's pretty new, and so not a lot of people have the answer</u>... it kind of comes down to making my own [decision], and to know what's going to be good, or risk compliance issues."*

Ability

👉 Key upshot

Practitioners often **lack a holistic view of the data pipeline**.

Practitioners **lack guidance** but must rely on individual judgment.

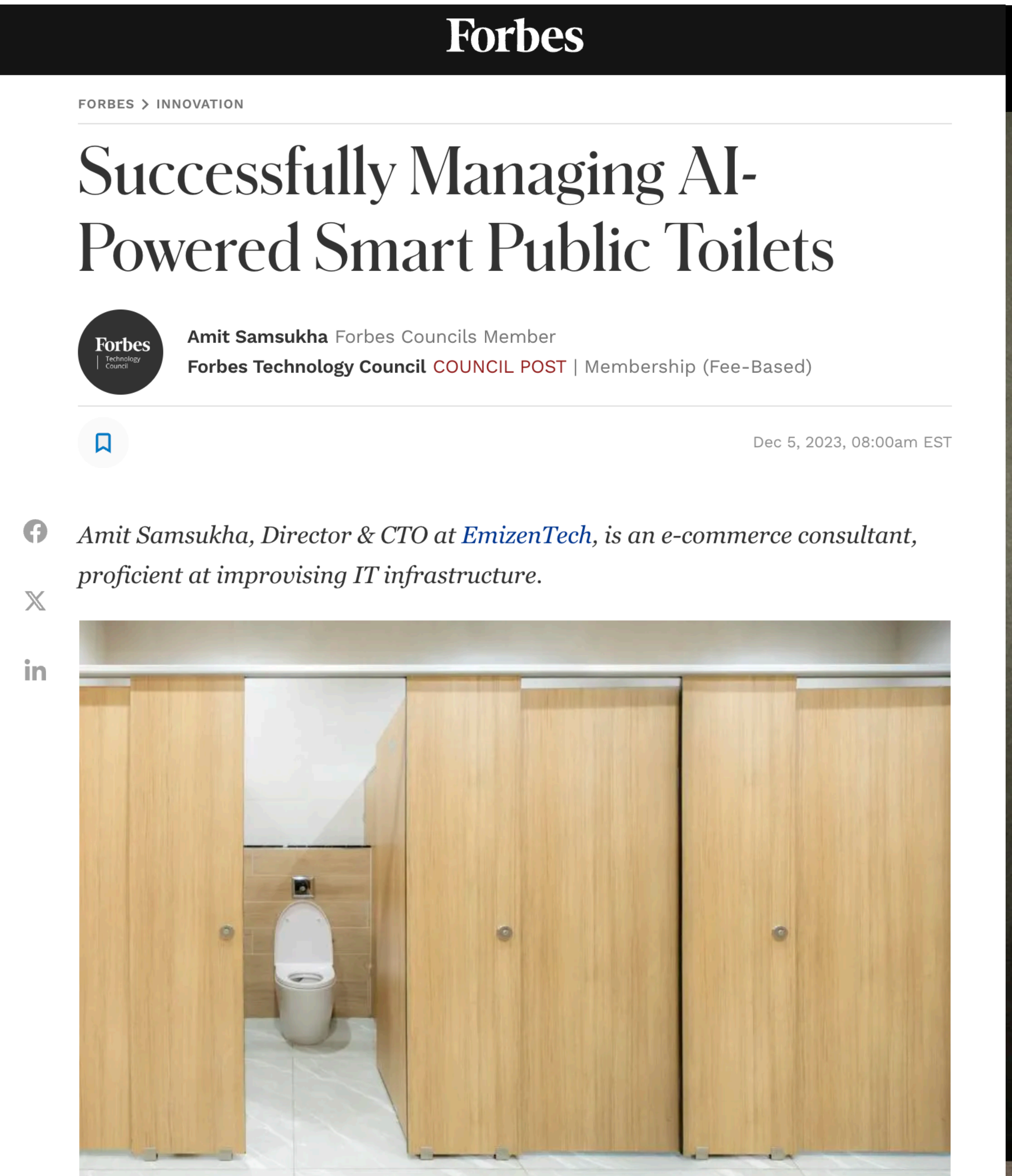
Practitioners lacked the tools, resources, and support needed to approach AI privacy work

