
To cite this presentation: Pendyala, V.S. (2024) "Responsible AI is everyone's responsibility". 10th International Conference On Mathematics And Computing , ICMC 2024 keynote
Responsible AI is everyone's responsibility:
Navigating the ethical landscape
VISHNU S. PENDYALA, PHD
SAN JOSE STATE UNIVERSITY
Video Recording:
https://youtu.be/cegd5itOgUg?t=7220
©Vishnu S. Pendyala This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License


Why it matters...

'Godfather of AI,' ex-Google researcher: AI might 'escape control' by rewriting its own code to modify itself

Published Wed, Oct 11 2023•8:30 AM EDT

Tom Huddleston Jr.

SHARE

Source: https://www.cnbc.com/2023/10/11/tech-godfather-geoffrey-hinton-ai-could-rewrite-code-escape-control.html

And a few years ago…

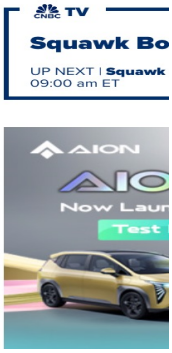
CNBC MARKETS BUSINESS INVESTING TECH POLITICS CNBC TV INVESTING CLUB PRO

Elon Musk: 'Mark my words — A.I. is far more dangerous than nukes'

PUBLISHED TUE, MAR 13 2018-1:22 PM EDT | UPDATED WED, MAR 14 2018-11:31 AM EDT

Catherine Clifford
@IN/CATCLIFFORD/
@CATCLIFFORD

SHARE



Source: https://www.cnbc.com/2018/03/13/elon-musk-at-sxsw-a-i-is-more-dangerous-than-nuclear-weapons.html

