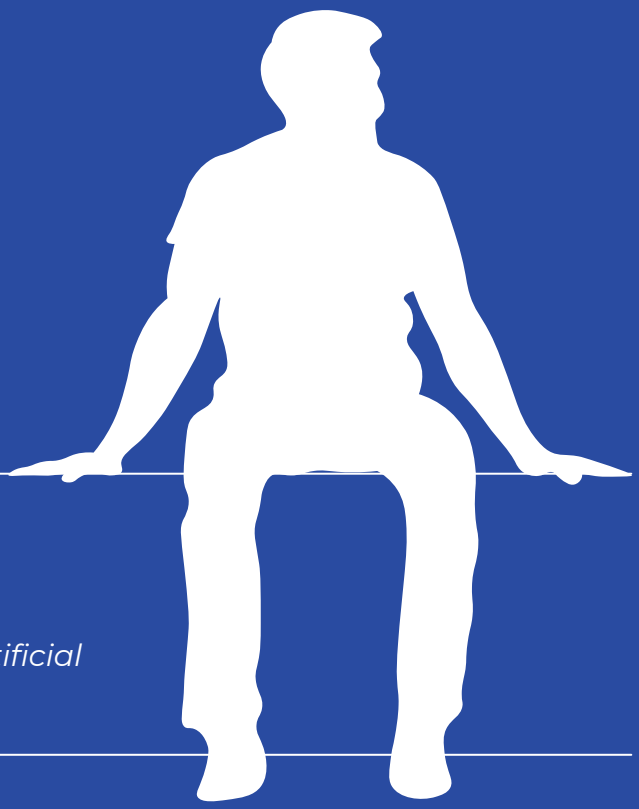
2023

U.S. PUBLIC ASSEMBLY ON

HIGH RISK AI



HOW TO CITE

U.S. Public Assembly on High Risk Artificial Intelligence 2023 Event Report

Center for New Democratic Processes

Syracuse University
Maxwell School of
Citizenship & Public Affairs

Published by the Center for New Democratic Processes
Authors: Sarah Atwood & Kyle Bozentko

CONTENTS

WITH THANKS

This project would not have been possible without the generosity and contributions of numerous organizations and individuals. Our work has been met with curiosity and enthusiastic support by those who have recognized a need for meaningful public engagement on AI.

Most importantly, we thank the Assembly participants who contributed to this work.

During a time when discourse about political division seeps into nearly every facet of public life, this project demonstrates that a diverse body of U.S. residents can not only come together to learn about a complex topic, but also meaningfully deliberate about a positive future that we all might share.

Participants took on the difficult task of learning about AI and placed an additional responsibility upon themselves to address these issues with diligence and care. They did not disappoint.

AI's ongoing impact on society is difficult to grasp, and planning for the future feels increasingly illusive as AI technology and its applications proliferate exponentially. However, we remain steadfast in our belief that public participation is integral to charting a future where AI systems operate for the benefit of all while mitigating risk and harm.

- Sarah Atwood
- Kyle Bozentko
- Kate Mays
- Baobao Zhang

PROJECT **OVERVIEW**

INTRODUCTION

In October 2023, Dr. Baobao Zhang (Maxwell Dean Assistant Professor of the Politics of AI at the Maxwell School of Citizenship and Public Affairs at Syracuse University) partnered with the Center for New Democratic Processes (CNDP) to convene the U.S. Public Assembly on High Risk Artificial Intelligence.

The U.S. Public Assembly on High Risk Artificial Intelligence (AI Assembly) met virtually over eight days. Forty Assembly participants were randomly selected from across the United States to include a broadly representative mix of residents in terms of age, gender, race / ethnicity, educational attainment, employment status, political affiliation, and geographic distribution. Participants were compensated for their time ($1,200 for the entire Assembly) and received technical support in advance of and throughout the event.

Findings from the AI Assembly will be circulated among stakeholders in order to shape future policy discussions on AI governance, policy development, risk and regulatory frameworks, and responsible and trustworthy AI. The AI Assembly will also serve as a starting point for exploring how to more effectively involve the public in future discussions regarding AI and emerging technologies.

The AI Assembly explored public attitudes about risk and uses of artificial intelligence across multiple domains, including administrative, health, search, and face recognition. Assembly participants heard from experts and deliberated about examples of AI systems regarding:

- assessing risk
- exploring accountability and responsibility
- determining harms associated with various uses of AI

CORE TEAM

Dr. Sarah Atwood is Head of Research & Engagement at CNDP. She draws on over 15 years of experience as a facilitator, scholar, and public historian to oversee the design and implementation of CNDP's engagement work and organizational research activities. Sarah received her PhD in American Studies from the University of Minnesota (Twin Cities) and currently serves on the Advisory Board of the Institute for Advanced Studies (UMN).

Kyle Bozentko is Executive Director of CNDP. He has contributed directly to the design, implementation, and evaluation of over fifty deliberative engagement projects (citizens' jury, citizens' assembly, and other mini-publics) in the U.S. and globally. Kyle holds an MS in Gerontology (SCSU) and an MTS (Boston University). Kyle currently serves on the MN Office of Collaboration and Dispute Resolution Advisory Committee and the Bloomberg New Economy Health Council.

Dr. Kate Mays is an Assistant Professor in the Department of Community Development and Applied Economics at the University of Vermont. She completed her PhD in Emerging Media Studies at Boston University's College of Communication. She was a Graduate Fellow at BU's Rafik B. Hariri Institute for Computing and Computational Science and Engineering and a postdoctoral researcher with the Autonomous Systems Policy Institute at Syracuse University.

Dr. Baobao Zhang is Maxwell Dean Assistant Professor of the Politics of AI in the Political Science Department at the Maxwell School of Citizenship and Public Affairs, Syracuse University. She is currently a Schmidt Futures AI2050 Early Career Fellow and a research affiliate with the Centre for the Governance of AI. Dr. Zhang graduated with a PhD in political science (2020) and an MA in statistics (2015) from Yale University.

PARTNERS AND FUNDERS

Center for New Democratic Processes

The Center for New Democratic Processes is a nonpartisan, nonprofit civic engagement and research organization committed to solving complex problems through rigorous, purposefully designed deliberative engagement initiatives.

Maxwell Syracuse University

The Maxwell School of Citizenship and Public Affairs is a community of scholars who believe that responsible citizenship is not just a concept; it's the active, ongoing pursuit to empower a more healthy and inclusive society.

SCHMIDT FUTURES

Rockefeller Brothers Fund
Philanthropy for an Interdependent World

CIFAR

The AI Assembly was funded by Schmidt Futures through an AI2050 Early Career Fellowship awarded to Dr. Zhang. Project outreach activities conducted by CNDP were supported by the Rockefeller Brothers Fund. Survey costs and research team support was provided by the Canadian Institute for Advanced Research (CIFAR).

PROJECT SCHEME

The U.S. Public Assembly on High Risk AI was convened as part of an ongoing research collaborative.

This document is a report in progress - a means to simultaneously share about the AI Assembly - as well as the first publicly available report on the initial outcomes of the Assembly members' work. To date, this work has included three key components:

01

Establishing Project Framework

The project team established the research and engagement framework for exploring public attitudes about AI and risk.

02

Survey

A 3,000-person, nationwide survey was conducted in August 2023. Nearly 2,100 respondents opted-in to join the applicant pool for the AI Assembly.

03

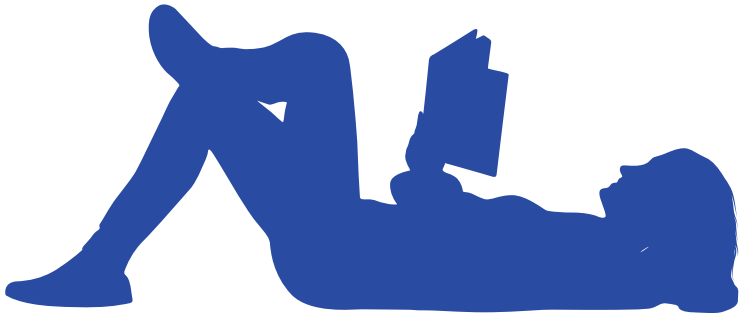
Public Assembly

A 40-person national panel was convened to learn and deliberate about AI and risk. The project was implemented with Institutional Review Board approval.

READING THIS REPORT

This report is a summary of the AI Assembly participants' work. It includes full results from all Assembly questions, which participants completed following expert witness presentations (including question and answer sessions) and subsequent deliberations with their colleagues.

The report also contains a diverse sample of rationale and other statements from participants' deliberations. Participants authorized report authors to edit these statements for clarity and grammar on their behalf. Report authors have taken care to ensure that, to the best of their ability, edits have been minimal.



ASSEMBLY
DEMOGRAPHICS

PARTICIPANT RECRUITMENT

Assembly participants were recruited to represent a broad cross-section of the U.S. public. Nearly 2,100 respondents from a 3,000-person nationwide survey opted-in to join the pool of Assembly applicants. Applicant information was pseudonymised, and 40 participants (plus 3 alternates) were randomly selected to join the Assembly using a sortition algorithm (Panelot).[1]

> Being a part of this assembly was one of the most interesting things that I have ever done. This experience made me realize just how big the world is and how much more there is to learn about what our future holds for us, especially with AI. The most important part to me about this Assembly was being able to meet such a great and diverse group of people.

1. Due to extenuating circumstances, three participants were unable to complete all eight days of the AI Assembly.

PARTICIPANT ONBOARDING

All AI Assembly members received comprehensive orientation and onboarding support to ensure that they were prepared to participate to the best of their ability.

Project onboarding included orientation about informed consent requirements and procedures, explanation of project funding, research goals, and questions, the Assembly Oversight Panel, as well as an introduction to the the nonpartisan deliberative process, itself.

Participants also received technical support prior to and during the Assembly in order to lower any technical barriers and ensure corresponding challenges were mitigated as necessary (including provision of technology such as hardware).

DEMOGRAPHICS

LOCATION

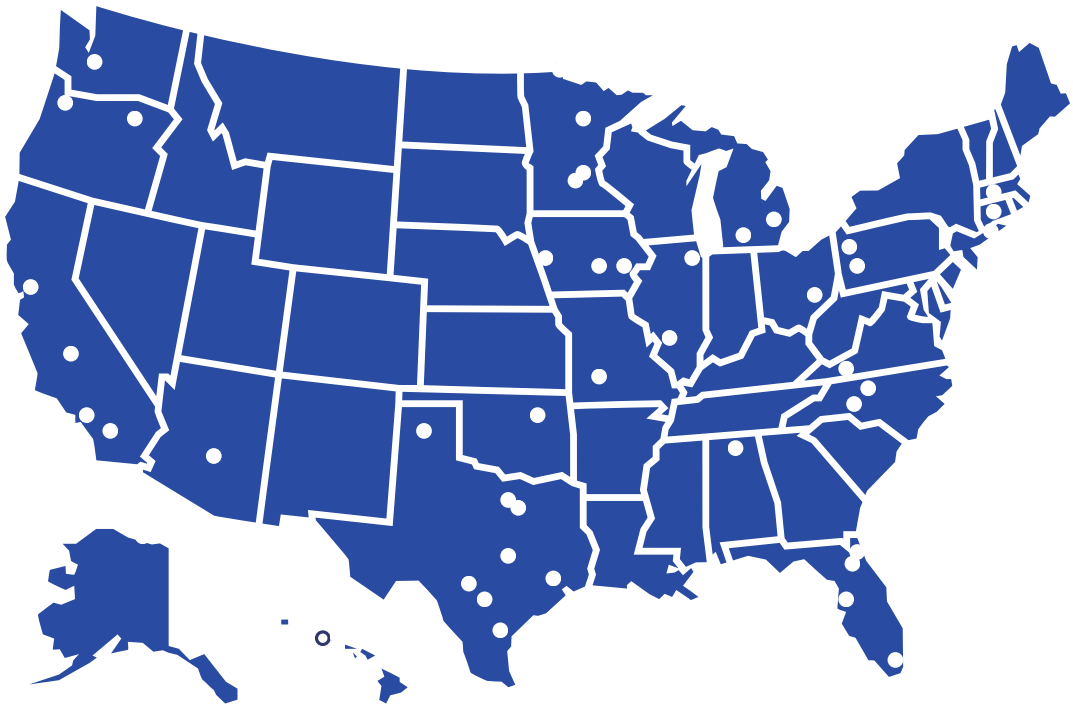
	Target	Actual
Rural	7 - 9	7
Urban	31 - 33	33

- Target categories from U.S. Census, 2020

REGIONS

	Target	Actual
Midwest	7 - 10	9
Northeast	5 - 8	6
South	14 - 17	16
West	8 - 11	9

- Target categories from U.S. Census, 2020



AGE

	Target	Actual
18 - 34	10 - 13	12
35 - 64	18 - 21	20
65+	7 - 10	8

- Target categories from American Community Survey, 2016 - 2021

GENDER

	Target	Actual
Female	19 - 20	19
Male	19 - 20	19
Non-binary / Transgender	1 - 2	2

- Target categories from American Community Survey, 2016 - 2021

EDUCATION

	Target	Actual
High school or less	14 - 17	15
Some college or Associate's	10 - 13	13
Bachelor's or higher	12 - 15	12

- Target categories from American Community Survey, 2016 - 2021

EMPLOYMENT

	Target	Actual
In labor force - full-time	23 - 26	23
In labor force - part-time	3 - 6	3
In labor force - unemployed	1 - 3	1
Not in labor force	13 - 16	13

- Target categories from American Community Survey / US Census

PARTY AFFILIATION

	Target	Actual
Independent, none, don't know	10 - 13	8
Moderate Democrat, lean Dem	7 - 10	9
Moderate Republican, lean Rep	8 - 11	11
Strong Democrat	5 - 8	7
Strong Republican	4 - 7	5

- Target categories Reuters / Ipsos Poll: Issues Survey, 2023

RACE / ETHNICITY

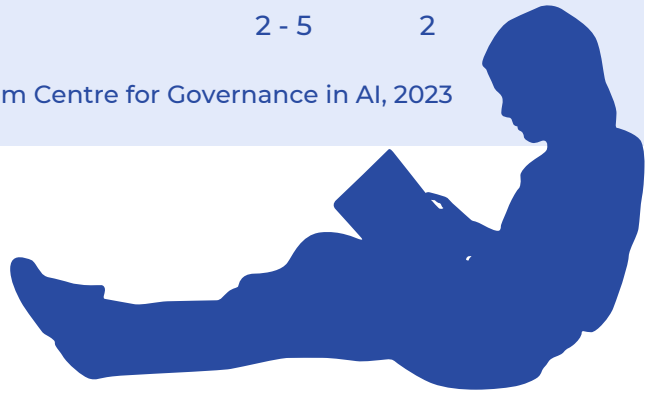
	Target	Actual
Asian	1 - 5	2
Black or African American	3 - 7	3
Hispanic or Latino	7 - 10	9
Indigenous	1 - 3	1
White / European-American	22 - 25	23
Two or More Races / Other	4 - 8	4

- Target categories from American Community Survey, 2016 - 2021

TECH / AI KNOWLEDGE

	Target	Actual
Closely follows AI-related news	4 - 7	8
Formal AI education or work	1 - 4	4
Heard about AI	27 - 30	26
Never heard about AI	2 - 5	2

- Target categories from Centre for Governance in AI, 2023



ASSEMBLY CONTENT

ASSEMBLY CONTENT

Participants engaged with a wide range of expert witnesses. Presentation topics included an introduction to AI, how AI systems are built and trained, how data is used within AI systems, AI governance, ethics, and risk frameworks, as well as the use of search / browser histories, health records, facial images, and administrative records within AI systems.

Expert witnesses were asked to respond to a series of prompts for the AI Assembly, and submitted draft presentations for review by the Oversight Panel.

Oversight Panel members then reviewed draft presentations for bias, jargon, general clarity, etc. Expert witnesses were then asked to amend their slides as appropriate before their presentations were distributed to Assembly participants.

Individuals who contributed to the Oversight Panel and as Expert Witnesses for the AI Assembly were asked to contribute based on their personal perspective and professional experience and were not asked to represent the views of their affiliated employers / organizations.

OVERSIGHT PANEL

The Oversight Panel provided feedback on the AI Assembly specification documents and reviewed expert witness presentation materials for clarity, accuracy, and breadth.

