

AI GOVERNANCE: ARE WE RISING TO THE CHALLENGE?

Daniel S. Schiff

Purdue University, Department of Political Science

Transitioning from Digital Transformation (DX) to the Age of AI (AX): The Role of Humanity in this New Era | Seoul, Korea

Korea Research Institute for Vocational Education and Training, the Korea Ministry of Education, and The Korea Economic Daily

Debates, dilemmas, and decisions

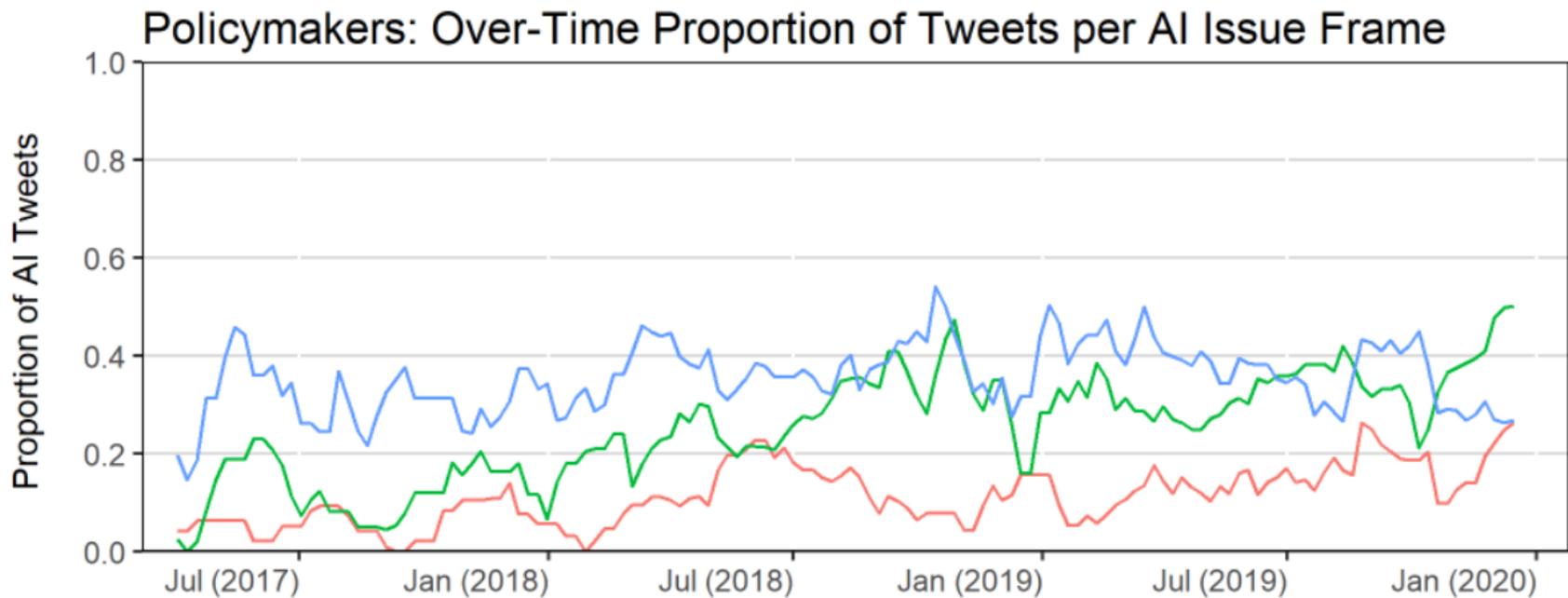
- **Competition or cooperation** (international politics)
- **Centralization or decentralization** (vertical or horizontal regulation)
- **Formal regulation or industry self-regulation** (shared governance)
- **Diverse stakeholder participation or expert leadership** (participatory democracy)

Anything else during Q&A is welcome (AI governance, ethics, AI in HR, education, future of work)!

1. Competition or Cooperation?

Competition Rhetoric 2017-2020

Keyword — Competition — Ethics — Innovation



Schiff (2022)



Marco Rubio @marcorubio

While we spend our time fighting each other a China-centric technology future is taking shape. If we don't wake up & get moving #China will be setting global standards for AI, gene editing ethics, 5G, & controlling global data flows & digital commerce.