Improving practitioners’ awareness of AI-exacerbated privacy threats

Many practitioners are not aware of the potential risks to privacy because the lack of **educational materials on AI-specific privacy topics**

Improving practitioners’ awareness of AI-exacerbated privacy threats

Many practitioners are not aware of the potential risks to privacy because the lack of **educational materials on AI-specific privacy topics**



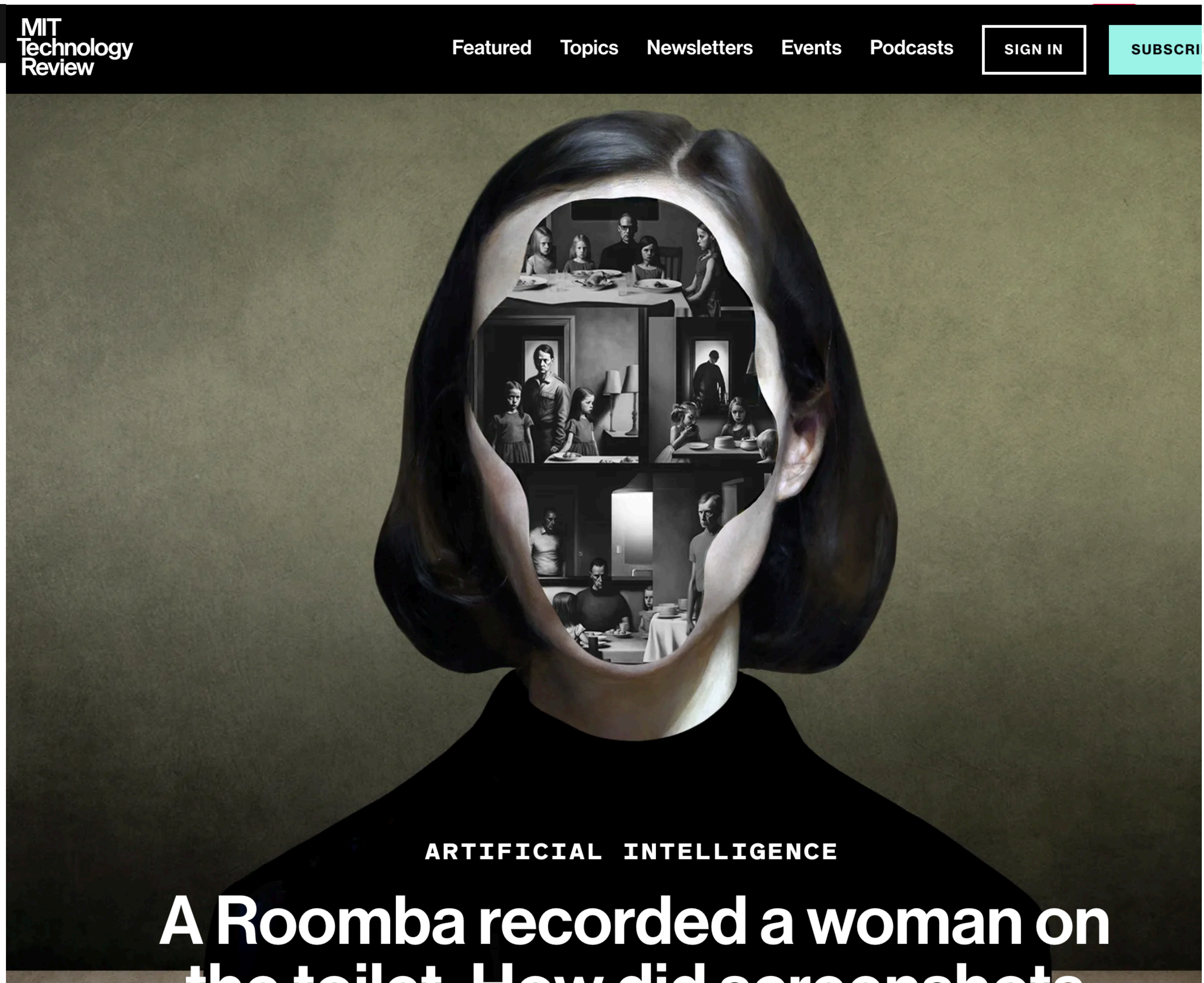
FORBES > INNOVATION

Successfully Managing AI-Powered Smart Public Toilets

Amit Samsukha Forbes Councils Member
Forbes Technology Council COUNCIL POST | Membership (Fee-Based)