Dr. Catherine Aiken
Director of Data Science and Research at the Center for Security and Emerging Technology at Georgetown University

Courtney Lang
Vice President of Policy, Trust, Data, and Technology, Information Technology Industry Council

Dr. Michael Miller
Managing Director, Moynihan Center, City College of NY – formerly Social Science Research Center

Dr. Tina Nabatchi
Director of Program for Advancement of Research on Conflict and Collaboration, Strasser Endowed Professor of Public Administration, Maxwell School at Syracuse University

Brian Scarpelli
Executive Director, Connected Health Initiative, ACT: The App Association

Dr. Allan Tucker
Head of Intelligent Data Analysis Group, Reader in the Department of Computer Science at Brunel University

ASSEMBLY PRESENTATIONS

01

Introduction to Artificial Intelligence

Dr. Solon Barocas, Principal Researcher, Microsoft Research

02

Anatomy of AI Systems

Dr. Katrina Ligett, Professor of Computer Science, Head, MATAR Program on the Interfaces of Technology, Society, and Networks, Hebrew University

03

All About Data

Dr. Gissella Bejarano, Assistant Professor of Computer Science, Marist University

04

Ethical, Regulatory, and AI Risk Frameworks

Dr. Brandie Nonnecke, Associate Research Professor, Goldman School of Public Policy; Director, CITRIS Policy Lab, UC Berkeley

05

Browser & Search History

Dr. J. Nathan Matias, Assistant Professor, Cornell University, Founder, Citizens and Technology Lab

06

Health Record

Dr. Marzyeh Ghassemi, Assistant Professor, AI Chair & Canada Research Chair, MIT, CIFAR

07

Face Image

Patrick Grother, Scientist, National Institute of Standards and Technology (NIST), U.S. Department of Commerce

08

Administrative Record

Dr. Chris Meserole, former Director, AI and Emerging Tech Initiative, Brookings Institution; current Executive Director, Frontier Model Forum

INTRO PRESENTATIONS

Background witnesses provided an overview of AI, introduced to how AI systems are developed, explained how AI systems utilize data, and discussed current ethical and regulatory frameworks pertaining to AI.

Introduction to Artificial Intelligence

Dr. Solon Barocas delivered an introductory presentation on AI. This included an overview of algorithms, machine learning, different types of AI systems, automated decision making, as well as how humans shape the AI the lifecycle.

Anatomy of AI Systems

Dr. Katrina Ligett delivered an introductory presentation on how different AI systems are built and trained. This included how AI can be used to make predictions, how accuracy / robustness is measured, interoperability, as well as explainability and transparency in AI systems.

All About Data

Dr. Gissella Bejarano delivered an introductory presentation on data and how AI systems employ data. This included what might constitute "data" in different AI systems, the management and movement of data, primary vs. secondary data use, as well as consent regarding data use.

Ethical, Regulatory, and AI Risk Frameworks

Dr. Brandie Nonnecke delivered an introductory presentation on ethical and regulatory frameworks and considerations. This included different ways of thinking about harms, fairness, risks, as well as guidance, policy, and regulatory frameworks related to AI governance.

BROWSER & **SEARCH HISTORY**

BROWSER / SEARCH HISTORY

Assembly participants examined how a browser / search history might be used in various AI systems including narrow AI, general purpose AI, as well as for secondary and future uses.

01

Uses for a Browser / Search History

- Used to deliver content, websites, ads, news, and / or pricing schemes across websites and applications
- Used to generate answers and explanations in response to an individual's queries, such that the queries and outputs are associated with the individual
- Used within additional datasets to train other models that will be employed across various sectors and industries for other purposes

02

Expert Witness

Dr. J. Nathan Matias delivered a background presentation about the use of browser / search histories in various AI systems and answered participant questions about how histories could be utilized throughout the AI lifecycle.

03

Deliberations & Voting

Assembly members deliberated about the potential benefits and harms of a browser / search history's use within different AI systems. Participants then registered their votes indicating which level of risk these uses might pose to various individuals, institutions / organizations, and society as a whole (including particular groups).

POTENTIAL BENEFITS

INDIVIDUALS

- Individuals may receive more relevant content tailored to their needs or wants (consumer goods, locations to visit, music or video recommendations, predictive search results, etc.); individuals may be directed to useful content regarding health, care, diagnosis, or have harmful content minimized
- Individuals may spend less time searching for specific information
- Individuals may benefit from the assistance of chatbots and receive more personalized and therefore relevant information
- Voice assistants can provide information or recommendations to individuals on a wide range of topics
- Individual researchers could benefit by accessing large amounts of information quickly and easily
- Wider adoption of these systems by individuals may make their lives more efficient, allowing more time for other activities
- Individuals may be exposed to new or novel ideas, concepts, or topics they otherwise might not have encountered
- The use of AI systems using browser / search histories may foster connections between individuals who have shared interests

POTENTIAL HARMS
INDIDVIDUALS

- Users may receive recommendations that are harmful to themselves or others (e.g. people experiencing addiction or a mental health crisis could receive results that exacerbate harmful behaviors, people might be encouraged to commit acts of violence, etc.)
- Sensitive information may be captured, sold, or leaked without an individual's knowledge
- Recommending systems using a browser / search history may direct users to bad information, surveil users' online activity, limit access to some views, opinions, and / or products, services and content, or simply waste the user's time by providing unhelpful results
- Personal information could be used in ways that an individual does not approve of if / when their data is distributed, shared, or sold
- Individuals may rely on AI that is still in development that may have a higher likelihood of providing incorrect suggestions, or otherwise biased or inaccurate recommendations (e.g. a chatbot could provide a recommendation to a product based on popularity that is potentially unsafe or ineffective for an individual)
- Multi-user kiosks (e.g. libraries) could have multiple profiles' worth of data to influence results and a user may not be able to delete their browser / search history or consent to its capture
- Some users may not want their information tracked but lack control over who has this data and, therefore, have no idea how to manage its current and future uses or whom to hold legally responsible / accountable for harm caused
- Individuals' browser profiles may be used in various ways that would cause future harm through AI applications (such as the promotion of dangerous products, targeting of children, or nefarious use of datasets) to themselves or others and may not know who to hold legally responsible for any harm caused
- The volume of data from search / browser history and user profiles may start to invade a user's privacy when they receive targeted recommendations

POTENTIAL BENEFITS

INSTITUTIONS

- Political parties could use AI systems to run targeted campaigns and increase voter engagement
- Organizations could capture and analyze large amounts of data for future use (e.g. setting prices for goods and services)
- Companies and data brokers may sell collected data to third parties
- AI systems using browser / search histories might limit the need for as many customer support staff and improve customer service experiences through the use of chat and other functions
- Institutions (e.g. companies, educational, and healthcare organizations) could have access to large datasets that can be used to create a more detailed understanding of the needs of users, consumers, etc.
- Organizations (e.g. businesses, hospitals, nonprofits, schools, social service organizations, veterans groups, etc.) could tailor goods / care, target advertising and marketing more effectively, customize messages for specific audiences, and identify customers, supporters, and volunteers more effectively
- Companies and organizations may be able to benefit by using shared technologies for cost effectiveness and profit (academic institutions, banks, charities, hospitals, fire and police departments, retailers)
- Both private and public institutions could utilize browser / search histories to create user profiles so that the future development of AI systems may lead to improved service for users / recipients
- Institutions would have access to large amounts of data to develop and train future models for accuracy

POTENTIAL HARMS
INSTITUTIONS

- Some businesses may be harmed if they become less likely to be recommended as consumers are pushed in the direction of select businesses
- Companies and organizations relying on recommending systems for employment screening or other similar uses may overlook viable candidates due to inconsistencies in data or misalignment between keywords and resumes, applications, etc.
- There is an increased potential for privacy breaches such as Health Insurance Portability and Accountability Act (HIPAA) violations and others through wide range sharing of personal information
- A business or governmental agency may be sued due to AI decision-making that is extrapolated from biased (such as racist or sexist) data made through using a browser / search history
- Health care systems and pharmaceutical companies may develop systems based on incorrect or biased data, leading to inaccurate recommendations and lower quality care and health outcomes
- Institutions may become too dependent on AI systems that use browser / search histories, resulting in significant service disruptions or practical challenges (e.g. if a system fails or is found to be inaccurate or unreliable)
- Storing increasing amounts of data about consumers / users from AI systems using browser / search histories may open organizations up to more frequent data breaches or hacking attempts
- Organizations may obtain or hold incorrect or unethical data from previous datasets (built from browser / search histories) and deploy AI systems built using this data due to limited transparency / traceability
- Smaller businesses may lack the ability to compete with the increase of AI systems to target consumers; they may lack the ability to update or improve technology or have a legacy system that is incompatible with newer systems

POTENTIAL BENEFITS

SOCIETY

- Society may benefit by recommending systems weighting / de-emphasizing toxic or inaccurate content in searches
- Society could benefit by consumers having more customized, easier customer experiences and easier access to products
- Increased profit resulting from businesses' use of AI systems utilizing browser / search histories would benefit the economy
- New technologies that could benefit society may be developed due to enhanced capabilities for capturing, processing, and analyzing vast amounts of data through browser / search histories
- Ease of access to information can result in faster advancements (particularly in science)
- Recommendation systems may benefit marginalized groups if they expose the broader public to a more diverse range of information and results
- AI systems using browser / search histories might improve lives in currently unforeseeable ways such as improving disaster preparedness and response, connecting people in new ways, and making more tasks easier for more people
- Sharing datasets across companies to build AI systems could lead to more diverse datasets and safer AI systems
- A wider range of data in a centralized space can expedite research and its subsequent dissemination

POTENTIAL HARMS

SOCIETY

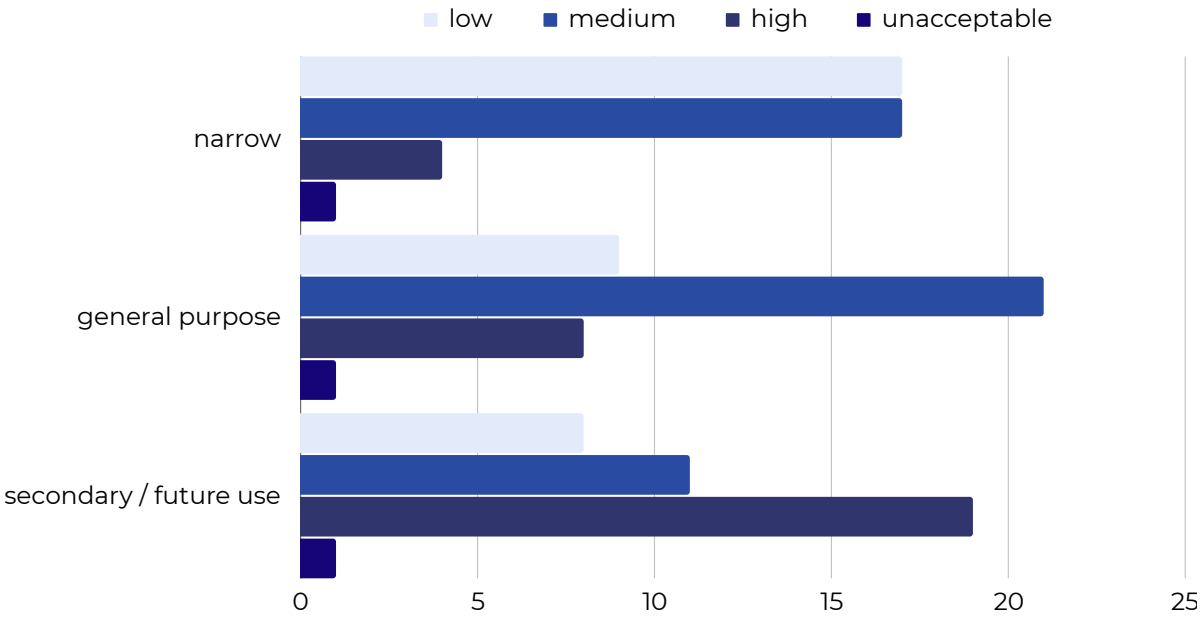
- These systems could be used to falsify data, create misinformation, or fuel disinformation
- Children may be targeted with inappropriate content on a range of platforms
- The development of AI-created content such as text or video may lead people to mistrust one another since they would be unsure whether information was genuine and / or trustworthy
- Incorrect information about political campaigns, elections, and political figures may have a detrimental impact on democracy
- False or misleading information can lead to discrimination based on race, ethnicity, sexual orientation, etc.
- Many people taking action based on incorrect, false, or inaccurate recommendations could be detrimental to society (e.g. public health impacts, etc.)
- Minority groups could be misrepresented or receive biased outcomes due to underrepresentation in data
- There may be a societal loss of trust in systems such as healthcare and policing due to errors, data breaches, or misinformation
- Certain job sectors may be replaced by AI built on browser / search histories

Q1 - LEVEL OF RISK

After deliberating about the potential benefits and harms of various AI systems using a browser / search history, as well as the potential risk of these systems to various parties, participants registered their votes.

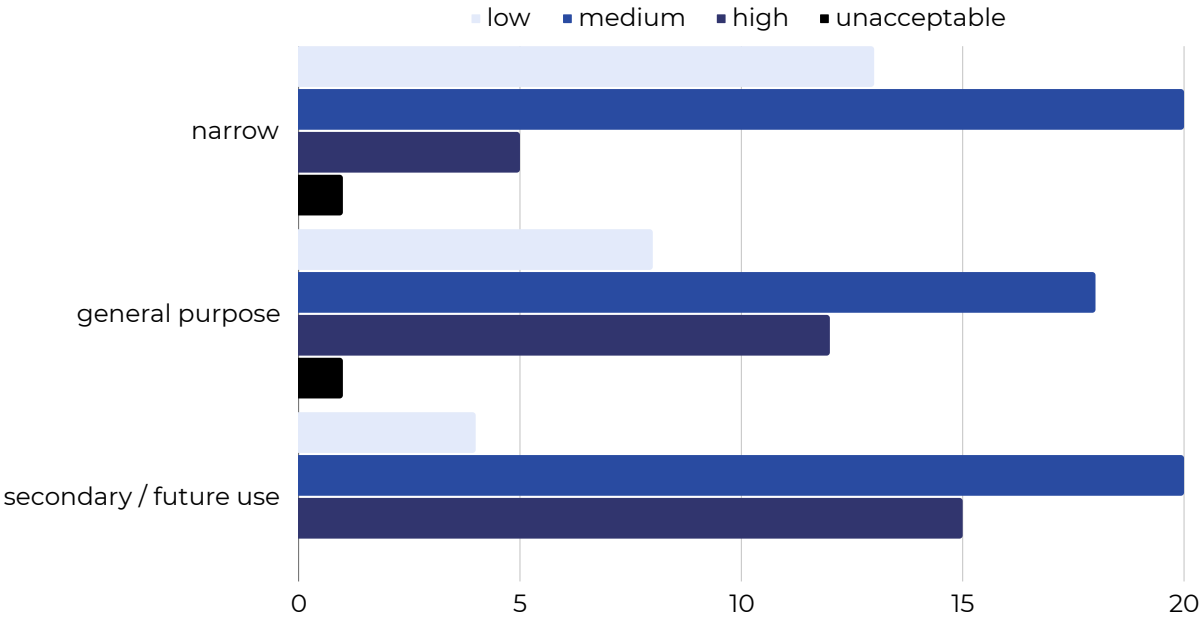
Given the possible benefits and possible harms associated with the various AI uses in “Browser / Search History” please indicate which level of risk you would apply to each narrow, general purpose, and secondary / future use as it pertains to individuals, institutions, organizations, and society as a whole (including particular groups).