Dysfunctional U.S. needs "Sputnik moment" on future tech
The U.S. isn't seeing the urgency of this Sputnik moment.
[axios.com](#)

7:11 AM · Jun 27, 2018 · Twitter for iPad

172 Retweets 41 Quote Tweets 447 Likes

Growing Action on Competition

Political rhetoric (U.S. and China) → action

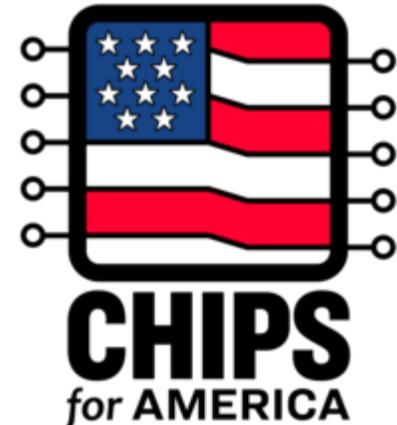
- See U.S. Innovation and Competition Act, CHIPS Act

Growing exports controls

- HPC and AI chips, semiconductor manufacturing equipment
- AI for surveillance, banned entity list
- Training datasets, licensing requirement

STEM talent

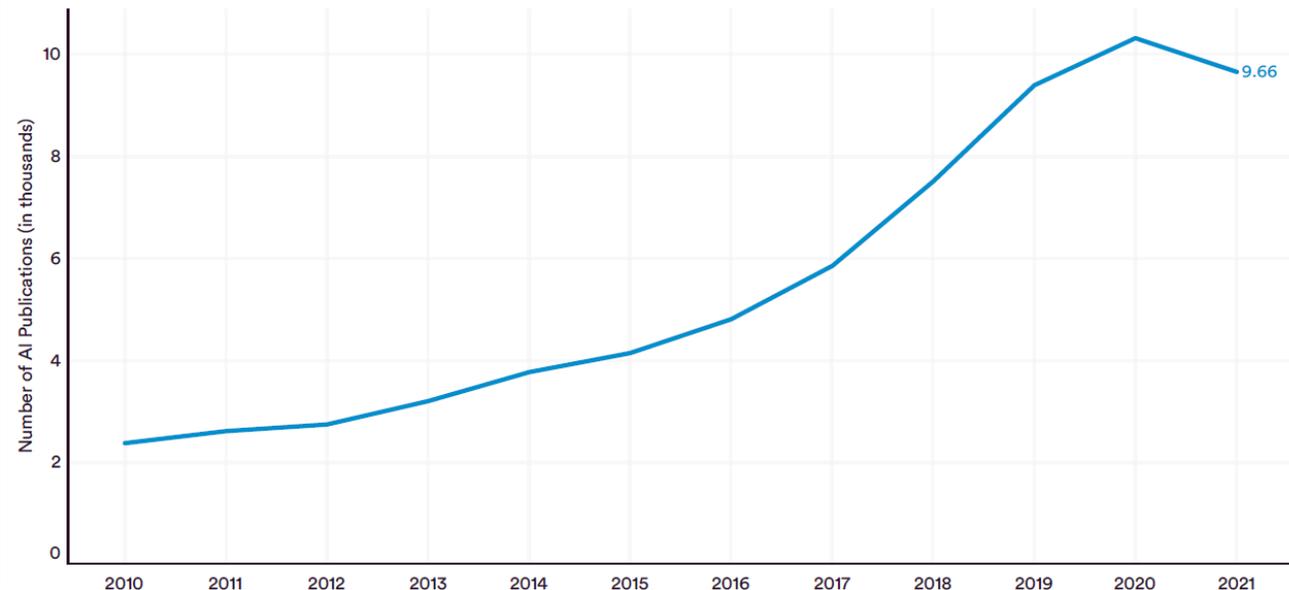
- Ubiquitous international efforts to compete over talent
- National Security Commission on AI
- Immigration/visa issues serious problem for 2/3 of AI researchers in the US (Zwetsloot et al. 2023)



But Cooperation Continues!