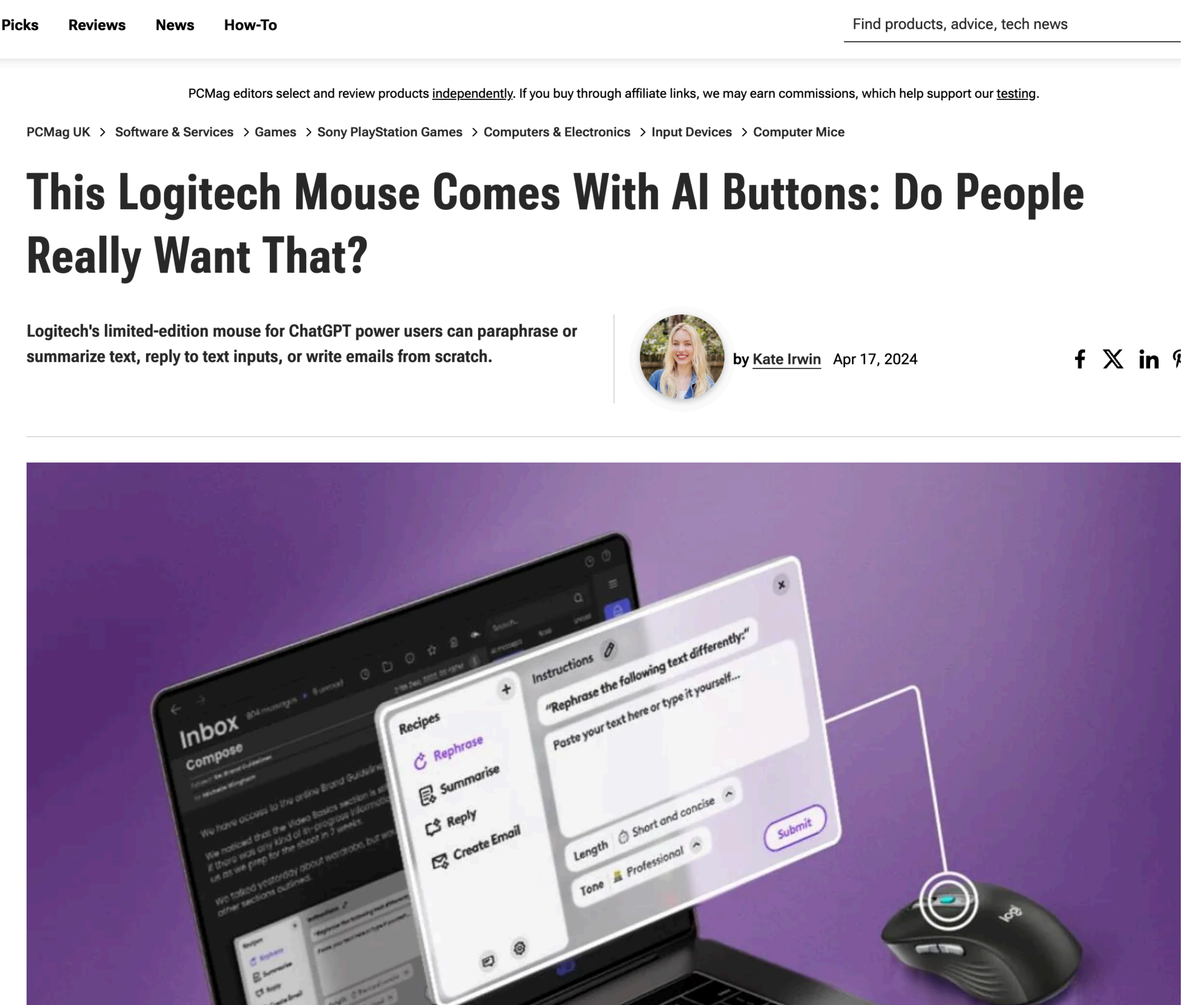
Dec 5, 2023, 08:00am EST

Amit Samsukha, Director & CTO at EmizenTech, is an e-commerce consultant, proficient at improvising IT infrastructure.



ARTIFICIAL INTELLIGENCE

A Roomba recorded a woman on the toilet. How did screenshots



PCMag editors select and review products independently. If you buy through affiliate links, we may earn commissions, which help support our testing.