Individuals

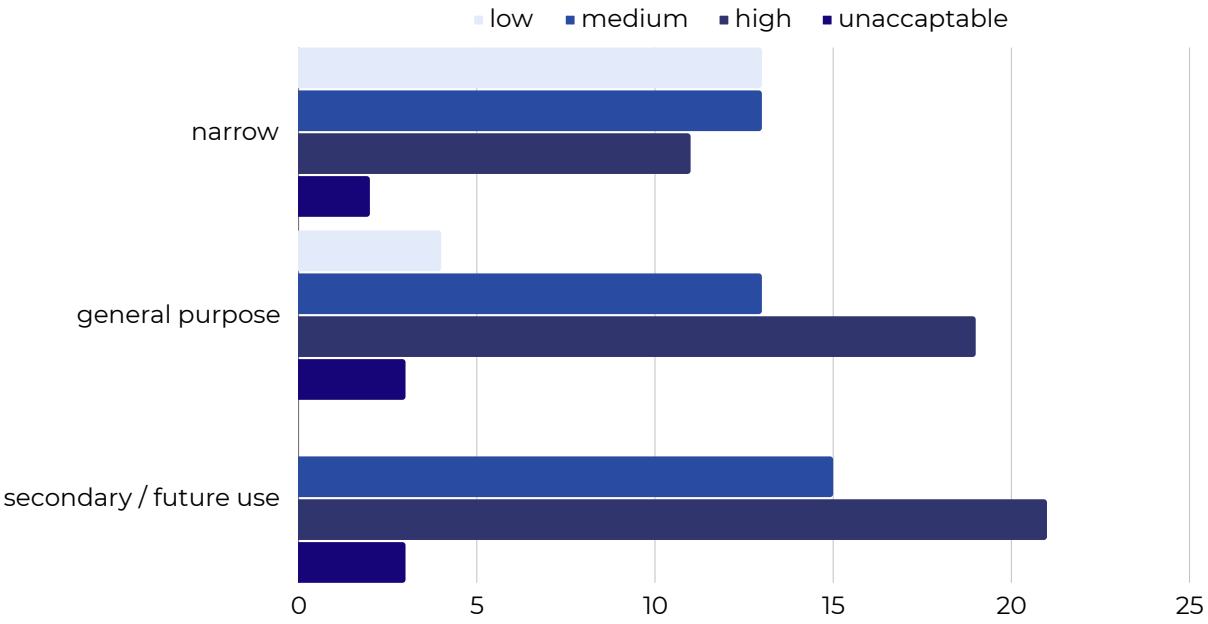


Q1 - LEVEL OF RISK

Institutions & Organizations



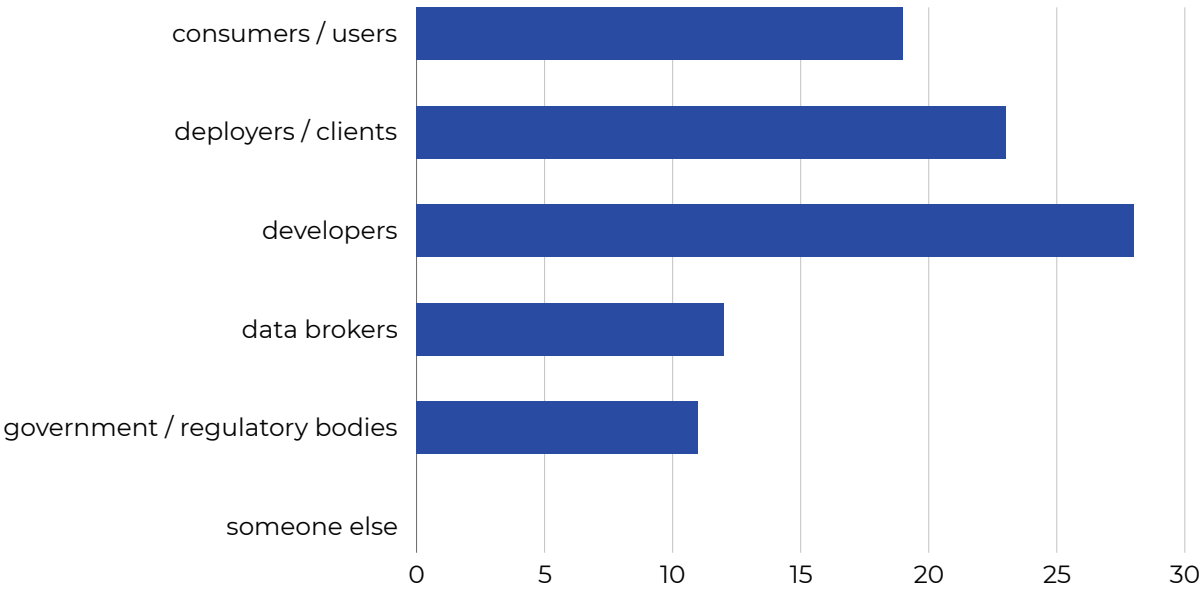
Society as a Whole



Q2.A - ACCOUNTABILITY

Participants deliberated about which parties or actors should be held accountable when an individual or group is harmed or an incorrect decision is made by an AI system using a browser / search history. They then registered their individual votes and recorded their rationale.

Which parties or actors in the AI lifecycle should be held accountable when an individual or group is harmed or an incorrect decision is made by an AI system using a browser / search history?



Q2.A. - RATIONALE

The user or operator of an AI system should be held accountable for how they apply AI technology. If they use it inappropriately or fail to follow guidelines they may be responsible for an outcome that results in harm. Companies and organizations that deploy AI systems have a responsibility to ensure these systems are used in a responsible manner. They should set guidelines / policies and monitor how an AI system performs.

Clients should be held responsible if they misuse an AI system or if they knowingly continue using a flawed system. Developers should be primarily responsible for any flaws in their systems that cause harm or produce incorrect decisions. Data brokers should be responsible if they knowingly provide or obtain biased or inaccurate data. Users of all types are most likely to be the victims of harm and should not be held responsible.

All parties should be held accountable when an incorrect decision is made or groups are harmed from AI systems. Partial blame may be on users for not doing their research and trusting one source of information. Developers and deployers may be held responsible for releasing systems for use prior to those systems being tested.

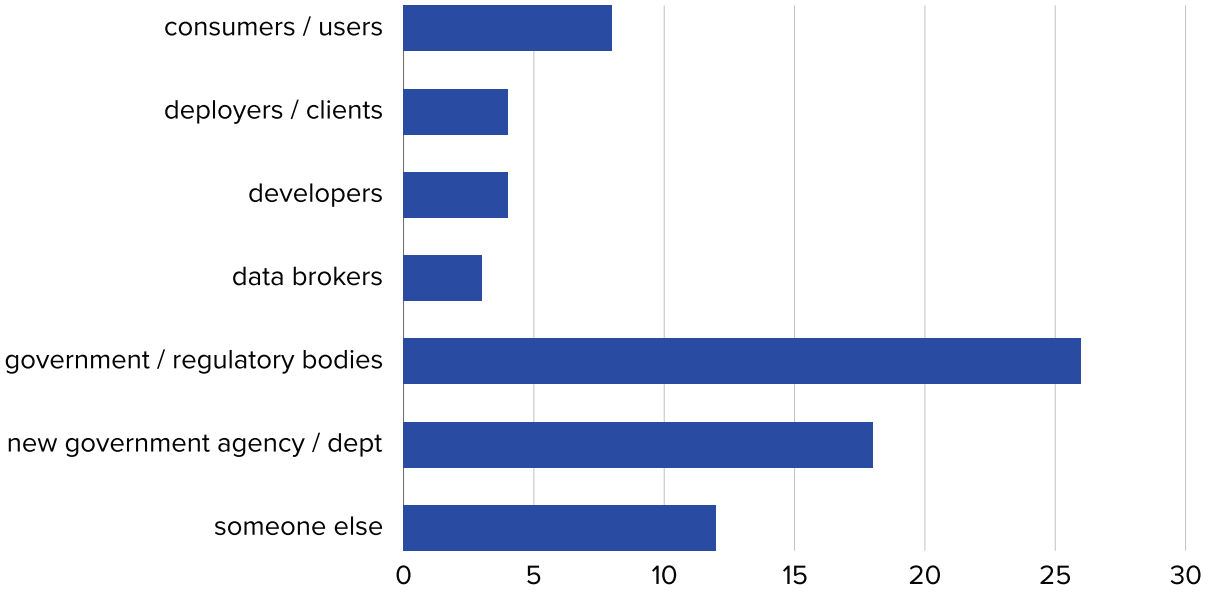
Users or consumers are directly responsible for the inputs within a search function. Developers are in charge of ensuring the search function works the way it is intended.

The inventor of a product has historically been held legally liable when their invention malfunctions and causes harm.

Q2.B - RESPONSIBILITY

Participants deliberated about who should determine which parties or actors should be held responsible when harm is done to an individual or group through an AI system using a browser / search history. They then registered their individual votes and recorded their rationale.

Who should determine which parties or actors should be held responsible when harm is done to an individual or group through an AI system using a browser / search history?



Q2.B. - RATIONALE

The best way to strike a balance between fairness and transparency is to make the committee focused around the public more than the government. AI is very easy to use for nefarious purposes, which is why so many citizens are nervous about it. Taking opinions from the people makes the process more trustworthy. Whoever was wronged should be involved (even if minimally / indirectly).

I don't believe anyone associated with the development or use of AI should be determining who should be held responsible as there is a conflict of interest for each of them. I believe that current government bodies should become more equipped to handle these situations while a new regulatory or government agency is put in place that specializes in AI. If multiple parties are at fault, the court systems are always an option where each party may make their case.

If it were developers / brokers / deployers they would be, in effect, self-policing. There needs to be a different power to report to. A new agency is needed, but in the meantime, existing agencies need to develop and enforce rules and regulations to determine when harm has been done.

The government should create, maintain, and update the rules for proper creation and responsible use of AI systems.

Users should be able to report when harmed because of an AI system. Regulatory bodies and new agencies (specifically built to manage AI systems) need to work together to maintain and decrease possible harms. If enough people report being harmed, it should be enough evidence to open up a legal case against the company / organization using the AI system.

HEALTH RECORD

HEALTH RECORD

Assembly participants examined how a health record might be used in various AI systems including narrow AI, general purpose AI, as well as for secondary and future uses.

01

Uses for a Health Record

- Used to determine / guide patient care and treatment
- Used to generate responses about an individual's health and care
- Used within additional datasets to train other models that will be employed across various sectors and industries for other purposes

02

Expert Witness

Dr. Marzyeh Ghassemi delivered a background presentation about the use of health records in various AI systems and answered participant questions about how these records could be utilized throughout the AI lifecycle.

03

Deliberations & Voting

Assembly members deliberated about the potential benefits and harms of a health record's use within different AI systems. Participants then registered their votes indicating which level of risk these uses might pose to various individuals, institutions / organizations, and society as a whole (including particular groups).

POTENTIAL BENEFITS

INDIVIDUALS

- Patients will be able to receive more personalized, relevant care and improved health outcomes
- Patients may receive more hands-on time with doctors if administrative tasks are simplified by using a healthcare record
- Healthcare workers will be able to more easily access records, develop care plans, and manage treatments
- AI systems using a healthcare record may enable patients to receive a diagnosis / treatment plan more quickly while reducing costs
- Individuals looking for care support may find helpful information on medical conditions by using general purpose AI (outside of a healthcare setting)
- Patients may receive more accurate diagnoses if AI systems help to reduce technical mistakes made by care providers
- Predictive systems using a healthcare record could help patients seek out and receive prophylactic treatment
- Larger amounts of data would assist in the diagnosis and treatment of future patients
- Patients may benefit from AI systems using a healthcare record by ensuring that treatment plans are increasingly safe (e.g. by further minimizing drug interactions)

POTENTIAL HARMS

INDIDVIDUALS

- Patients may be misdiagnosed, given incorrect information, or denied care care for their needs based on data from others (e.g. they may have abnormalities that an AI system cannot account for, or datasets may be biased / include flawed data leading to inaccuracies)
- Patients may not receive the time they feel is necessary with their physician or may feel interactions with health providers lack empathy if care is augmented by AI systems
- Individual care providers may be open to lawsuits based on malpractice using AI systems
- Patients seeking care may be misdiagnosed and / or delayed from reaching an actual professional when interacting with a chatbot helping to direct their care path
- Sensitive patient information could be leaked and have future ramifications for patient privacy
- An individual may have coverage denied or their insurance rate increased based an exchange with an AI system
- Individuals, especially those who are uninformed or unable to advocate for themselves, may be convinced to agree with treatment plans because they are legitimized through AI
- A patient may find that technologies actually increase the price that they pay for some services / care or may limit their access to treatments
- Future AI systems using health records could be unethical if the data has unclear origin (“provenance”); incomplete samples or biases in data could also result in the misuse of models that produce improper diagnoses, treatment plans, etc.

POTENTIAL BENEFITS

INSTITUTIONS

- Health care providers are able to share information about patients, saving both resources and time
- Medical systems and insurance companies will be able to capture and analyze patient data more efficiently, including health history, medication lists, and approvals / billing
- Efficiencies gained by using AI systems could reduce the cost of care for health providers, insurers, etc.
- Care providers and health systems may benefit from AI systems using health records to reduce misdiagnosis and mistreatment of patients
- Health systems may be benefit by reducing the the time needed to diagnose or treat patients while providing additional resources to patients for managing their illness / issue
- Healthcare systems will be able to generate data for both immediate and long-term purposes (e.g. care plans, research, and resource allocation)
- More data provides the opportunity for innovation and refinement of current health products / services among public and private institutions
- Pharmaceutical companies and researchers can use AI systems to study and develop new treatments, accelerating drug discovery and development
- Organizations can be better prepared to respond to public health needs (distribute resources fairly, project trends, and receive advance warnings)

POTENTIAL HARMS
INSTITUTIONS

- Health systems may become too dependent on AI (e.g. care systems may be at risk for unintended or malicious leaks of data, there may be less personalized standard of care for patients, the use of AI could cause a company or institution to become unreliable and / or subject to lawsuits)
- Healthcare systems may displace workers due to automation using AI systems
- Healthcare systems using AI in treatment schemes may find that staff are unable to explain how decisions about billing, care, diagnosis, treatment, etc. are made using patient records
- Insurance companies may use AI systems to deny coverage or procedures due to errors in data entry (via healthcare records)
- Health systems could inadvertently establish care precedents and protocols based on biased data
- Hospitals may use AI systems to cut costs but not improve care
- A healthcare system's model / algorithm audits may fail to identify / prevent skewed outputs and those systems could still cause medical harm and patient trauma
- Small pharmacies in low income communities may not be able to compete with larger providers that use AI to manage care and reduce costs
- Organizations that employ health records to develop AI systems may lose public trust if they continue to utilize biased data, employ flawed models, or produce inaccurate results

POTENTIAL BENEFITS

SOCIETY

- There is a potential to reduce bias in care
- People with fewer resources may gain increased access to care
- There could be an overall improvement in both the efficiency and quality of care
- The rising costs of healthcare may be slowed
- Aggregated information could provide a centralized database that benefits a wide range of people; this could be used for early detection of outbreaks such as Covid, expand opportunities for Federal and State funding, or development of treatments
- Large-scale prediction for preventative care could result in better health outcomes for target populations
- AI systems could help improve public health outcomes (better detection of trends, ability to understand environmental factors as they relate to community health, forecasting treatment needs such as vaccines, etc.)
- AI systems using health records might create new opportunities for expanding care models such as telehealth for low-to-no cost or expanding preventative care to underserved communities
- Ideally, future / secondary use of health records would result in less discriminatory systems while enabling more equal access to care

POTENTIAL HARMS

SOCIETY

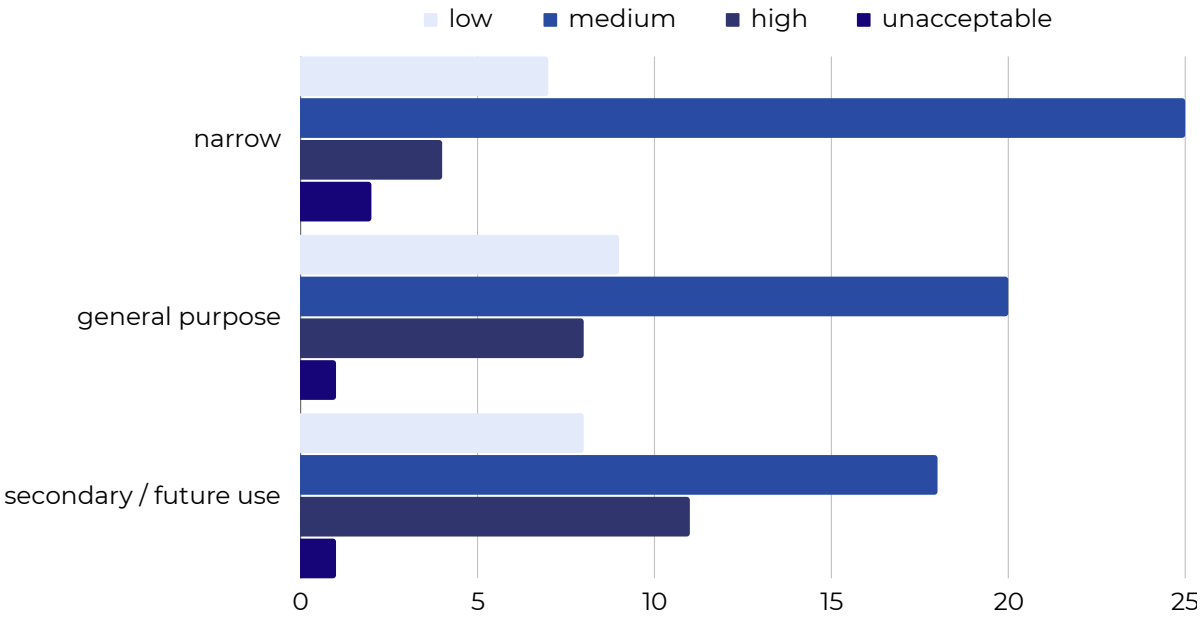
- Less compassionate care may become increasingly common throughout the health system as AI is more frequently used
- If an AI system is populated with unethical / biased data or employed without oversight, individuals in groups underrepresented in the data pool may lose trust in physicians, hospitals etc., leading to decreased trust in the healthcare system
- Underrepresented communities may be at risk of mistreatment or misdiagnoses based on flawed data (people of color, low income patients, women, etc.); this could continue to perpetuate disproportionate care and unequal treatment
- The public may become increasingly reliant upon AI systems used outside of health systems and become less trusting of care providers
- General purpose AI systems that draw on health records may not factor all aspects of individuals' humanity into decisions and recommendations
- Large databases of private health information could become available to hackers
- Inequity between medical systems with access to modern technologies vs. medical systems that do not have the capacity to use systems (e.g. NYC hospital vs. rural hospital in the midwest) could perpetuate disproportionate care among different populations
- Predictive healthcare insights from AI systems using health records could be seen as a breach of patient privacy
- Society may be harmed if AI systems that utilize health records are built upon inaccurate or biased datasets or preexisting systems, therefore perpetuating a cycle of flawed AI

Q1 - LEVEL OF RISK

After deliberating about the potential benefits and harms of various AI systems using a health record, as well as the potential risk of these systems to various parties, participants registered their votes.