UNITED STATES and CHINA COLLABORATIONS in AI PUBLICATIONS, 2010–21

Source: Center for Security and Emerging Technology, 2021 | Chart: 2022 AI Index Report



ETO AGORA beta

[Documentation](#)

Welcome to ETO AGORA (AI GOvernance and Regulatory Archive), a living collection of AI-relevant laws, regulations, standards, and other governance documents from the United States and around the world. Updated regularly, AGORA includes summaries, document text, thematic tags, and filters to help you quickly discover and analyze key developments in AI governance.

 AGORA is in beta! We are actively adding new data and features. Data currently in AGORA are incomplete and subject to change. You can [share your feedback](#) to help us improve or [read our roadmap](#) to see what we're currently working on. 

DOCUMENTS COLLECTIONS

497 documents found [DOWNLOAD DATA](#)

FILTERS APPLIED: --none-- SORT BY ASSESSED SIGNIFICANCE 

FY2021 NDAA
Authority: United States Congress | Last activity: Enacted | Last activity: 2021-01-01
 Tags in progress
[View full text](#) | [View source](#) | [Report an issue](#)

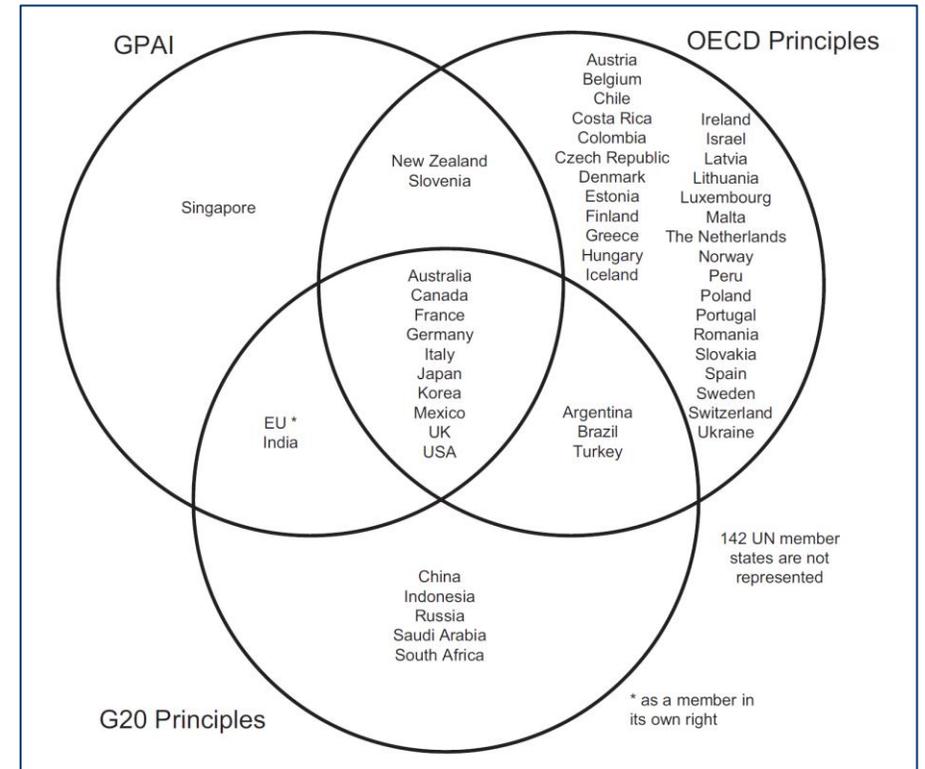
Safe and Secure Innovation for Frontier AI Act (SB 1047)
Authority: California | Last activity: Defunct | Last activity: 2024-09-30
 Requires developers of frontier AI models to implement comprehensive safety and security protocols, including cybersecurity

<https://agora.eto.tech>

Numerous Multilateral & Bilateral Efforts

- OECD, UN, APEC, Council of Europe, Vatican, African Union
- Global Partnership on AI (GPAI) (G7+)
- US-EU Tech-Trade Council (TTC)
- UK-Singapore, Canada-France, US-Japan, US-UK, India-UK, Africa-France
- The AI Safety Institutes