PCMag UK > Software & Services > Games > Sony PlayStation Games > Computers & Electronics > Input Devices > Computer Mice

This Logitech Mouse Comes With AI Buttons: Do People Really Want That?

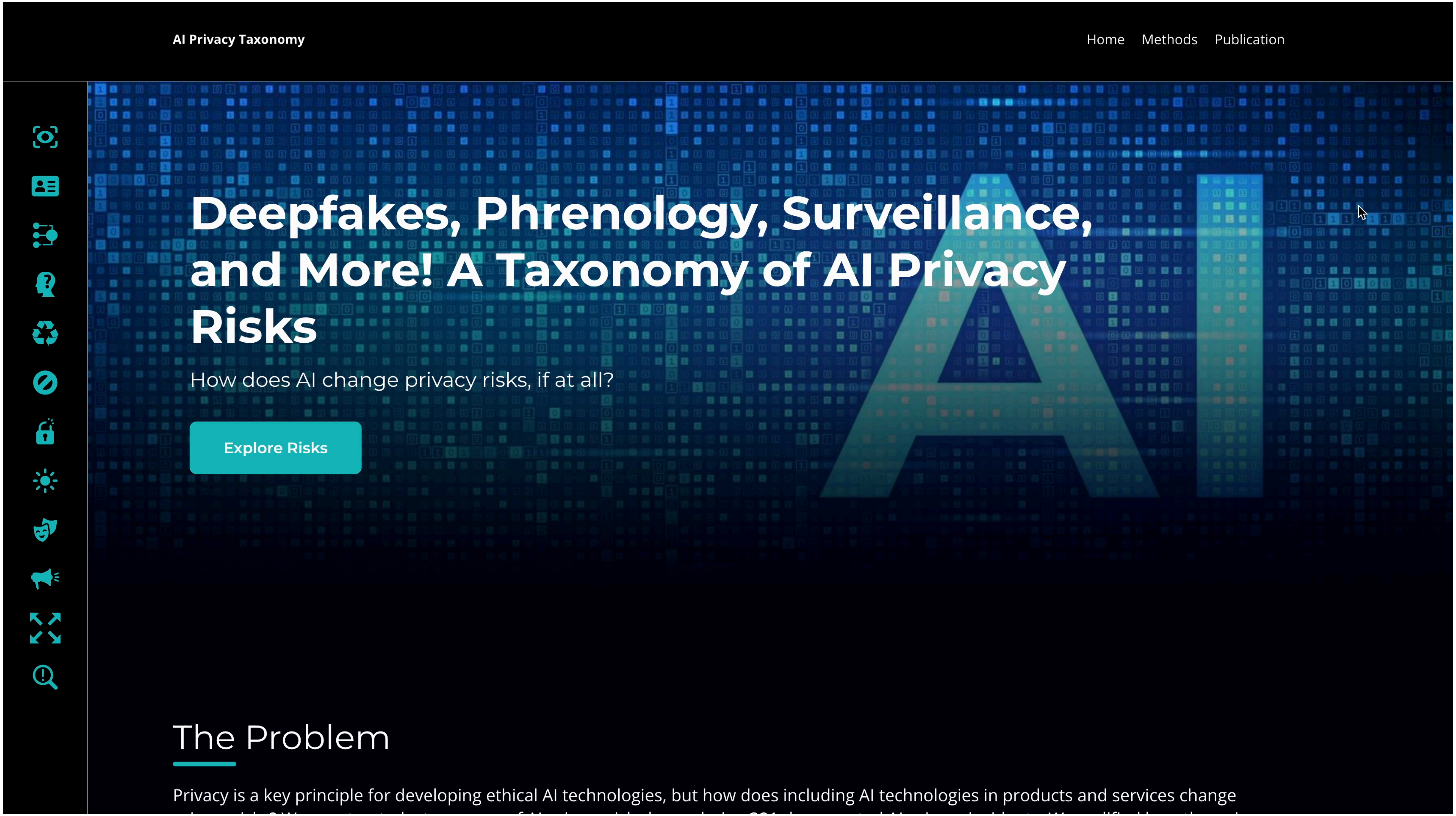
Logitech's limited-edition mouse for ChatGPT power users can paraphrase or summarize text, reply to text inputs, or write emails from scratch.