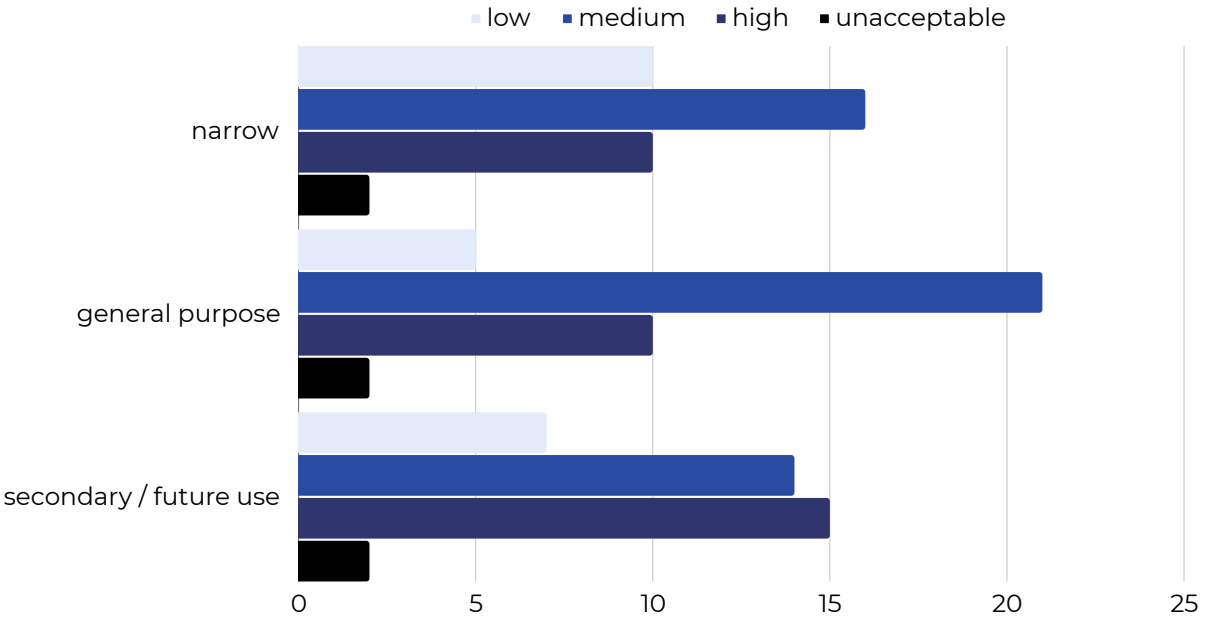
Given the possible benefits and possible harms associated with the various AI uses in “Health Record” please indicate which level of risk you would apply to each narrow, general purpose, and secondary / future use as it pertains to individuals, institutions, organizations, and society as a whole (including particular groups).

Individuals

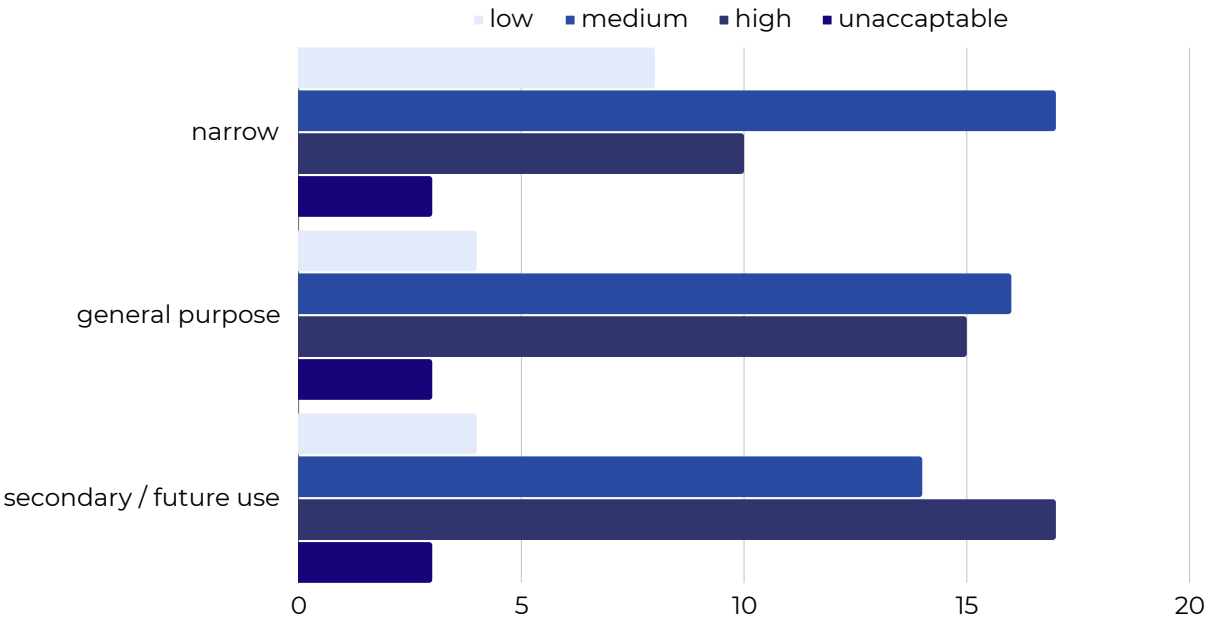


Q1 - LEVEL OF RISK

Institutions & Organizations



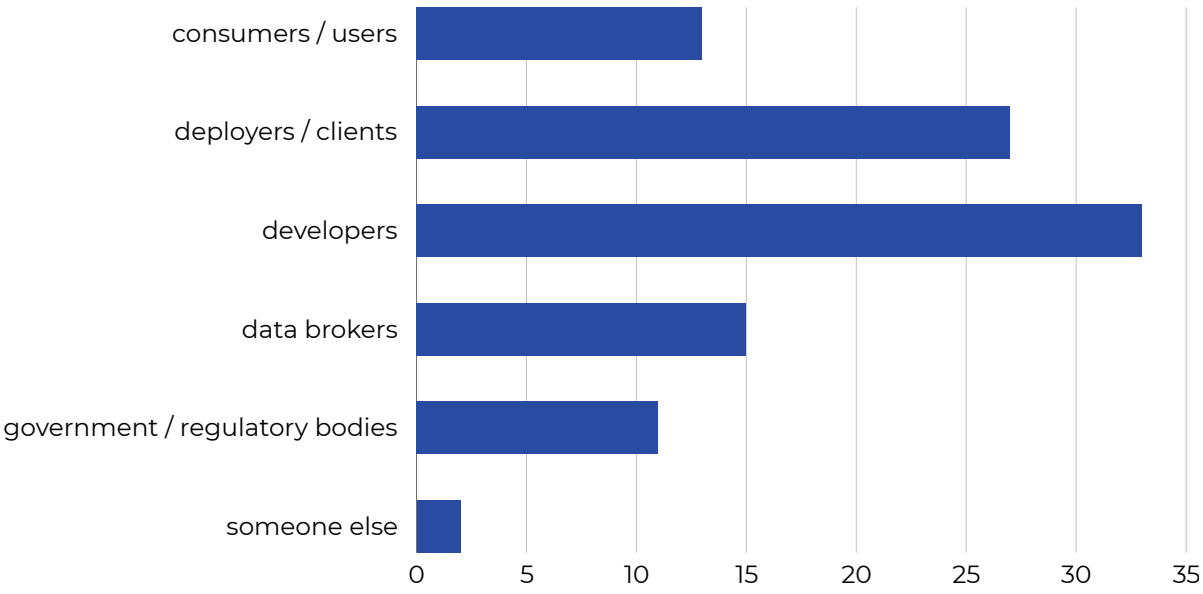
Society as a Whole



Q2.A - ACCOUNTABILITY

Participants deliberated about which parties or actors should be held accountable when an individual or group is harmed or an incorrect decision is made by an AI system using a health record. They then registered their individual votes and recorded their rationale.

Which parties or actors in the AI lifecycle should be held accountable when an individual or group is harmed or an incorrect decision is made by an AI system using a health record?



Q2.A. - RATIONALE

Developers who are responsible for making sure their systems are safe and effective should be held accountable. An organization (like a hospital) using AI should be responsible for making sure its staff knows how to use it properly and that the staff, which are the users, don't misuse the system.

Human oversight must be in place to avoid harm when allowing non-human decision making in an area that is life and death in nature.

Clients should be held responsible if they misuse a system or continue to use a flawed system. If an AI system is misreading X-rays, the doctor or medical system should be held responsible for continuing to use it. The same goes for insurance companies making wrong coverage decisions. Developers should be held responsible for bias or errors in their systems. Government should only be held responsible to the extent that they are clients.

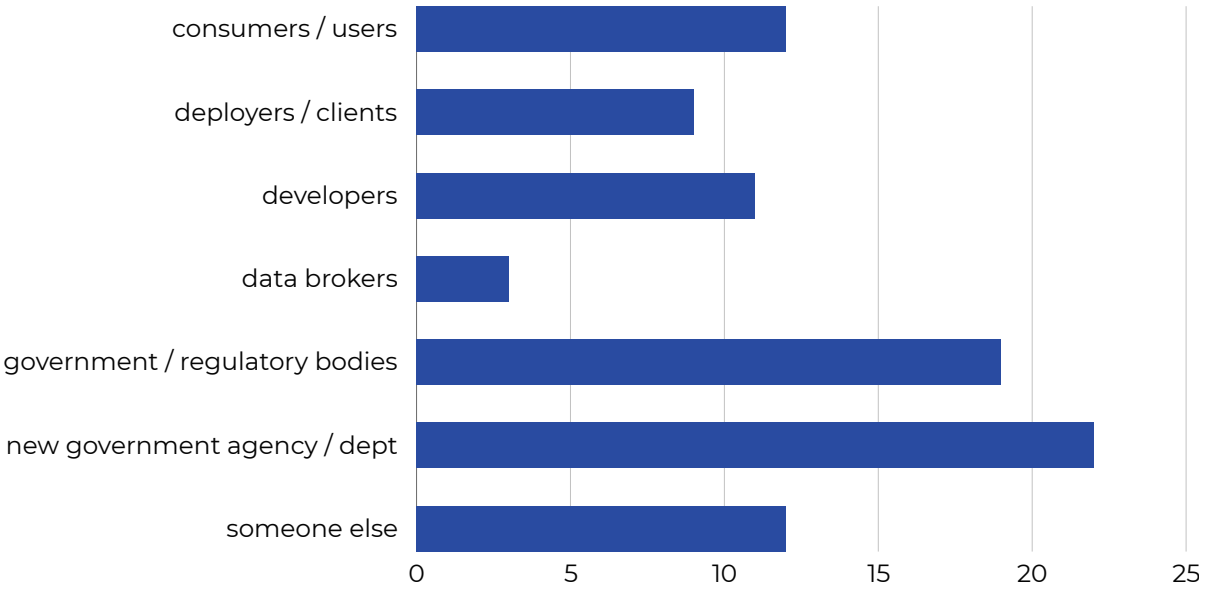
If a group is harmed or an incorrect recommendation is made by an AI system related to a patient's health, the consumer / user should be accountable. Unless AI systems are proven to make the correct decisions, there should always be a doctor or other trained personnel reviewing and approving the AI's recommendation.

If anyone in the health domain is most responsible it would be doctors who should have to manually verify or screen decisions about patients instead of blindly following AI results. Developers and deployers should also be responsible for building faulty systems or not performing complete testing. Governments and regulatory bodies should not allow systems which are not accurate or may cause harm to be deployed.

Q2.B - RESPONSIBILITY

Participants deliberated about who should determine which parties or actors should be held responsible when harm is done to an individual or group through an AI system using a health record. They then registered their individual votes and recorded their rationale.

Who should determine which parties or actors should be held responsible when harm is done to an individual or group through an AI system using a health record?



Q2.B - RATIONALE

Health care providers and insurers can use internal control systems to identify harmful or inaccurate systems and stop using them. Developers can identify the reasons for mistakes or harm and fix internal problems or cut off clients who are incorrectly using the system. Regulatory bodies can decide why harm was caused and hold the responsible party accountable.

There has to be oversight, especially with something as serious as a HIPAA violation or a data breach. A regulatory body should be in place to hold the people who developed a faulty AI system accountable. Law enforcement (including the judicial system) should be able to hold lawbreaking accountable.

Who should determine whom is responsible for harm varies case-by-case. If a nurse or doctor is misusing AI, it would be up to the hospital while hospital misuse would be up to medical regulatory boards. Deciding who's responsible for all of the possible harms is a huge task.

AI consumers / users, deployers / clients, and developers all have a conflict of interest when it comes to the use of AI systems. I do not find it wise, appropriate, or ethical to allow them to regulate themselves. Court systems should be brought in, if and when, they are needed.

Existing government / regulatory bodies, including courts, should hold parties accountable for harm done. A new agency would hopefully prevent some future wrong actions / misuse of data / harm to patients. Questions remain about how a new agency might be structured (e.g. Would it replace existing regulatory and legal systems? Would it just give guidelines and regulations? Would it have the ability to fine or penalize offending parties?).

IMAGE OF A FACE

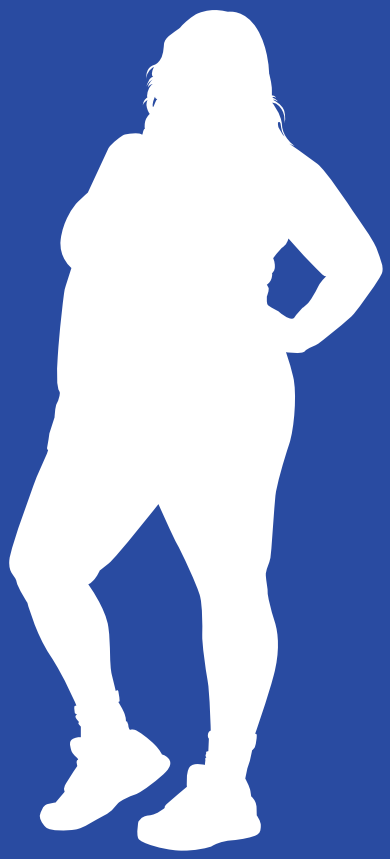


IMAGE OF A FACE

Assembly participants examined how an image of a face might be used in various AI systems including narrow AI, general purpose AI, as well as for secondary and future uses.

01

Uses for an Image of a Face

- Used to determine access to a physical space / device / digital app / account
- Used to generate an image or video using an individual's image / likeness
- Used within additional datasets to train other models that will be employed across various sectors and industries for other purposes

02

Expert Witness

Patrick Grother delivered a background presentation about the use of an image of a face in various AI systems and answered participant questions about how these images could be utilized throughout the AI lifecycle.

03

Deliberations & Voting

Assembly members deliberated about the potential benefits and harms of an image of a face's use within different AI systems. Participants then registered their votes indicating which level of risk these uses might pose to various individuals, institutions / organizations, and society as a whole (including particular groups).