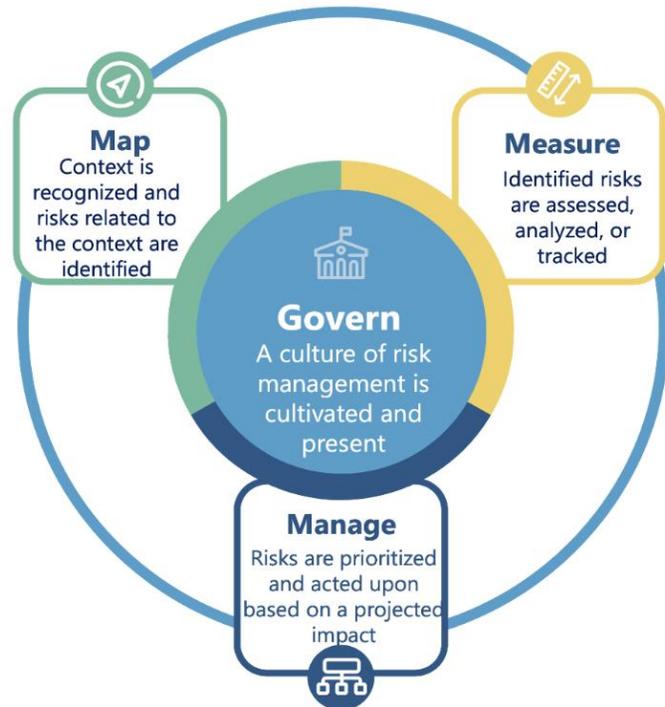
Intensive efforts toward cooperation ongoing!



Cihon et al. (2020)

2. Centralization or Decentralization?

Decentralized Governance Advancing in US (Centralized in EU)



US NIST AI RMF (2023)

Action by US sectoral regulators: finance, transportation, health, etc.

- FTC, EEOC, DOC, NHTSA, FDA, SEC, FAA, DOJ

NIST AI Risk Management Framework (RMF) is paradigmatic here

- Also see White House Blueprint for on AI Bill of Rights

Horizontal governance still in play:

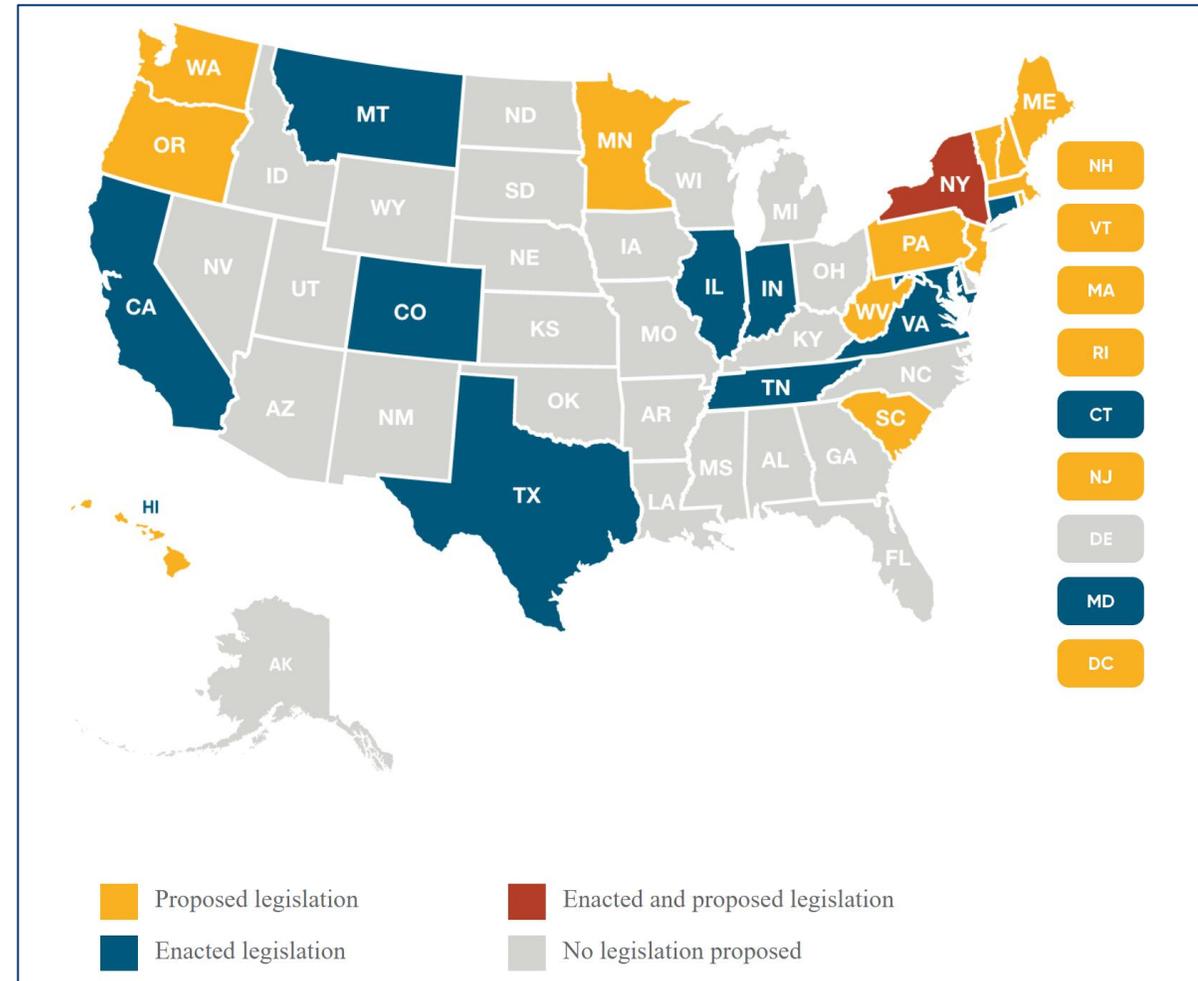
- AI Incidents/National Algorithms Safety Board
- Federal Digital Platform Commission
- Algorithmic Accountability Act