by Kate Irwin Apr 17, 2024

Improving practitioners' awareness of AI-exacerbated privacy threats

Many practitioners are not aware of the potential risks to privacy because the lack of **educational materials on AI-specific privacy topics**

AI-specific training campaigns may be effective at raising practitioners' awareness of how AI technologies might entail privacy threats for a consumer product



Deepfakes, Phrenology, Surveillance, and More! A Taxonomy of AI Privacy Risks

How does AI change privacy risks, if at all?

The Problem

Privacy is a key principle for developing ethical AI technologies, but how does including AI technologies in products and services change

https://AIPrivacyTaxonomy.com/

Improving practitioners’ ability to address AI-exacerbated privacy threats

Practitioners were required to use generic tools and methods

Privacy concerns and challenges
unique or exacerbated by AI

Mismatch

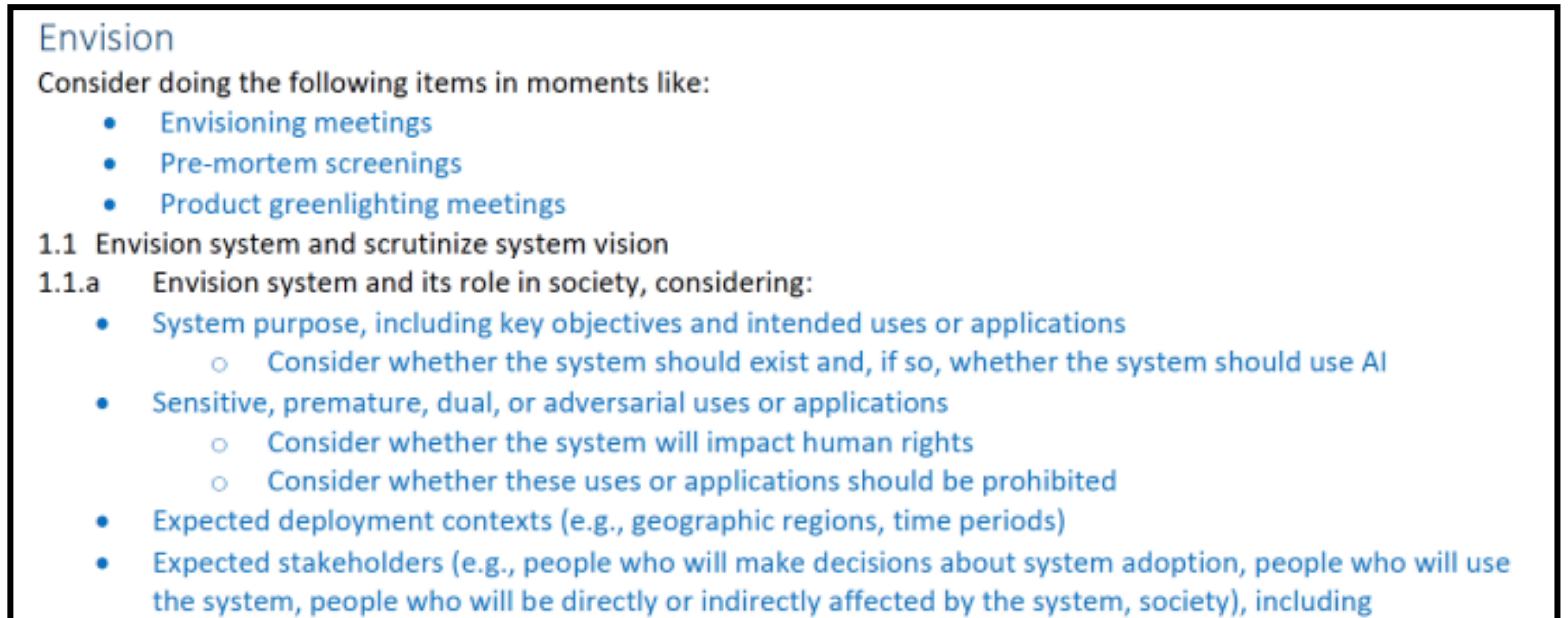
Tools and procedures for privacy
generic and not tailored toward AI

Improving practitioners' ability to address AI-exacerbated privacy threats

Practitioners were required to adapt generic tools and methods

Practitioners need **AI-specific privacy design assessment and design tools**

AI fairness checklist (Madaio et al., 2020)



Envision

Consider doing the following items in moments like:

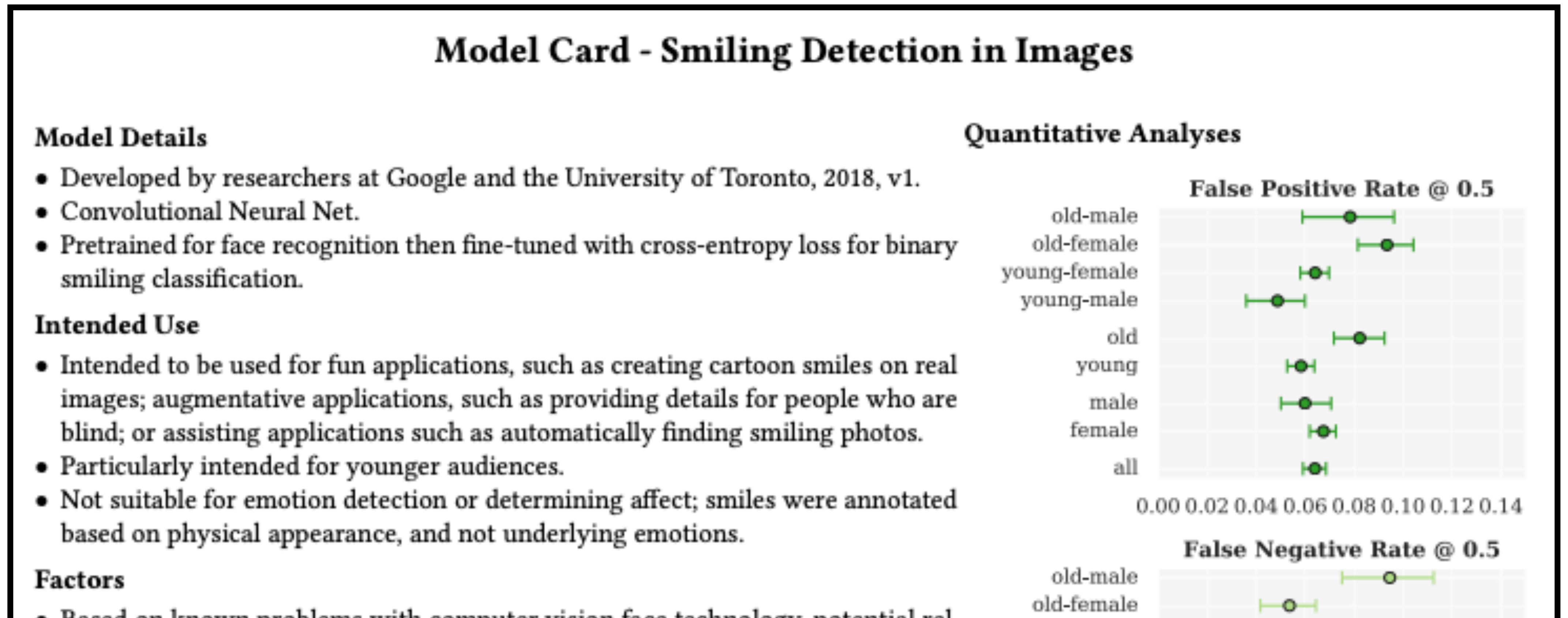
- Envisioning meetings
- Pre-mortem screenings
- Product greenlighting meetings