POTENTIAL BENEFITS

INDIVIDUALS

- Individuals may find AI systems that require an image of a face convenient when used to access a device, application, or physical location
- AI systems using an image of a person's face could streamline travel (e.g. alleviating long lines or assisting individuals with language barriers via face recognition screening / clearance)
- Users may benefit from an additional layer of security
- Individuals can use AI systems to create images or videos for entertainment or other uses
- Artists could use AI systems that use images of faces to generate new work or enhance existing art (image or video)
- A missing person might benefit from images generated to model their appearance and assist in identifying them years later
- Individuals can use AI to create avatars to protect their privacy in a digital environment
- Face recognition may alleviate the need for individuals to carry currency (cash, credit card) or identification
- Individuals may no longer need to rely on multiple passwords or other security measures if biometric identification is used instead

POTENTIAL HARMS
INDIDVIDUALS

- Individuals may have their right to privacy violated if a face recognition system is accessed inappropriately and / or leads to identity theft
- An individual may be harmed if they are misdiagnosed / treated using a faulty AI system that employs images of faces
- Individuals may be misidentified and have to deal with the repercussions / negative outcomes (e.g. emotional, financial, reputational harm)
- Individuals may have their likeness used in deepfakes (e.g. AI-generated revenge porn) or used for malicious purposes such as catfishing or extortion
- A person's likeness could be appropriated within different AI models (including for training purposes) without their knowledge
- An artist or actor can be harmed if their work, likeness, image, or video is used without their consent, to produce content without compensating them, or used in violation of copyright
- An individual may ultimately be prevented from anonymity and unable to avoid surveillance
- Biometric data (i.e. an image of a person's face) may be used by anyone for any purpose, removing a person's ability to consent to its secondary / future use and potentially violating their privacy and / or security
- Increased use of face recognition technology may result in more frequent misidentification of people (e.g. incorrectly charging someone as a suspect), the use of digital twins, or impersonators

POTENTIAL BENEFITS

INSTITUTIONS

- Organizations could use face recognition to streamline user access to account information, applications / programs, devices, and locations
- Organizations may be able to cut costs through AI systems using face recognition (e.g. schools and other venues may no longer need staff / security screening at the entrance to facilities)
- Governments, law enforcement agencies, and Homeland Security could use face recognition for preventing crime, locating missing persons, etc.
- Synthetically created images could be used by organizations (both public and private) to generate content for a wide range of purposes
- Law enforcement agencies may use AI-generated images to model persons of interest
- Using general purpose AI systems to produce images or videos may allow companies to reduce costs by replacing staff
- Organizations training AI systems using images of faces may continue to improve the quality of their models leading to higher accuracy (e.g. tighter security controls, more realistic synthetic content, etc.)
- Organizations and institutions may benefit by identifying new methods for distinguishing between real images and AI-generated content
- Organizations could continue to improve their face recognition technology so that misidentification of individuals and groups is less frequent

POTENTIAL HARMS

INSTITUTIONS

- Potential advances in deepfake technology may weaken the security of systems using face recognition and lead to fraud / data breaches for organizations using this technology
- Because biometric data cannot be easily changed and is identifiable, organizations retaining this data may experience increased distrust and scrutiny if there is a data breach or data leak
- Organizations using AI systems that underrepresent people in datasets could be more likely to misidentify individuals within these groups, resulting in lack of trust by those affected populations
- Organizations using biased face recognition technology may damage their reputation or be subject to lawsuits if this becomes publicly known
- Companies could experience harm due to actors using deepfakes to spread unauthorized information or misrepresenting messages about them or supposedly from them
- AI systems used by law enforcement agencies may misidentify individuals, hindering their ability to do their job
- Law enforcement agencies may have their reputation harmed and lose public confidence if AI systems falsely flag people on a regular basis
- Tech companies and AI developers may lose public trust if they are not able to explain how bias is managed, how models are developed, or systems are trained
- Organizations and companies using face recognition for security may be at risk of future breaches if hackers are able to continually override security measures

POTENTIAL BENEFITS

SOCIETY

- AI systems requiring a face to unlock a device could be used to keep children off of age-restricted websites
- AI systems using images of faces could continue to improve efficiency of movement in public
- AI systems requiring face recognition or using images of faces may reduce fraud on a large scale
- AI systems using biometric data (including images of faces) could benefit society if used to help find missing persons or suspects at large
- Society could benefit if AI systems using images of faces could identify (and stop) people at borders who may pose a threat or who should not be allowed to enter
- AI systems using images of faces may be used to enhance augmented reality experiences
- Society may benefit from the use of generative AI technologies that are built upon technologies that employed face recognition to recreate or restore lost or damaged artistic, cultural, or historical artifacts
- Future uses of face recognition technologies could lead to breakthroughs in scientific research (such as modeling environments, extinct species, etc.)
- Future uses of AI technologies that use images of faces may enhance national security and reduce terrorist activity

POTENTIAL HARMS

SOCIETY

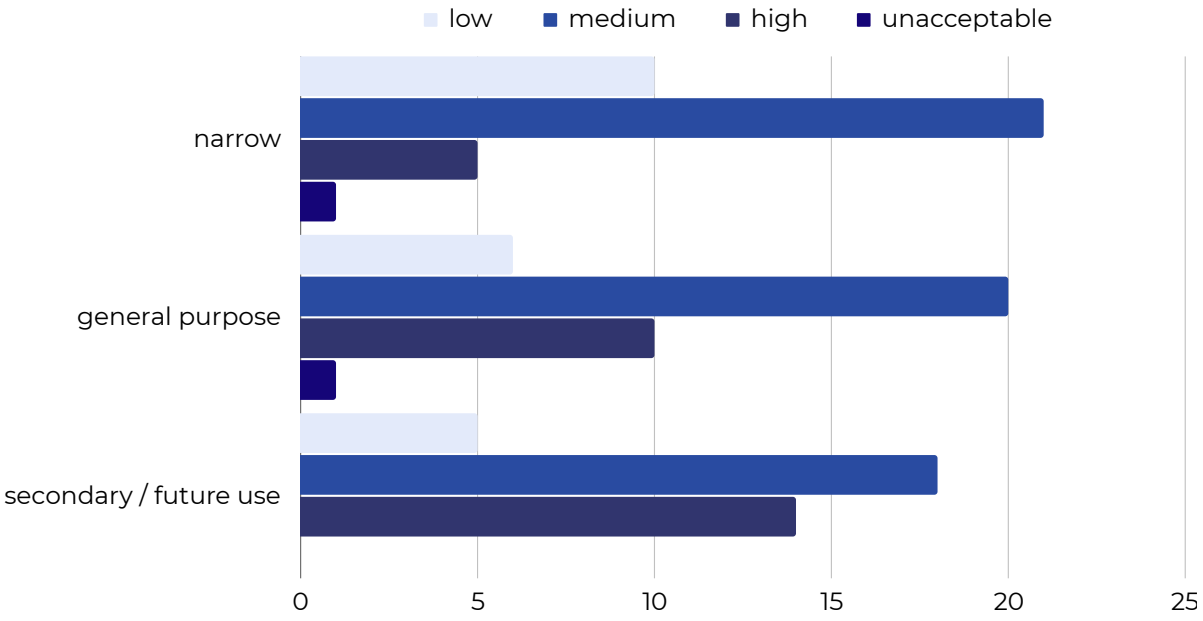
- Face recognition can produce a false sense of security
- Communities prone to misidentification (e.g. people with brown skin or people who are transgender) by AI systems may suffer long-term effects
- Failure or misuse of these AI systems could lead to loss of trust in AI, which may pause or limit AI's progression
- Marginalized groups of people could be targeted through distribution of AI-generated content
- Proliferation of deepfakes and AI-generated content could make it impossible for the general public to distinguish between genuine and AI-generated content
- The creation of artificial entities could be used to spread misinformation or inflammatory speech with little-to-no accountability / liability
- Face recognition by despotic governments may harm populations and individuals (e.g. people who have been seen at / involved with certain events such as political demonstrations may be blacklisted)
- Large-scale AI deepfakes could manipulate public opinion, sway political races, or cause mass hysteria
- The ability to generate synthetic imagery (e.g. static images or video) could reduce employment opportunities across many industries

Q1 - LEVEL OF RISK

After deliberating about the potential benefits and harms of various AI systems using an image of a face, as well as the potential risk of these systems to various parties, participants registered their votes.

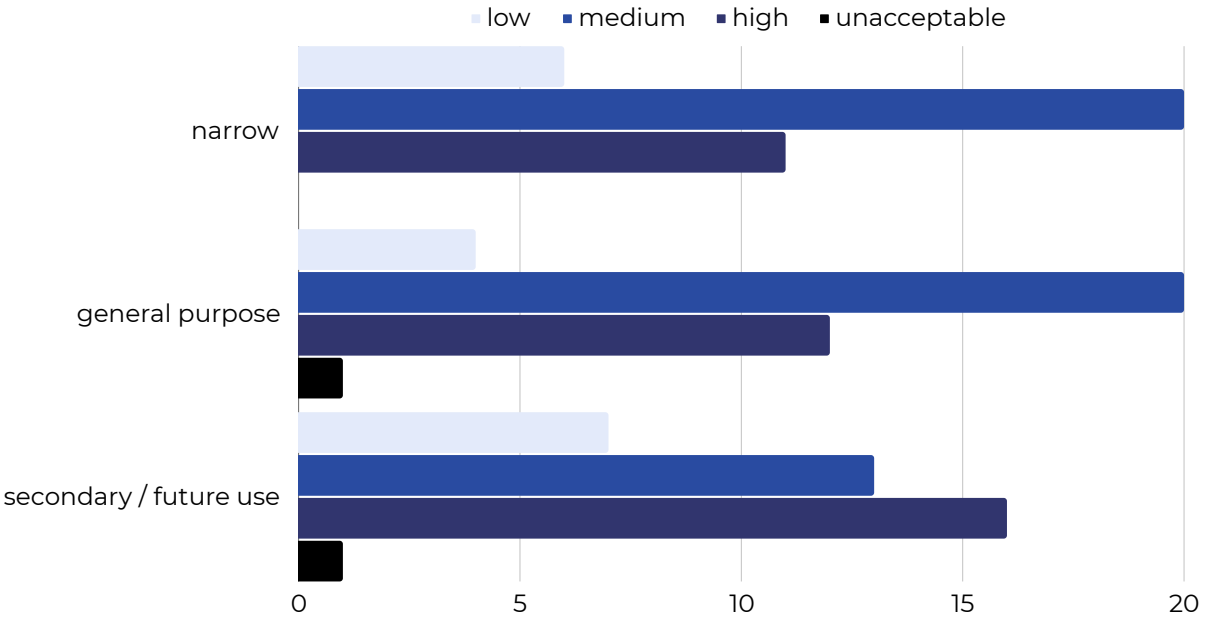
Given the possible benefits and possible harms associated with the various AI uses in “An Image of a Face” please indicate which level of risk you would apply to each narrow, general purpose, and secondary / future use as it pertains to individuals, institutions, organizations, and society as a whole (including particular groups).

Individuals

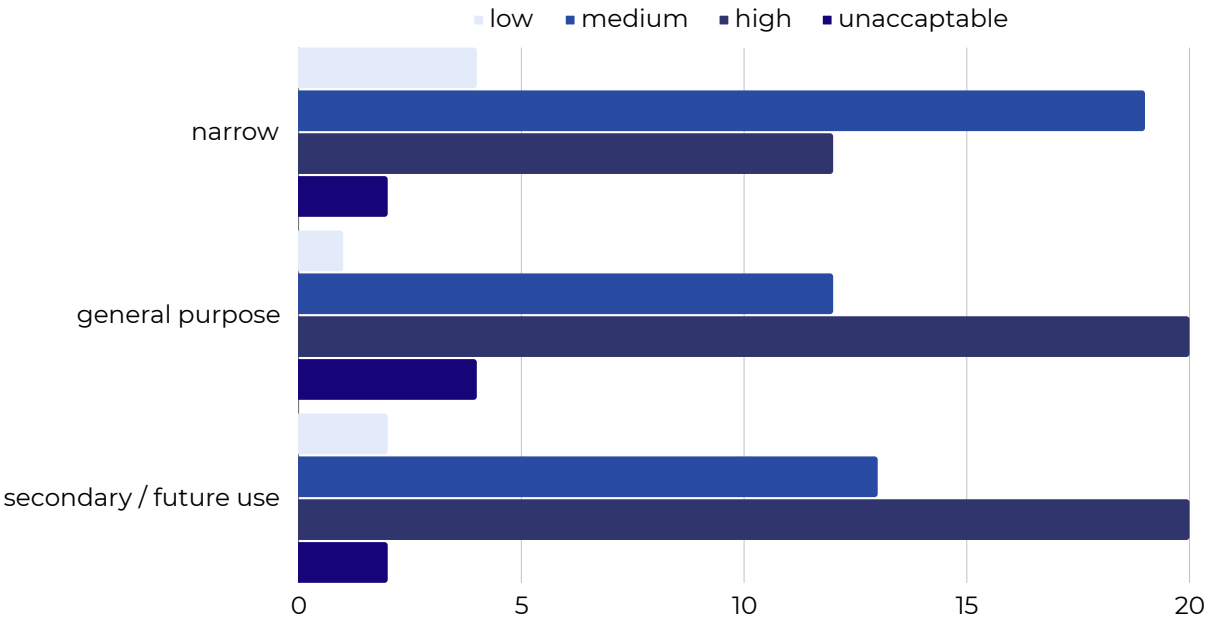


Q1 - LEVEL OF RISK

Institutions & Organizations



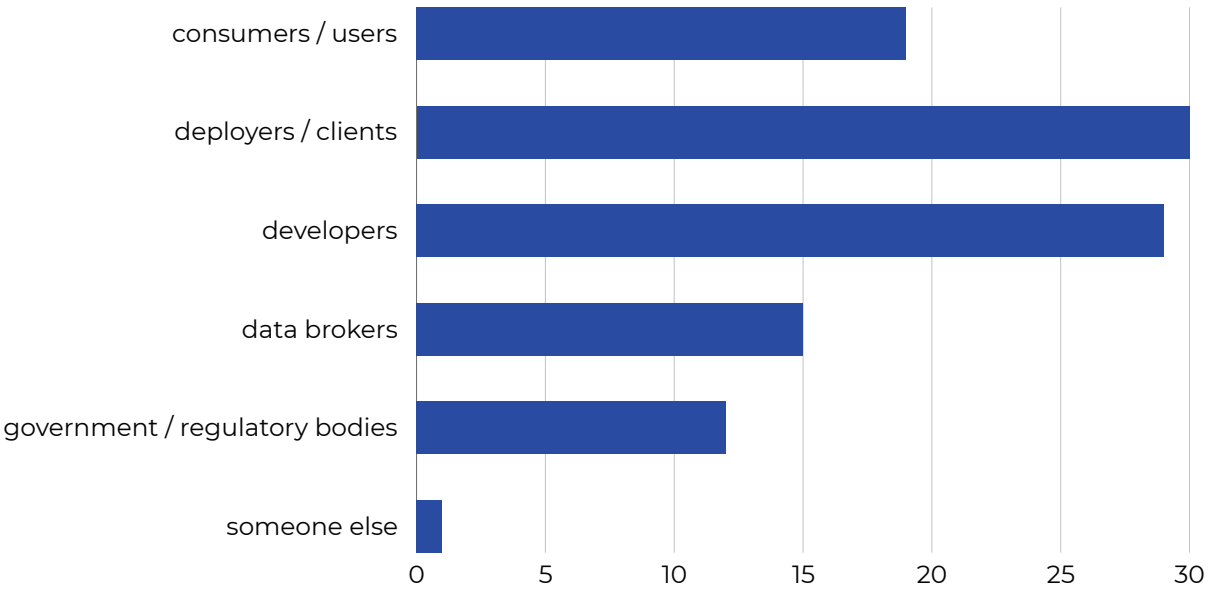
Society as a Whole



Q2.A - ACCOUNTABILITY

Participants deliberated about which parties or actors should be held accountable when an individual or group is harmed or an incorrect decision is made by an AI system using an image of a face. They then registered their individual votes and recorded their rationale.

Which parties or actors in the AI lifecycle should be held accountable when an individual or group is harmed or an incorrect decision is made by an AI system using an image of a face?



Q2.A - RATIONALE

The companies or developers who create AI face recognition technology should be held accountable if the technology is flawed, biased, or just not working.

Deployers and developers should be held responsible because it's technology they chose to use / develop that is creating problems for users. If the user is misusing the technology, however, they should have to answer for it, too.

This technology is heavily flawed and anyone utilizing it should be held responsible for harm done. As a user, you have no say in whether the system works as is claimed.