Governance not 'solved' but many advancements in responsible AI, AI ethics standards, corporate and design practices, etc.

Subnational Governance Advancing On Many Issues

Local actors are also filling gaps

- Facial recognition, biometrics, privacy
- Employment & surveillance
- Model inventories
- Platform regulation
- Workforce & education
- Impact assessments



3. Formal Regulation or Industry Self-Governance

Not a Binary

- Historically, self-regulation and formal regulation can complement each other
- Govs don't (currently/ever?) have capacity to audit every system (hence risk tiering)
- Industry still has special knowledge and access
 - Especially upstream developers
- Leading companies are active in shaping safety practices, responsible AI, etc.

TECH POLICY

AI companies form new safety body, while Congress plays catch-up

The Frontier Model Forum is the latest sign that industry is racing ahead of policymakers who want to rein in AI

By [Cat Zakrzewski](#) and [Nitasha Tiku](#)
July 26, 2023 at 6:00 a.m. EDT



President Biden delivers remarks on Artificial Intelligence in the Roosevelt Room of the White House. (Demetrius Freeman/The Washington Post)

MOST READ TECHNOLOGY >



1 Credit card fraud isn't going anywhere. Here's how to protect yourself.

2 Agence France-Presse takes 

Growth of the AI Ethics/Governance Audit Ecosystem

EAIDB Responsible AI Startup Ecosystem July 2023

The grid displays logos for various AI startups, categorized into five main groups:

- Data for AI:** FAIRGEN, visual, DEDOMENA, illumr, RYVER AI, gretel, citibeats, datagen, TRUATA, MOSTLY-AI, GOX, syntheticData, SIMERSE, Visym Labs, tomTA, MDCLONE, Datovillage, CloudTDS, replis, Humans in the Loop, Syntheticus, Galileo, Sarus, LINKAI, ZUMOLABS, clearbox, prifina, Airbloc, Unbiased, Panalyt, StageZero, betterdata, destined ai, SYNDATA, SYNTHO, PRIVATEAI, SYNTGRA, inrupt, SYNTHESIZED, modulos, mindtech, Sourcepoint, Mirry.AI, SYNTRIC AI, co-census, VISIONAI, Synthesis.ai, tasq.ai, TOMIC, Regulation, Diveplane, onetrust, Snorkel, privyca, YData, Datomize, APHERIS, TripleBlind, EDER LABS, OCTOPIZE, hazu.
- ML & ModelOps:** Galileo.XAI, LatticeFlow, XAI by Aiyas, SOLASAI, deepchecks, THEMIS AI, Verta, Etiq AI, Akira.AI, aporia, ZETANE, DeepBrainz AI, daios, dotatron, SigmaRed, ModelOp, mona, DEEPLY, WHYLABS, arize, nannyML, trubrics, fairplay, Arthur, truera, trstai, KOSA.ai, censius, fiddler, mind FOUNDRY.
- AI Security:** AISHield, TROJ.AI, Protect AI, ADVERSA, HIDDENLAYER, CERAMIC, MITRAL SECURITY, CALYPSO AI.
- Alternative ML:** FedML, DynamoFL, Flower, UMN.AI, causalens, integrate.ai.
- AI GRC:** fairnow, KomplyAI, tenzai, Mission Control AI, Saidot.ai, anch.AI, KONFER, YOOI, code4thought, QUANTPI, ETHICS GRADE, TITANIUM, TRUSTIBLE, Giskard, ROBUST INTELLIGENCE, supervise, PureML, 2021 AI, ARMILLA, MONITOUR, RELYANCE AI, calvinrisk, VALIDATOR, Vera, preamble, FAIRLY, credo ai, CertKOR AI, SECLEA, C-THRU.ai, proddago, Holistic AI, Adalan AI.
- Open Source:** Rebuff AI, Guardrails.ai, ALIBI EXPLAIN, lens, SDV, FIDLES Auditor, PHOENIX, AVID, Fairlearn, PiML, OpenDP, MI DATALAB.
- Model Builders:** seekout, gallio, black, VALOSSA, algoface, Reejig, knockri, MATHISON, multitudes, PIPELINE, syntonym, Headai, sapia, QuadFi, checkstep, PROVEN BASE, Stratyfy, textio, LELAPA AI, Recast AI, Lauretta.io, Exparang, MODULATE, actium, TRUSTLAB, CROSSCHA, impress, PickleTech, brighter AI, piktid, cohere, Kanarys, abodoo, PAVE, ZELROS, Ambient.ai, WHITEBOX AI, Tengai, digiLab, Alva, evispot, Virtuosus AI, HUMANLY, EQUALTURE, spectrum labs, BrightHire, Bodyguard.ai, DPTIMAL, BAST, ALIGNED AI, Celantur, adtlede, BRA, retorio, Cangrade, XOPA AI, KNAC, eskalera, DIVERSIO.
- Consulting:** AI Governance, Ulysses.ai, AlEthica, BNH AI, Social Tech Lab, Cantellus Group, innovethic, ORCAA, eticas, ETHICALLY ALIGNED, Adaptive, Best Practice AI, hal, h, DEXAI, lewand.ai, ASSESSED INTELLIGENCE, INQ, AlyData, BUI CONSULTING, Anekanta AI, VIRTUE CONSULTING, INTELLIGENT TECHNOLOGIES, VERAI, Kairoi.

- Governance consulting
- SaaS
- Technical audits
- Intermediary orgs promoting accreditation
- ForHumanity, Responsible AI Institute, IEEE CertifAled
- Ecosystem growing!