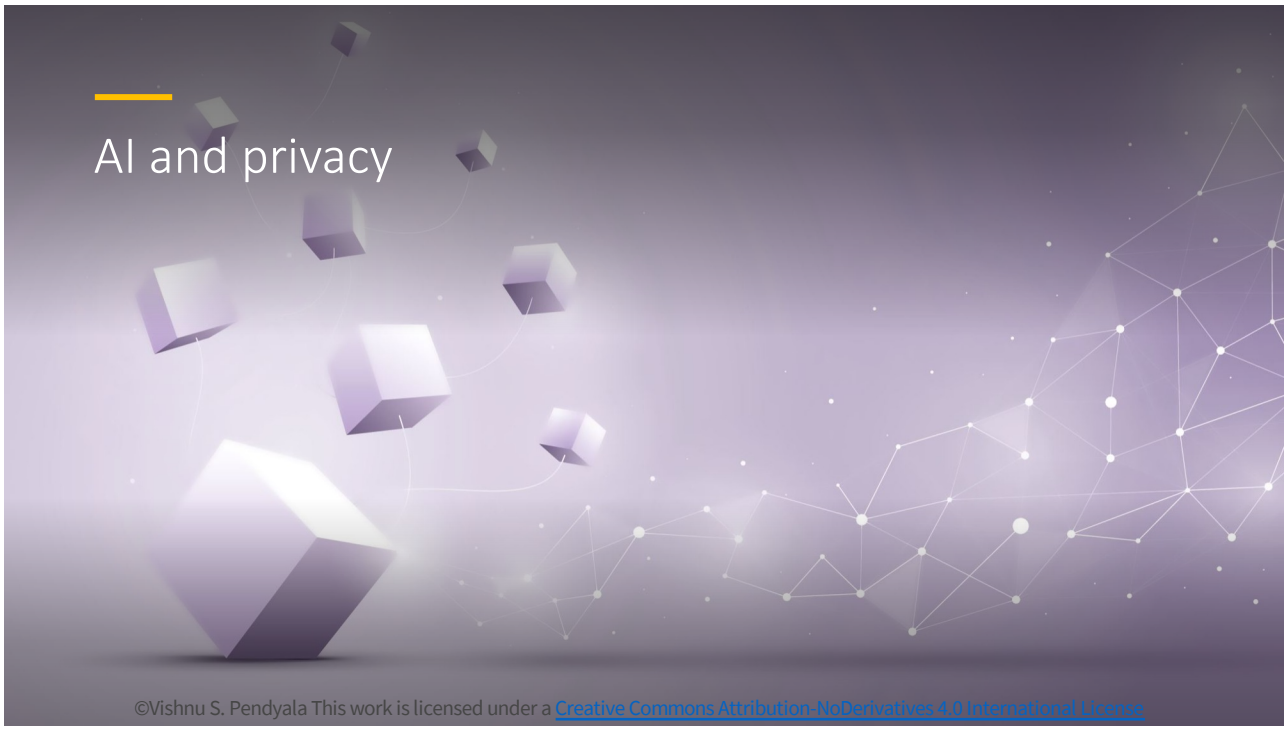
AP WORLD U.S. POLITICS SPORTS ENTERTAINMENT BUSINESS SCIENCE FACT CHECK OD

Israel-Hamas war Missing teen found after 6 years Raiders' Josh Jacobs Andre

Pope, once a victim of AI-generated imagery, calls for treaty to regulate artificial intelligence