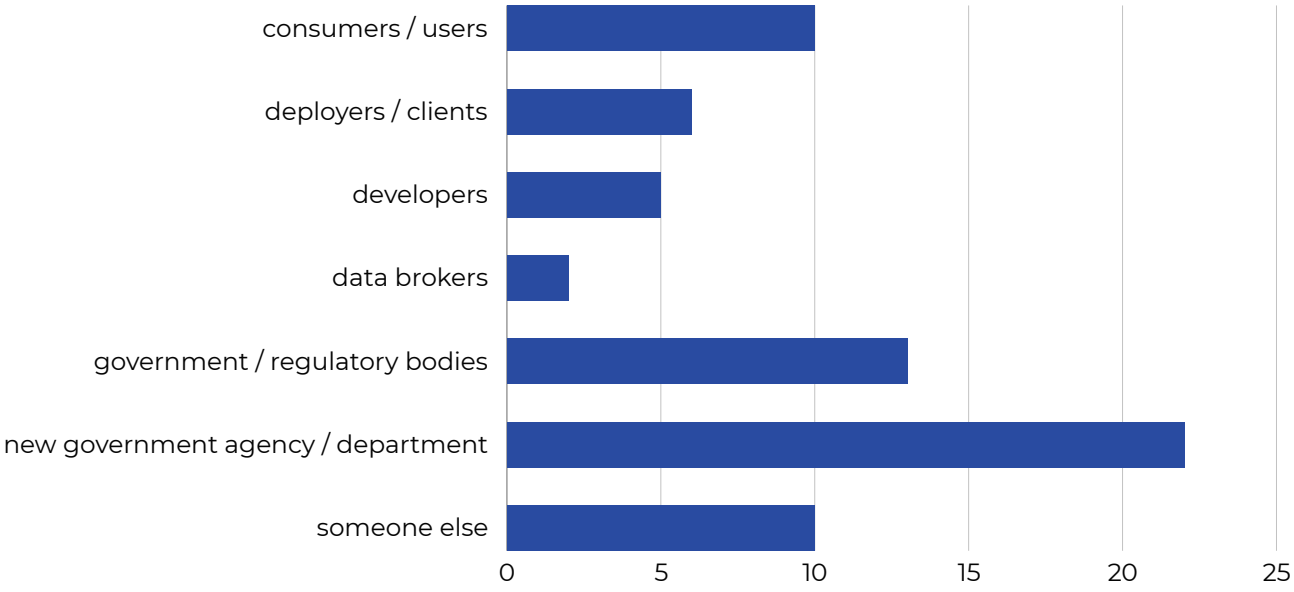
Many apps have terms & conditions stating that if you upload an image of your face, you agree to let them use it however they like (including in their training). Partial blame can be put on the user for this, but they don't often know or understand what they are agreeing to and it is important for users whose faces are being incorporated into these datasets to be aware of the terms. The majority of the blame, however, should be on developers.

The creators and / or purveyors of an AI-generated image should be held responsible for their actions since they are the ones who make an image public.

Q2.B - RESPONSIBILITY

Participants deliberated about who should determine which parties or actors should be held responsible when harm is done to an individual or group through an AI system using an image of a face. They then registered their individual votes and recorded their rationale.

Who should determine which parties or actors should be held responsible when harm is done to an individual or group through an AI system using an image of a face?



Q2.B - RATIONALE

Clients can hold developers responsible by stopping the use of their (developers') systems. Developers are able to cut off users who are misusing the system. Regulators can assess and address harm in those areas over which they have jurisdiction.

AI deployers and clients have a big responsibility to not misuse our facial data, but we need national regulations that will help determine who can be held responsible. Many times, I think, multiple parties could be responsible for harm.

Users have a right to hold the developers and deployers responsible for using a system that causes harm. Regulatory agencies exist for this purpose.

Government Agencies, regulatory bodies, judicial systems, and possible new government agencies should be responsible for determining who should be held responsible for harm done by face recognition systems.

As the one being harmed, the user is in the best position to determine blame.

ADMINISTRATIVE RECORD

US PUBLIC ASSEMBLY ON HIGH RISK AI | 2023

ADMINISTRATIVE RECORD

Assembly participants examined how an administrative record might be used in various AI systems including narrow AI, general purpose AI, as well as for secondary and future uses.

01

Uses for an Administrative Record

- Used to determine / guide eligibility for financial products, housing (renting / mortgage / refinancing), government programs, insurance, etc.
- Used to generate responses about navigating an individual's financial, legal, housing, benefits, or other matters
- Used within additional datasets to train other models that will be employed across various sectors and industries for other purposes

02

Expert Witness

Dr. Chris Meserole delivered a background presentation about the use of administrative records in various AI systems and answered participant questions about how these records could be utilized throughout the AI lifecycle.

03

Deliberations & Voting

Assembly members deliberated about the potential benefits and harms of an administrative record's use within different AI systems. Participants then registered their votes indicating which level of risk these uses might pose to various individuals, institutions / organizations, and society as a whole (including particular groups).

POTENTIAL BENEFITS

INDIDVIDUALS

- Individuals such as refugees, students, veterans, and others may benefit from AI systems that use administrative records that could provide efficient eligibility processing for various programs (governmental assistance, educational grants, etc.)
- Individuals seeking financial and other services may experience faster processing times for applications, credit checks, etc.
- AI systems utilizing administrative records could streamline application processes and simplify procedures for users in a wide range of settings
- Language or technical literacy barriers might be mitigated by AI systems using administrative records when applying for citizenship, identification, or other documents
- Individuals who are eligible for social programs such as welfare could be auto-qualified and / or enrolled by AI systems using administrative records
- Individuals working within different governmental / bureaucratic systems may have their work streamlined, alleviating the need to enter repeat information
- Widespread adoption of AI systems across government could make it easier for individuals to interact with government agencies or access information from government bodies
- Individual consumers may benefit from easier, more efficient financial transactions (such as applying for a mortgage, enrolling for insurance, or purchasing a vehicle) through increased adoption of AI systems using administrative records
- Individual employees could save time by having access to large amounts of data to support decision-making processes

POTENTIAL HARMS
INDIDVIDUALS

- Individuals may be denied benefits because of biased, discriminatory, or otherwise flawed AI systems that use administrative records
- Individuals may be at higher risk of having sensitive personal information exposed if data is made accessible and / or cross-referenced across multiple systems
- Individuals may be unaware of which administrative records about them are being used, by whom, and for what purposes; there may be a lack of transparency and consent about data collection and AI system use
- Mistakes in data entry and / or flawed datasets may lead to permanent disqualification for services (where otherwise these mistakes could have been flagged / checked by a human)
- AI systems using administrative records may not be trained to take an individuals's background information into account when decisions are made
- Algorithms built using administrative records that prioritize keywords in their models (as opposed to qualitative factors such as applicant personality, verifiable skills, "fit," etc.) may overlook or deprioritize some otherwise qualified applicants
- Individuals may find that AI systems using administrative records are slower or more complicated to navigate
- If personally identifiable data is not adequately protected within and across AI systems using administrative records, fraud, identity theft, reidentification, or other misuse of data about an individual may occur
- If current systems are developed using outdated or biased data, future AI models that are trained on this data could lead to incorrect decisions about an individual

POTENTIAL BENEFITS

INSTITUTIONS

- Public agencies may be able to utilize AI systems that employ administrative records to review / process applications for programs more efficiently (e.g. Veterans Affairs could provide benefits more quickly while saving individuals and the agency valuable time and expenses)
- Private institutions could use AI systems to process applications (e.g. eligibility for credit, loans, or other financial services), improve reporting, manage risk, and monitor client accounts
- Research institutions and organizations (such as libraries) could adopt AI systems using administrative records to gather, manage, and share larger sets of data and improve ability to conduct research and provide access to information
- Court systems may be able to utilize general purpose AI to provide information to parties who may otherwise not have access (through translation, chatbots, etc.)
- Organizations could utilize chat functions to control costs, increase worker efficiency, manage data, and enhance employee performance
- Institutions could use AI systems to organize administrative data and compile these datasets for ease of access and use by employees, which might improve claims processing, marketing, lead / sales generation, etc.
- Organizations such as businesses, hospitals, and schools may benefit from lower operating costs by using AI assistance to process and maintain data / records or aid in performing other administrative tasks (data processing, digitization, fraud monitoring, generating reports, etc.)
- The federal government and policymakers may use AI systems for administrative recordkeeping and data analysis (such as census projections and resource allocation), which could improve policy and decision making
- AI developers and tech companies could benefit from the use of administrative records in their efforts to develop new, innovative AI systems

POTENTIAL HARMS
INSTITUTIONS

- Organizations' application processes could be subject to massive fraud
- Institutions may be at risk for reputational harm, tort, and liability litigation as a result of discrimination within AI systems that utilize administrative records
- Organizations using flawed administrative data may incorrectly allocate resources, unjustly deny or disqualify applicants, or reject suitable candidates
- Companies using chatbots could lose out on sales, enrollment, or fail to meet other goals if they utilize AI systems that are ineffective or utilize outdated or biased administrative records
- Institutions or organizations using general purpose AI that produces false or misleading information based on administrative records could alienate users, jeopardize public trust, or lead to lawsuits
- Organizations or institutions that develop an AI chatbot or other feature using administrative records may not be able to keep up with the lifecycle management necessary to ensure the system remains accurate and up-to-date
- New AI systems that are created using flawed or biased administrative data could harm a company or organization by wasting money and producing bad outcomes
- Organizations that adopt AI systems without adequate data protection measures could be at risk for increased data breaches, cyber crimes, hacking, or other data security incidents that could be costly to address
- Overuse of general purpose AI systems by companies or organizations could lead to an overreliance on automated systems without human talent available to address issues as they arise

POTENTIAL BENEFITS

SOCIETY

- Expanded government use of AI systems that use administrative records could improve service delivery
- Particular groups of people who receive government and / or social services (such as veterans, people experiencing domestic violence, or people who are unhoused) may benefit from the use of AI systems which provide more responsive services and programs
- Use of AI systems using administrative records to provide more effective and efficient allocation of benefits and resources may allow for cost savings to be utilized for other programs and/or services (serving specific groups who would benefit from additional resources)
- The general public (constituents, consumers, and users) may benefit from AI systems that use administrative records by having easier, more consistent access to information
- Society's overall needs may be better met by government and social service agencies using AI systems that employ administrative records
- Communities could be empowered to make more informed decisions using AI that uses administrative records to increase access to information on a wide range of issues that affect them
- AI systems using administrative records could be used to stabilize housing markets (by assisting with tasks such filling vacancies or flagging violations of fair housing practices)
- AI systems using administrative records might help to reduce unemployment by matching job seekers with open positions that match their skill sets
- Faster and more accurate AI systems using administrative records could lead to a more efficient judicial system

POTENTIAL HARMS
SOCIETY

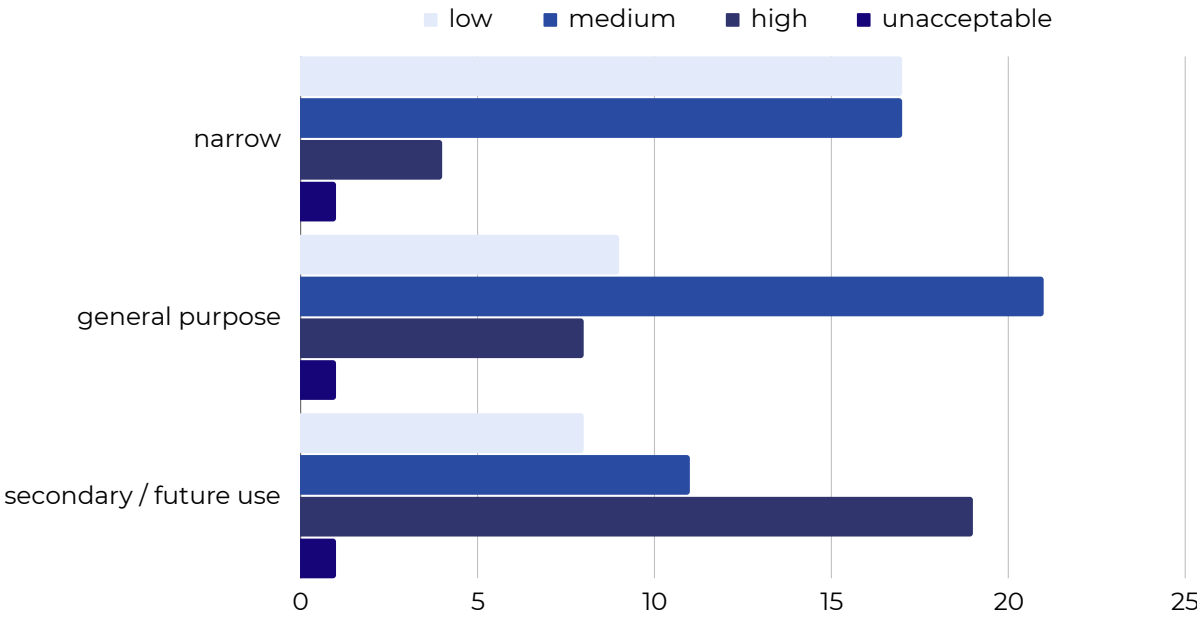
- A lack of transparency and / or accountability regarding how administrative records are used in AI systems may harm society
- Trust in government could be damaged if there is no transparency as to whether or not AI systems are biased in their decision making or lead to discriminatory outcomes and decisions (erroneous decisions about benefit eligibility, redlining, etc.)
- A larger percentage of society could be at risk of cyber crime, data breaches, and hacking when more and more administrative datasets are used by a wider range of parties
- If AI systems are used to replace human staff (e.g. processing applications for government benefits) there could be a loss of empathy amongst people and / or decreased chance to address issues from a human perspective
- General purpose AI systems using administrative records may provide discriminatory, false, or incorrect information due to biased, outdated, or poor quality data, which could decrease trust between individuals, communities, and society as a whole
- Companies that use general purpose AI that is flawed, inaccurate, or unhelpful (e.g. poor customer service) could lead to broad consumer dissatisfaction
- The sharing of data between too many systems could lead to “Big Brother,” with people fearful that they no longer have privacy
- Consolidation of administrative data by the government may lead to breaches of sensitive information about individuals, pose major national security issues, or create a black market for administrative records to be sold and used, any of which could decrease confidence in government
- Federal agencies that are required to publish inventories of their AI systems may fail to do so and this lack of transparency and accountability about how AI is used by the government could erode trust in our institutions

Q1 - LEVEL OF RISK

After deliberating about the potential benefits and harms of various AI systems using an administrative record, as well as the potential risk of these systems to various parties, participants registered their votes.

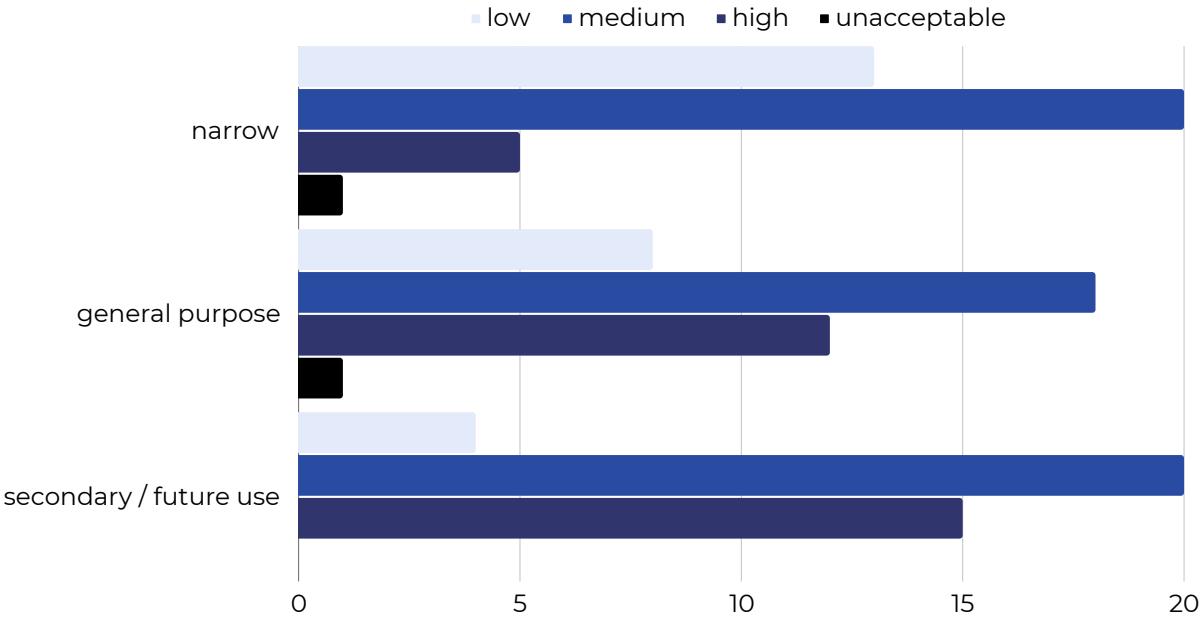
Given the possible benefits and possible harms associated with the various AI uses in “Administrative Record” please indicate which level of risk you would apply to each narrow, general purpose, and secondary / future use as it pertains to individuals, institutions, organizations, and society as a whole (including particular groups).