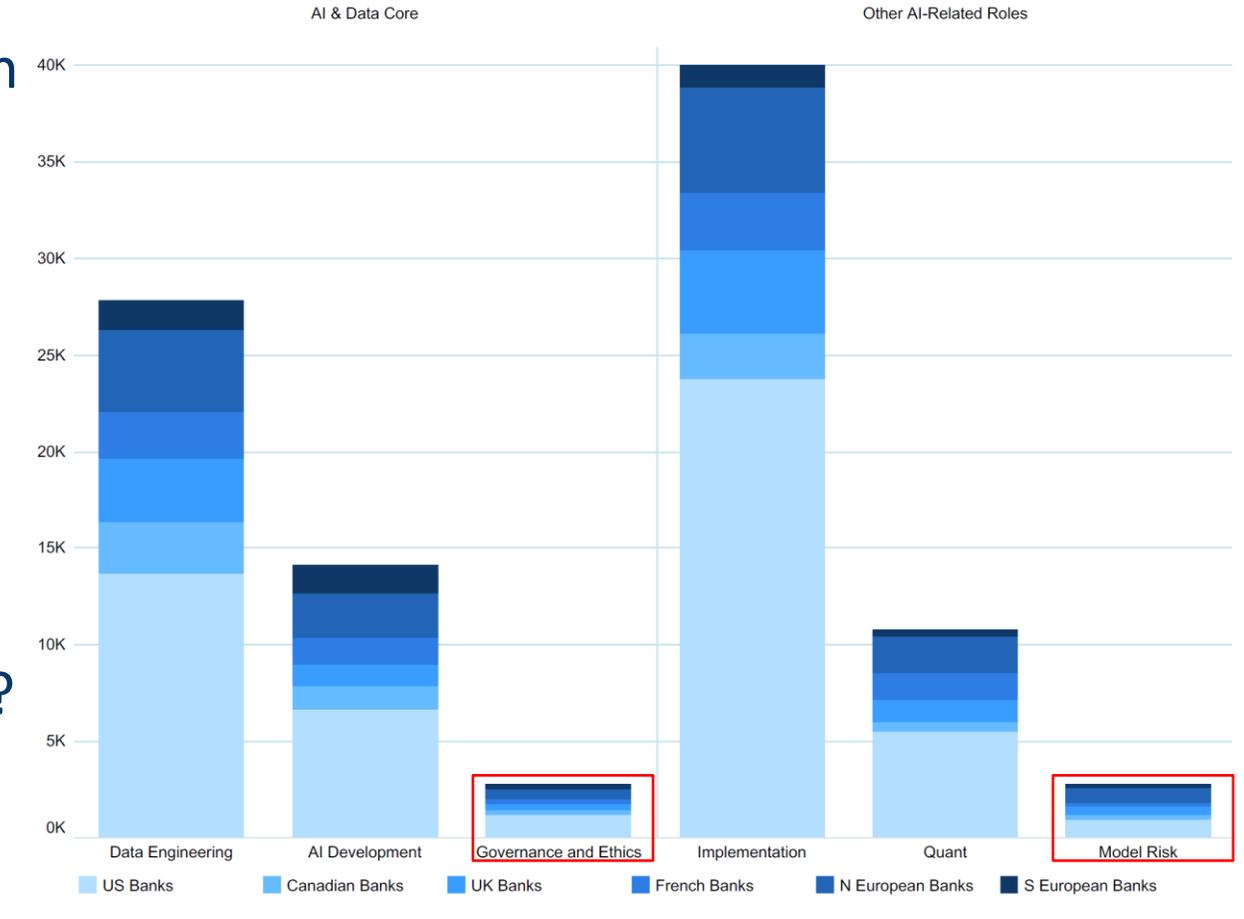
4. Diverse Stakeholder Participation or Expert Leadership ?

Consistent, Widespread Calls for Diverse Participation

- Disciplinary diversity: AI ethics, social sciences, humanities
- Lived experience diversity: gender, sex, race/ethnicity, disability, income
- Global diversity: low-income countries, linguistic diversity
- Public participation: general public, civil society, vulnerable groups

Progress or Rhetoric on Interdisciplinary Diversity?

- Observation: focus on STEM expertise in hiring, fellowships, etc.
- Vs. ethics or governance
- What # or ratio of experts in ethics, design, policy, history, etc., do we need? What team structure?