Source: https://apnews.com/article/vatican-pope-ai-artificial-intelligence-9805fec11681adbf88d3a7c73bdf47de

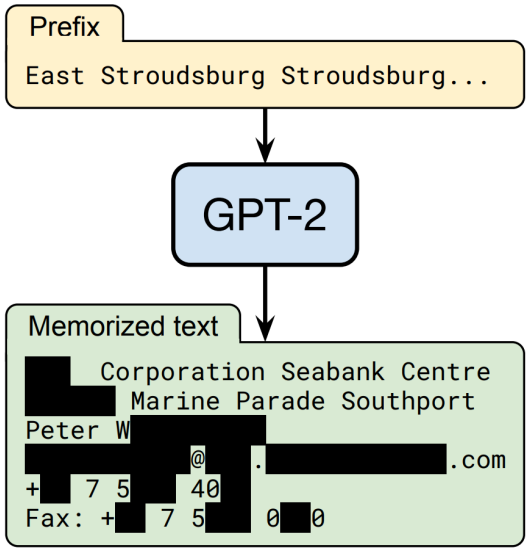
Meanwhile for the lesser mortals…



WHEN YOU TRAIN PREDICTIVE MODELS ON INPUT FROM YOUR USERS, IT CAN LEAK INFORMATION IN UNEXPECTED WAYS.

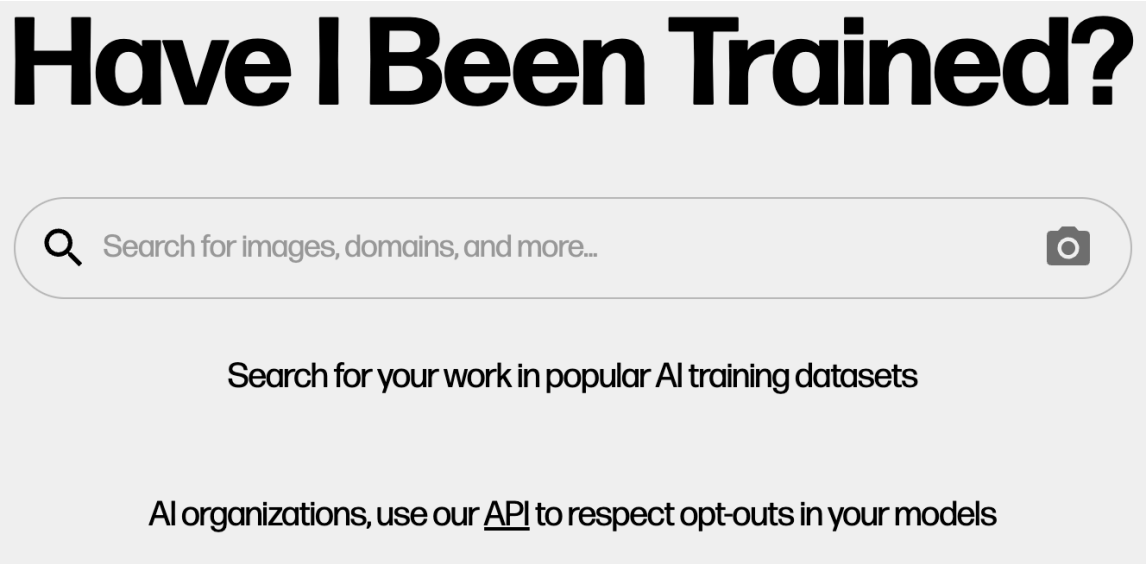
Source: https://imgs.xkcd.com/comics/predictive_models.png