Individuals

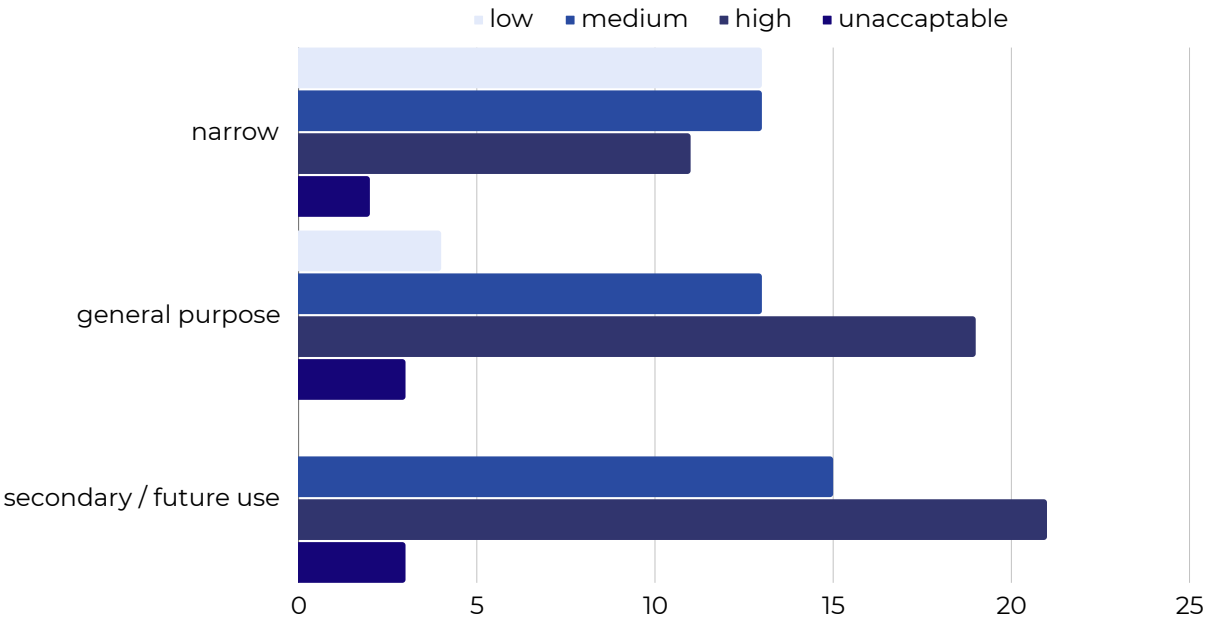


Q1 - LEVEL OF RISK

Institutions & Organizations



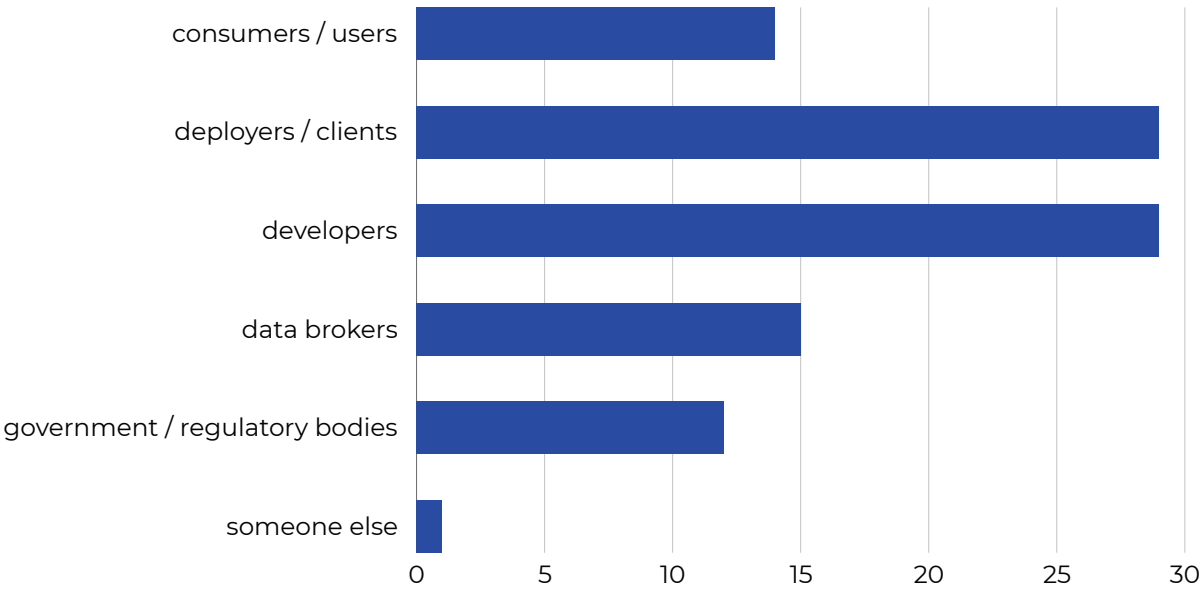
Society as a Whole



Q2.A - ACCOUNTABILITY

Participants deliberated about which parties or actors should be held accountable when an individual or group is harmed or an incorrect decision is made by an AI system using an administrative record. They then registered their individual votes and recorded their rationale.

Which parties or actors in the AI lifecycle should be held accountable when an individual or group is harmed or an incorrect decision is made by an AI system using an administrative record?



Q2.A - RATIONALE

It is everyone's responsibility to make sure that an AI system is safe, unbiased, risk-free of harm, and used in an ethical manner.

Data brokers should be responsible if the data they're peddling is inaccurate and they know it, or if it is sourced in a sketchy manner. Developers and deployers are responsible for the design of an AI system, for doing enough testing, and putting it to work safely. If they don't do this they should be responsible. Clients should be responsible if they purposely want a system that's flawed or if they keep using a system they know is faulty.

System deployers and developers should be held accountable when administrative decisions are made that are harmful to an individual or a group.

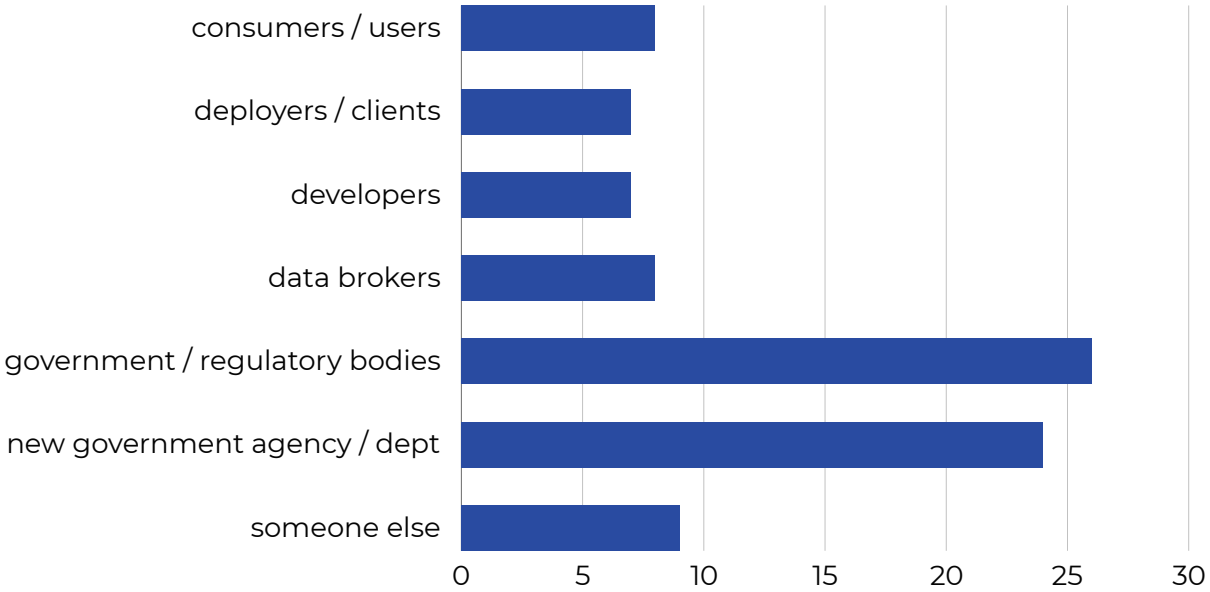
Almost everyone can be held accountable: the clients for employing an unjust system, the developers for creating an inaccurate system, the data brokers for sharing / acquiring data without consent, and the government bodies because there are not any regulations set in stone to prevent harm.

All parties should be held accountable regarding the use of administrative AI. We willingly give our data out and trust that brokers and developers will use it correctly.

Q2.B - RESPONSIBILITY

Participants deliberated about who should determine which parties or actors should be held responsible when harm is done to an individual or group through an AI system using an administrative record. They then registered their individual votes and recorded their rationale.

Who should determine which parties or actors should be held responsible when harm is done to an individual or group through an AI system using an administrative record?



Q2.B - RATIONALE

If there is not a new agency, then current government agencies (e.g. Department of Health & Human Services, Attorney Generals) need to make the public well aware of their rights and who to contact should there be a case of harm. In addition, these agencies should provide a public record of offending AI companies, developers, or programs.

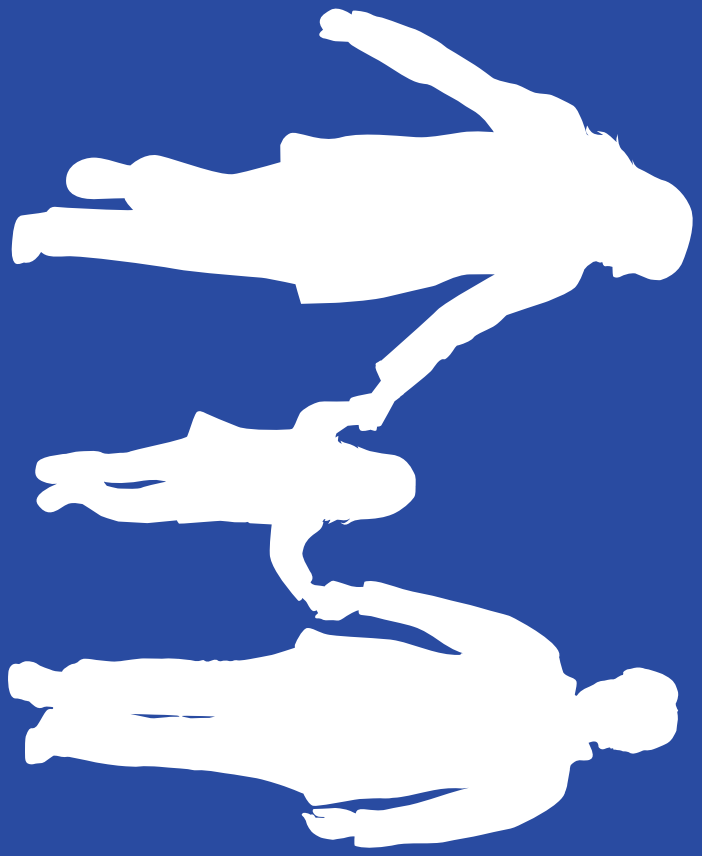
There needs to be major regulations in place so that harm from AI systems can be prosecuted and stopped. This is up to our government, and if they have to start a new agency specifically for AI, then so be it. The court systems would also play a part in finding who's at fault.

Users should always have the right to seek damages when an AI system causes harm. Regulatory agencies have a responsibility to ensure these systems function without causing harm and enforce penalties when they do.

A nonpartisan committee made up of AI experts and average U.S. citizens would be the best way to govern AI in the way the public agrees it should happen. This could prevent potentially biased agendas from influencing decisions.

Users cannot stop the sharing of their personal data, so everything must be done in order to prevent harm and, when it does, those accountable need to actually take responsibility and resolve the situation. All of these parties can possibly play a role in preventing harm and, when it does happen, they need to be held accountable.

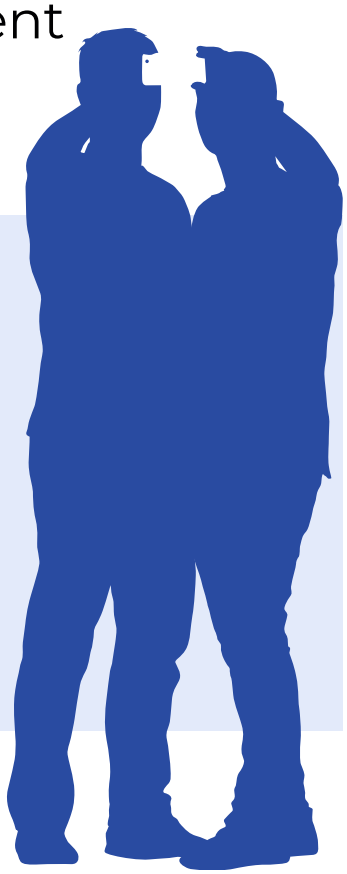
DETERMINING **HARMS**



DETERMINING HARMS

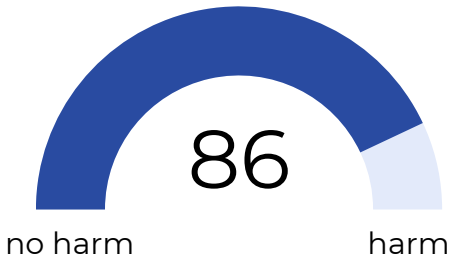
Participants deliberated about what constitutes harm across a range of scenarios. These included, but were not limited to: collection, sharing, and use of data across / within AI systems, AI system design and training, AI-supported decision making, and AI-generated content. Using a scale of 0 (no harm) to 100 (harm), participants registered their individual votes and recorded their rationale. Figures shown represent the average score.

> Through our deliberations, we have decided in our panel that while AI is a great surge forward in technology and presents many advantages and new conveniences, it is, like virtually all new technology, likewise fraught with many drawbacks and potential harms. Thus, it is important to approach this new system with caution, and to take special care to consider all of its implications, both positive and negative, when preparing to use it.



Q3.A - DETERMINING HARM

An AI system makes a technically incorrect / erroneous decision about an individual that has an adverse material impact on the recipient's life.



If the system makes an incorrect decision, and the impact is adverse and material, the harm is almost certain to be high. The fact that it is adverse and material raises both the likelihood of harm and severity of the harm.

An individual should have the right to seek reparations when an incorrect decision is made about them.

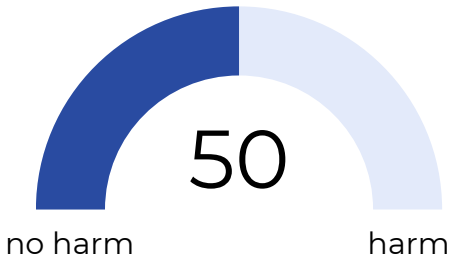
Level of harm is subjective in a lot of circumstances. What might be an adverse impact to one person might not matter at all to another. Decisions affecting mental or physical health, matters of life and death, or financial issues are more significant harms than others.

An AI system has no idea how detrimental an impact a decision could have on a person's life if, for instance, a loan isn't approved or a line of credit decreases. In some cases, it could prevent a person from putting groceries on the table, paying rent, etc. while they wait for a paycheck to be deposited.

False arrest (e.g. criminal record), application denials (e.g. rental, financial cards), and other various incorrect decisions cannot be erased and could lead to someone's life being greatly impacted in a negative way.

Q3.B - DETERMINING HARM

An AI system makes a technically correct / accurate decision about an individual that has a perceived adverse material impact on the recipient's life.



A decision may have a perceived adverse impact but if it is technically correct, in line with the decision a trained human would make and is done with oversight by a human, I see no harm. There should, however, be an avenue for arbitration.

If the correct data is used and the outcomes are correct, this has lower likelihood of harm. One might not agree with the AI decision being made (e.g. being turned down for a car loan), but a human would probably have turned you down, too.