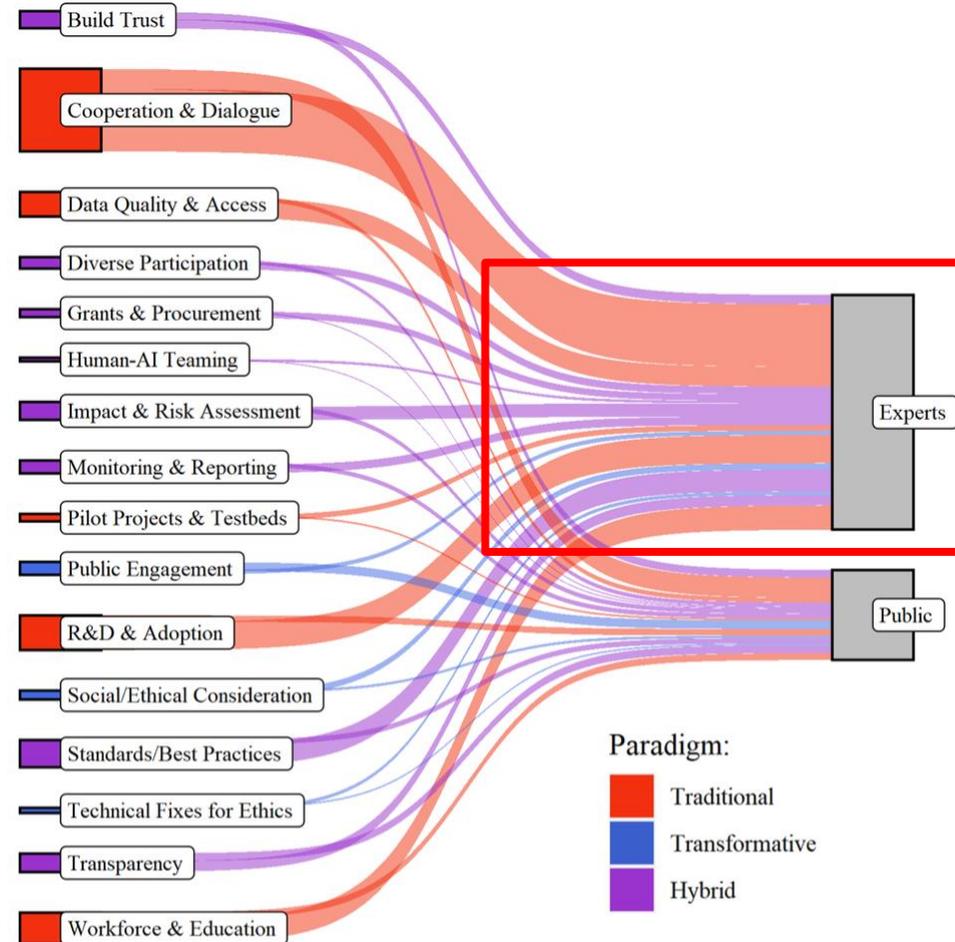


Evident AI Index (2023)

Public Participation Mostly Nominal

- Growing public opinion research
- Meaningful projects on participatory governance

but....



Schiff (2023)

Room to Foster Public Participation & Democratize AI

- Still opportunities to create channels, mechanisms, and learn from other disciplines and regions



Addressing Challenges in AI Public Opinion Research: Introducing the AI SHARE Database*

Chloe Ahn[†] Kaylyn Jackson Schiff[‡] Daniel S. Schiff[§]
Zachary Peskowitz[¶] Yu Lu^{||} Yunzhe Liu^{**}

October 23, 2024

Abstract

Public opinion research related to artificial intelligence (AI) has dramatically increased over the past several years, tracking the need to better understand attitudes and policy preferences toward a critical emerging technology. However, efforts to produce aggregate insights suffer from challenges related to conceptualization, measurement, data collection, and coordination. In response, we have compiled a comprehensive database of AI-related survey measures that we call AI SHARE, the AI Survey Hub for Attitudes and Research Exchange. To build AI SHARE, we collected and conceptually coded polls, surveys, and survey experiments ($N = 218$ at launch) from academic papers and non-academic sources released between 2014 and 2023. We find that extant research exhibits a geographical bias and fails to examine key sub-groups; often conceptualizes AI too generally and overlooks differences in AI knowledge; and lacks reliable and validated measures. We offer concrete recommendations to improve survey research and to enhance coordination between researchers, facilitated by AI SHARE.

Keywords: artificial intelligence, public opinion, survey, survey experiment, database

Where does this leave us in 2024?

Extensive efforts across sectors, regions, and scales

The work of standards bodies and auditing orgs will be critical

AI is still not highly politicized, and this may be desirable

Opportunities to advance meaningful public participation

Important: still risks of weak regulatory regimes, capture, competition, overly-narrow governance solutions, etc.

So the work & experimentation must continue!

THANK YOU!

Daniel S. Schiff
Purdue University, Department of Political Science

dschiff@purdue.edu | @dan_schiff