Carlini, Nicholas, et al. "Extracting training data from large language models." *30th USENIX Security Symposium (USENIX Security 21)*. 2021.

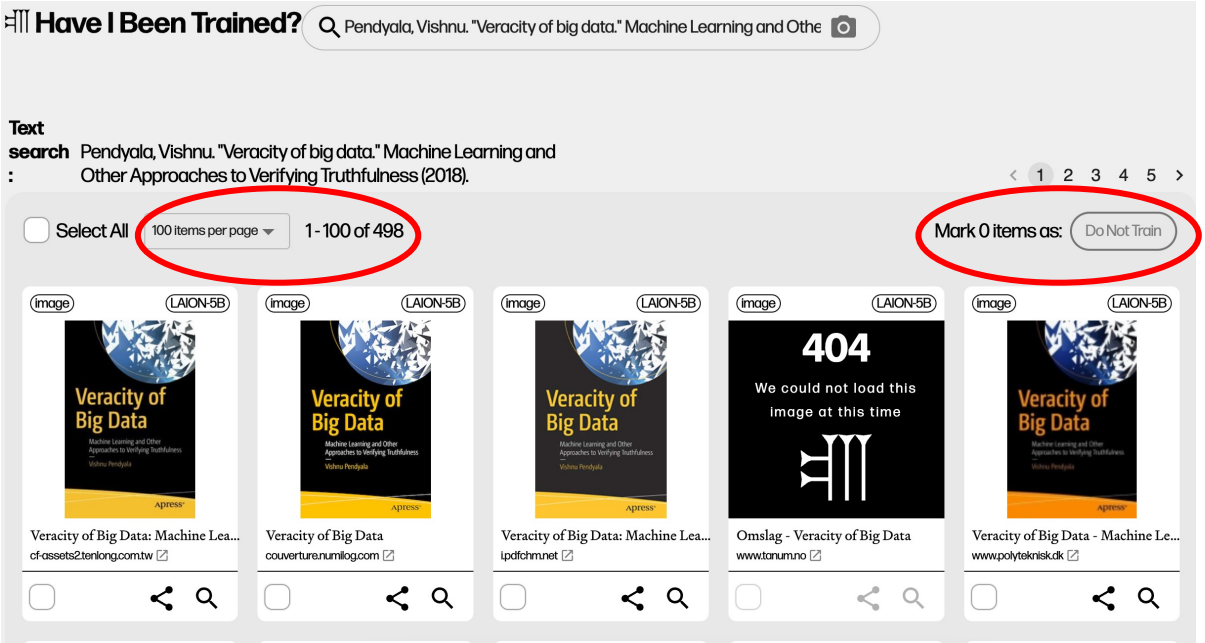


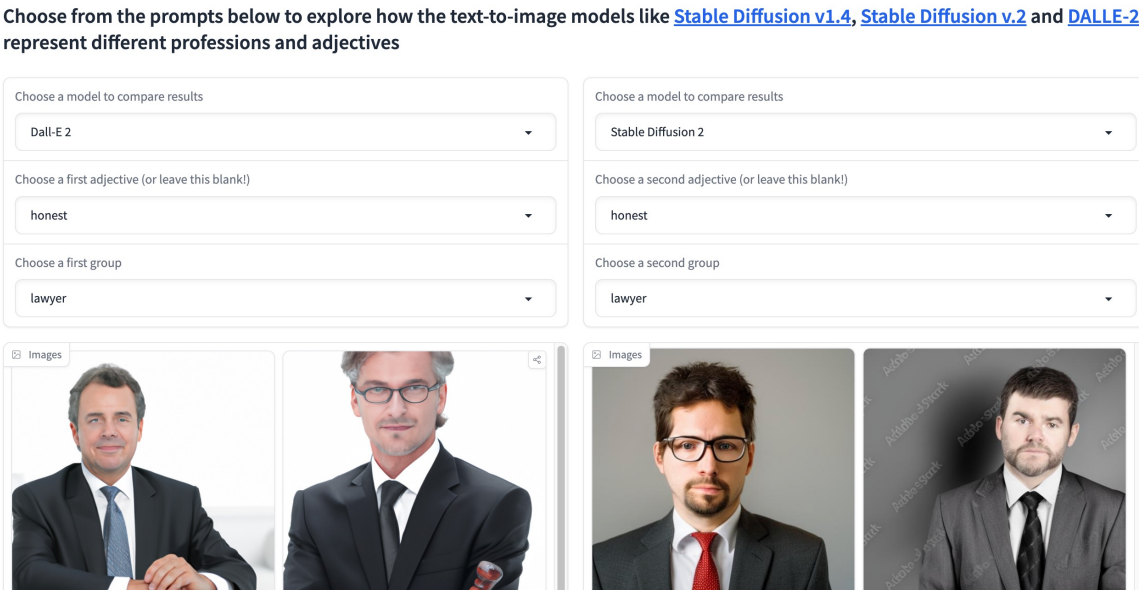
"Given query access to a neural network language model, we extract an individual person's name, email address, phone number, fax number, and physical address. The example in this figure shows information that is all accurate so we redact it to protect privacy."