1.1 Envision system and scrutinize system vision

1.1.a Envision system and its role in society, considering:

- System purpose, including key objectives and intended uses or applications
  - ○ Consider whether the system should exist and, if so, whether the system should use AI
- Sensitive, premature, dual, or adversarial uses or applications
  - ○ Consider whether the system will impact human rights
  - ○ Consider whether these uses or applications should be prohibited
- Expected deployment contexts (e.g., geographic regions, time periods)
- Expected stakeholders (e.g., people who will make decisions about system adoption, people who will use the system, people who will be directly or indirectly affected by the system, society), including

Model card (Mitchell et al., 2019)



False Positive Rate @ 0.5
old-male
old-female
young-female
young-male
old
young
male
female
all
0.00 0.02 0.04 0.06 0.08 0.10 0.12 0.14
False Negative Rate @ 0.5
old-male
old-female

Model Card - Smiling Detection in Images

Model Details

- Developed by researchers at Google and the University of Toronto, 2018, v1.
- Convolutional Neural Net.
- Pretrained for face recognition then fine-tuned with cross-entropy loss for binary smiling classification.

Intended Use

- Intended to be used for fun applications, such as creating cartoon smiles on real images; augmentative applications, such as providing details for people who are blind; or assisting applications such as automatically finding smiling photos.
- Particularly intended for younger audiences.
- Not suitable for emotion detection or determining affect; smiles were annotated based on physical appearance, and not underlying emotions.

Factors

- Based on known problems with computer vision face technology, potential rel

Quantitative Analyses

*Are practitioners, who are building AI technologies, equipped to **recognize** and **mitigate** the privacy risks introduced by AI?*

👀 **Awareness**: privacy is viewed as protecting users against **pre-defined intrusions** that could be exacerbated by AI.

💪 **Motivation**: practitioners faced more **inhibitors** than **motivators** for AI privacy work.

🔧 **Ability**: **tools** and **resources** that practitioners utilized in their privacy work were typically non-product and non-AI specific.

✉ haopingl@cs.cmu.edu
🐦 @hankhplee
⌂ https://hankhplee.com/

"I Don't Know If We're Doing Good. I Don't Know If We're Doing Bad"
Investigating How Practitioners Scope, Motivate, and Conduct Privacy Work When Developing AI Products

Hao-Ping (Hank) Lee, Lan Gao, Stephanie Yang, Jodi Forlizzi, Sauvik Das