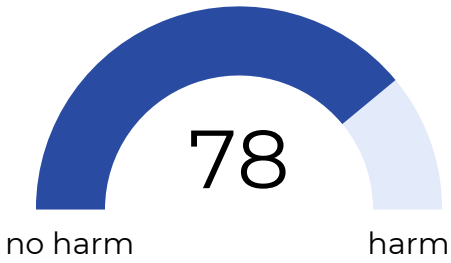
This would not be a harm (one can't always get what they want) as long as the AI system is unbiased and operates with full and fair data on a given situation. An AI system not granting a loan to someone with a poor credit score would not be a harm. If this loan denial is based on, say, an individual's ethnic background or educational history, etc., that would be a harm.

While this could be beneficial or have minimal harm most of the time, there are a few circumstances where it could be pretty harmful. Not everything is black and white, but AI is trained to output black or white answers, when in reality, a lot of things fall into gray areas.

Situations are not cut and dry. Someone may have poor credit due to poverty but have just gotten a new great job, so a human might see their potential, while an AI system might deny a credit or rental application. This can also happen in medicine if a patient cannot advocate for themselves when a doctor is using AI to assist with the diagnostic process.

Q3.C.I - DETERMINING HARM

An AI system makes a decision or produces an outcome that violates an individual's civil or human rights which is technically correct.



Violating civil rights is wrong whether it is done by a human or a machine.

Human and civil rights should never be violated! Even though they are violated all the time by people and institutions, having AI violate them, too, makes it even worse.

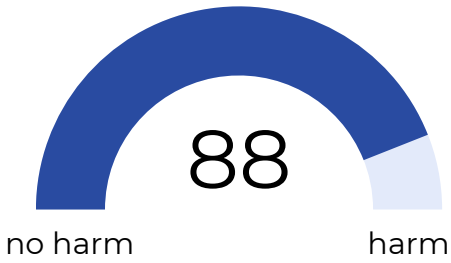
Are human rights violations by fellow humans not already considered uniformly harmful? I see no reason why this would change if the violation was perpetrated by a machine. Perhaps the blame falls more directly on the operators or programmers in this situation.

A violation of human rights is unacceptable, regardless as to whether or not the outcome of the AI system is technically correct.

There were no valid situations provided in our group deliberations where this would be okay.

Q3.C.II - DETERMINING HARM

An AI system makes a decision or produces an outcome that violates an individual's civil or human rights which is technically incorrect or erroneous.



Any violation of human or civil rights (intentional or not) is wrong, whether it be enacted by a human or a machine.

Even in an outcome with a technically correct decision, the violation of civil or human rights should not be tolerated under any circumstance. The scale would need to go beyond "100" to properly describe how horrible it would be were somebody's rights violated from an incorrect decision.

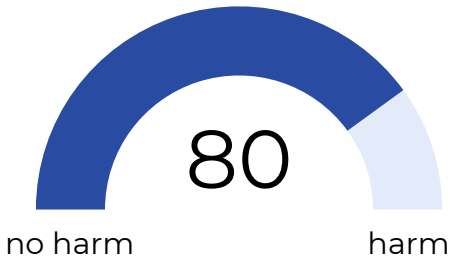
A violation of civil and human rights is just that: a violation. Data that is incorporated into AI models should adhere to regulations and federal laws and not violate civil and human rights.

This would be a disaster for a group or individuals as they may not have any recourse if an incorrect decision is made about them.

This is the fault of those that built the system and not the system's fault. Once it has been set up and operates with bias, it becomes harmful.

Q3.D - DETERMINING HARM

An AI system produces content (text, video, audio, etc.,) that is discriminatory or biased against particular communities or groups.



A system that produces discriminatory or biased content is extremely dangerous because it perpetuates biases and discrimination.

Any inflammatory, incorrect, biased information, or content is very harmful. Black, Indigenous, and People of color (both as communities and as individuals), various religious groups, and different segments of society can all be victims of this. AI systems start with people and develop due to information from human sources.

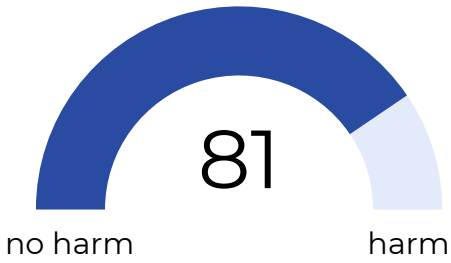
While AI being used, say, for credit or employment should never be discriminatory or biased, the act of producing content such as text or images is not as easily classified as "harmful." Hurtful, perhaps, condemnable and problematic, yes, but freedom of expression must also be upheld.

The purpose of using an AI system should be to prevent discrimination or bias. Having it produce content that does exactly the opposite should be frowned upon.

The content itself may not be harmful, but it can perpetuate, extend, and amplify biases, discrimination, and hate speech, because of the content generated.

Q3.E - DETERMINING HARM

An AI system appropriates an individual's likeness (image or audio) without their consent.



If an AI system appropriates a likeness it could or could not be harmful. But, in today's society, likeness and audio are always appropriated by numerous sites and systems without consent. It depends on how the imagery or audio is used that would actually constitute harm.

There are many public sources of images and likenesses, some of which are in the public domain. Ideally, there would always be consent given, but if someone's image is used in a derogatory way, there needs to be consequences.

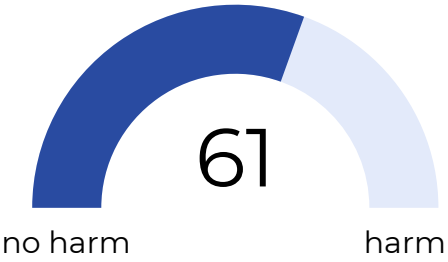
Appropriation of an individual's likeness without their consent could be very harmful if used for negative or illegal purposes.

Anything done without someone's consent is harmful. This can damage a person or company's good name or reputation.

All people should be protected under copyright law. While generative AI technology can create content using voice clips of Taylor Swift to generate a completely new song, being famous does not mean someone doesn't deserve copyright protections. This could also be used to create deepfake pornography or deepfake content in other situations.

Q3.F - DETERMINING HARM

An AI system recommends or provides information and / or content that privileges the ideas, content, posts, and / or products and services of some groups over others through the use of an AI algorithm (recommends, privileges, siphons, sorts, directs, filters).



Companies, political candidates, nonprofits, or anyone else instituting an AI system is trying to further their own self interests. It is common knowledge, and common sense, to take a lot of this information with a grain of salt. Other aspects of AI are much more dangerous.

I'm pro-targeted advertising, pro-editorial freedom, and support opposing views. There is some danger of individuals being targeted with disinformation that is tailored to them specifically, perhaps to sway political views or actions. While the vast majority of commercial uses may be harmless (and in some cases beneficial), the potential concerns make this moderately harmful.

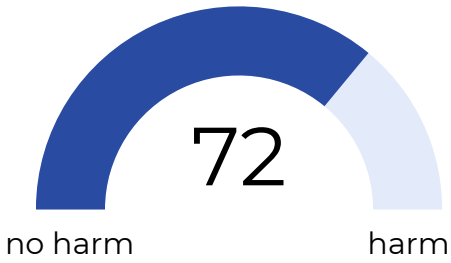
This can seem very harmless, but there are extreme cases where it can be very harmful. Beyond that, even some of the less-harmful aspects may arguably still be harmful. It seems convenient to have content similar to what you enjoy recommended, but it also risks putting you in an echo chamber where your thoughts are never challenged.

Discrimination is not always negative. It might be harmful to certain groups when a ranking system recommends some content over other options. Conversely, an AI system that ranks via customer service points might highlight and present new or better data to an individual.

We ask AI to do this.

Q3.G - DETERMINING HARM

An AI system appropriates an individual's data without their consent.



How was the information used? Did it cause harm by being used in an inappropriate way? AI systems are made to sort and generate information, business leads, or other data, but many people are unaware how to control the sharing of their information.

Everyone should have control over how their data is used, no matter how it is used. Being compensated for its usage would be a great change, as well.

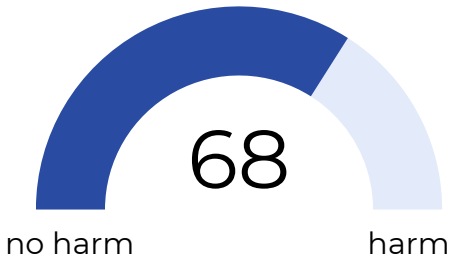
Right now I believe that we cannot, and will never be able to, give companies / institutions consent to use our data because data use and sharing is far beyond our control. My main questions when it comes to harm are "What is being done with the data?" and "What decision is being made?"

If the data is accurate, then no harm, no foul. If it means I have to reclaim my life because of falsehoods or lies, then identity theft might be happening [which would be harmful].

There is no reason any system should use an individual's data without consent, regardless of what it is used for.

Q3.H - DETERMINING HARM

An AI system is used without an individual's knowledge and / or consent to reach a decision about them.



This can be good and bad. When an AI system makes a decision about an individual without their knowledge or consent, they may have no opportunity to influence the outcome of that decision. This can lead to decisions being made that are not in the individual's best interest.

There should be a notification and / or legal paragraph of disclosure when AI is used to make decisions regarding certain decisions such as banking, credit, education, etc.

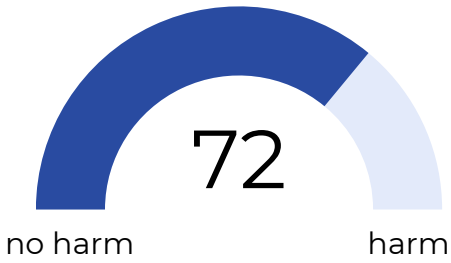
I find this to be a harm since it may be using data from social media accounts from years or decades ago. I'm a much different person now than I was then, so if I'm unable to contest the data and / or don't have some control over how my personal information is being used, I may be incorrectly judged.

Using an individual's data without their knowledge to reach a decision about them could range from being mildly harmful (not getting a loan) to very harmful (not getting life-saving healthcare).

This happens constantly. It can be subtle or more pronounced, but it is basically out of my hands.

Q3.1 - DETERMINING HARM

An AI system is used to make a decision about an individual but where an explanation by a human cannot be provided.



Decisions that affect me should be explainable. There needs to be a way to tell me why the decision was made the way it was so I can either challenge it or make changes so that it will come out differently next time.

These AI systems are made by humans and should be explainable by a human. If the system is producing results that cannot be explained, it should not be used.

This could be harmful if companies use "It's based on the algorithm," as an excuse to charge higher prices for services. Someone wanting an explanation about an increase in their utility bill and billing, if based on an algorithm, may have no way to investigate or have the changes explained (e.g. if they know they haven't increased their use of gas but get charged for more).

This can go both ways. If it is a good decision someone agrees with, they may not need / want an explanation. But, in the case of a denial or a bad decision, someone may want an explanation that can't be given.

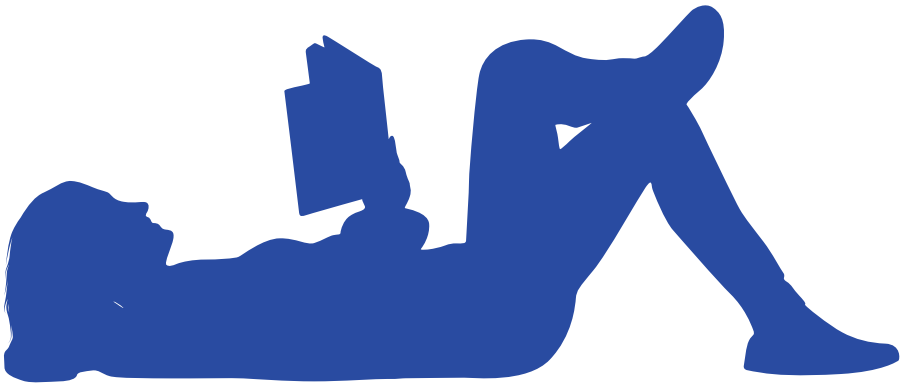
There is no harm if nothing can be done about it.

ASSEMBLY MEMBERS **ON THEIR WORK**

FROM ASSEMBLY MEMBERS

Participants share about their experience serving on the AI Assembly, what it was like working with others from across the U.S., and what they want the public to know about AI.

> The moral use of AI is an incredibly delicate balance. It can be so easily abused even though it simultaneously can be so beneficial. It's not just a little input / out table and it's not some malicious entity, either. AI is so vast and complex, there's no way to put a solid definition on it that encapsulates every possibility. There are so many possibilities, so we need to stay alert and vigilant against potential malice or dysfunction, but we should also be excited for the future and the ways that this might change the world as we know it.



This has truly been the experience of a lifetime. The opportunity to go on this journey, learn so much about something I knew nothing about, and engage with peers from various demographics and walks of life has been phenomenal. This experience has really shown me how differently and similarly we think and really brought to light that everyone's thoughts matter.

I had no clue what AI consisted of. I am still not an expert. I will not allow this to stop me from realizing that technology is constantly evolving. The unknown does not equal something bad. It just means we need to open our minds and constantly be seeking knowledge. We came together as a group of strangers and formed a community to work together for the better of society. I hope we made a positive impact and can see our thoughts and ideas somewhere in the near future through AI.

After coming together for two weeks with my fellow peers from across the States, we have learned so much about AI and it's uses and harms... It was an eye-opening experience for me to learn so much about AI and I would recommend everyone should learn more about it.

It makes me so happy that people from all over the country are on the same page no matter our race, religion, sex, etc. This hits home to me as a Marine.

It's been really exciting and encouraging to see, first-hand, how a cross section of people in the U.S. come together and not only learn more about AI, but also decide what is important to consider as AI continues to be used (for good and bad), developed, deployed, and regulated.

We have learned so much about AI these past two weeks and have shared our thoughts about what we have learned. The general public does not have enough information about AI and its uses, both good and bad. Those organizations involved in the future development and deployment of AI systems should have a greater focus on informing society, groups, and individuals about the systems and their specific use. Individual and group data should be better managed and protected.

Every part of this experience has been important. I would share that even though experiences are new and unknown, that shouldn't prevent you from trying it out. I left my comfort zone to participate in the Assembly. I almost did not attend. I am more than glad I did. Getting to know new people, hearing others' views, and learning how to work together to make a change was life-changing.

This experience was insightful and fulfilling. I was able to be a part of something amazing. I appreciated the small group and large group discussions we had where my fellow Assembly members may have amplified my stance or helped to change my mind by bringing up a different point of view.

I was pleasantly surprised to hear so many different points of view from my fellow members that I would not have considered. This, combined with the experts, allowed me to learn so much about AI and its potential.

AI is still a newly developing field, but it is developing at an incredibly fast pace. Those involved in AI, including the general public, need to understand the enormous potential for good that AI can bring but also the enormous harms it can cause. I hope our work together will provide insight to those involved in the creation, use, and regulation of AI as well as help the public better understand its benefits and risks.

As an optimist, I am excited about the innovation and usefulness of AI. I also realize the potential of AI for harm and misuse. To the developers involved in this field of science, I plead to you to do the right thing and govern yourself and peers until the rules get worked out and a regulatory body can be put in place that will protect us from the potential harm that this innovation's misuse might present while unregulated.

AI already affects almost everyone's life in the U.S. ...We represent the everyday people of America. Our concerns are very legitimate. Our neighbors and friends need to be aware of the things we have learned these past two weeks. I hope that the final report resulting from our group input will be the beginning of regulations and protections for all of us.

It's important to know that such a diverse group of people share a lot of the same concerns about the technology and where it may be heading. This experience helped me to better understand how things work. I'm glad I was a part of this and appreciate everyone that participated.

It has been an interesting and eye-opening experience. Learning, working, and deliberating with my fellow Assembly members has been really interesting and exciting. The expert witnesses were engaging and informative. It has been really great to be able to work and deliberate with a diverse group of people who were so willing to share their ideas and views. I have learned so much from everyone involved in this assembly.

I've learned a lot over the course of this Assembly and, more than anything, I realized how little I know. Many of my colleagues and I have some level of excitement for the future of this tech, but also a great deal of concern. Both the wider public and those drawing boundaries and enacting legislation need to know as much as possible. There is a prerogative to learn more on one own's volition, but I also think that more of this information out to be actively presented to the public.

It's important to know about the concerns we all have been raising throughout this Assembly session as they are important to address and remediate for future development and deployment of AI. We all came with some or little background about AI but we all learned a lot here and became aware of what others thought could be.

It has been one of the best experiences I've had in a long time...I'm glad that I stepped out of my comfort zone and did this. I wanted to try something new, and I want to thank everyone for helping me achieve that.

Center for New Democratic Processes

Syracuse University
Maxwell School of Citizenship & Public Affairs

Published by the Center for New Democratic Processes