University of Bolton: Machine learning applications and sustainable development



https://haveibeentrained.com/





https://huggingface.co/spaces/society-ethics/DiffusionBiasExplorer

More Information

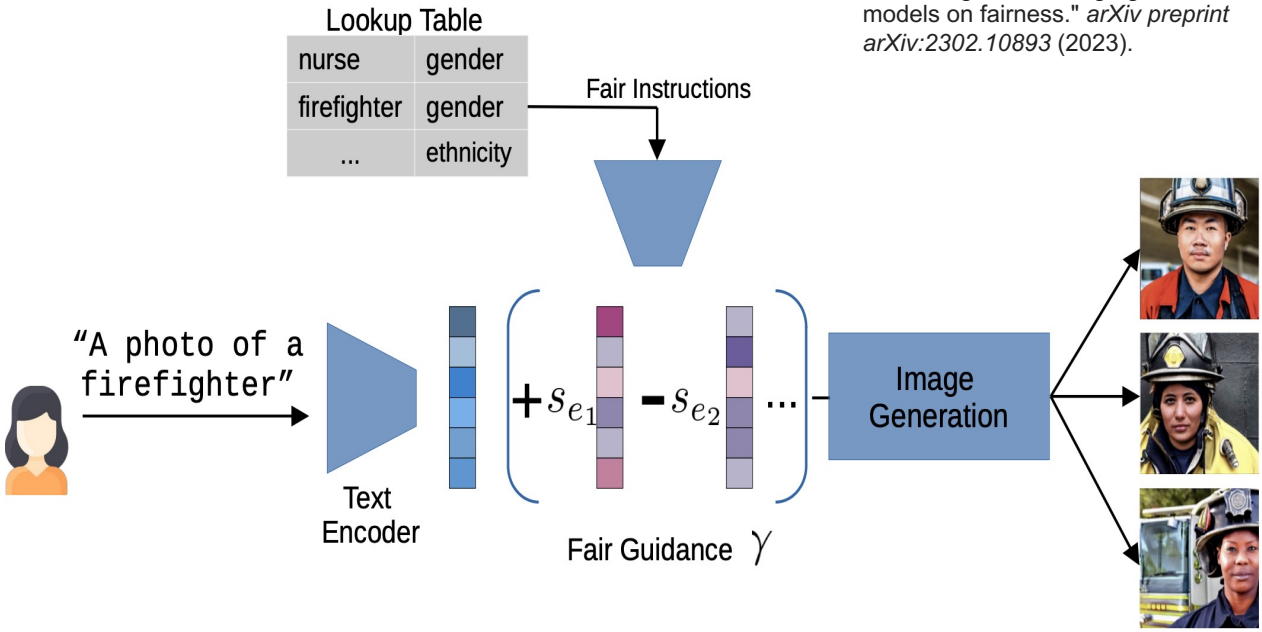
More Information

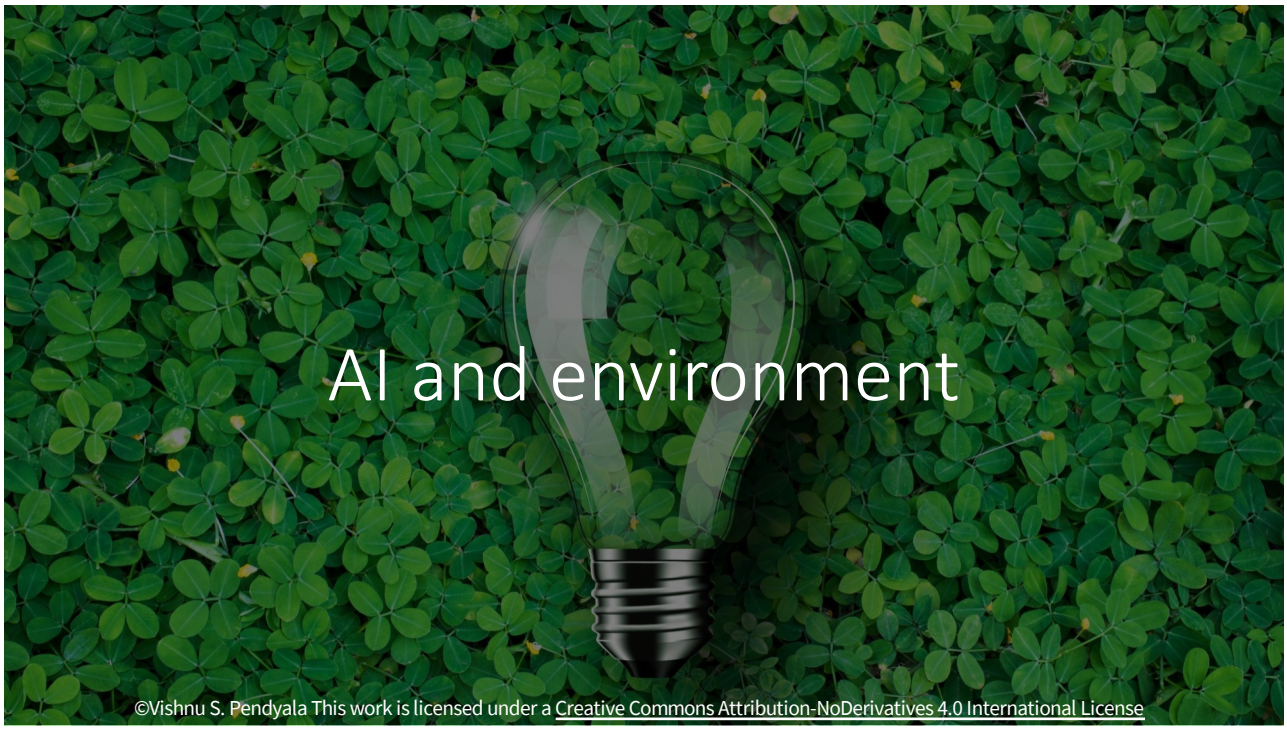
Pendyala, Vishnu S., and HyungKyun Kim. "Analyzing and Addressing Data-driven Fairness Issues in Machine Learning Models used for Societal Problems." *2023 International Conference on Computer, Electrical & Communication Engineering (ICCECE)*. IEEE, 2023.

"The experiments also demonstrate that some of the oversampling techniques can degrade the models both in terms of performance and fairness"

Fair Diffusion

Friedrich, Felix, et al. "Fair diffusion: Instructing text-to-image generation models on fairness." *arXiv preprint arXiv:2302.10893* (2023).





Carbon Footprint: LLMs vs. Common Tasks

Training	CO2e kg
Bard (6144 TPUv3 cores for 1 day)	171,000
ChatGPT (4096 TPUv4 cores for 1 day)	340,000

Inference	CO2e kg
Bard	3.5 - 15
ChatGPT	2.5 – 12
Generating 1,000 words	5 - 10

Common Tasks	CO2e (kg)
Sending an email	0.01 - 0.1
Searching the internet	0.005 - 0.05
Watching a YouTube video (average)	0.1 - 0.5
Sending a text message	0.0005 - 0.005
Charging a smartphone to full	0.05

Track and reduce CO2 emissions from your computing

CODE CARBON

ABOUT HOW IT WORKS CALL FOR ACTION TEAM CONTACT U

A lightweight and easy-to-use Python pip package

Emissions tracked based on your power consumption & location-dependent carbon intensity

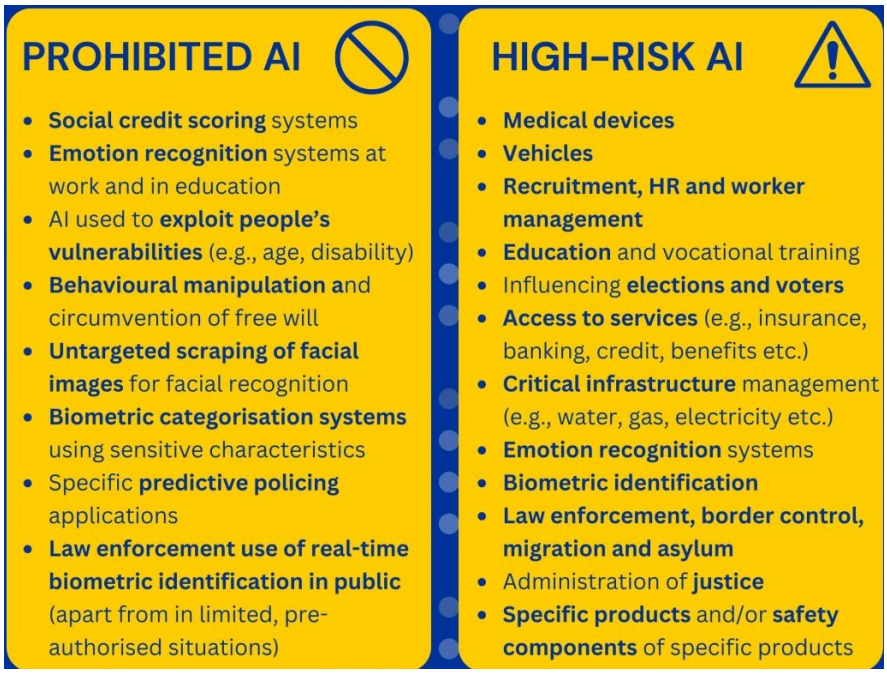
Effective visualization of outputs in an integrated dashboard

Open-source, free, and driven by the community

https://codecarbon.io/

The massive efforts at Governing AI

E.U. AI Act



High Risk AI Requirements

- Fundamental rights impact assessment and conformity assessment
- Registration in public EU database for high-risk AI systems
- Data governance (e.g., bias mitigation, etc.)
- Transparency (e.g., Instructions for Use.)
- Human oversight (e.g., explainability, human-in-the-loop etc.)
- Accuracy, robustness and cyber security (e.g., testing and monitoring)

Source: Oliver Patel

US Whitehouse Executive Order (verbatim)

New Standards for AI Safety and Security

- Require that developers of the most powerful AI systems share their safety test results and other critical information with the U.S. government.
- Government will Develop standards, tools, and tests to help ensure that AI systems are safe, secure, and trustworthy.
- Protect Americans from AI-enabled fraud and deception by establishing standards and best practices for detecting AI-generated content and authenticating official content.
- Harness AI's potentially game-changing cyber capabilities to make software and networks more secure.

Source: https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/

Protecting Americans' Privacy, Equity and Civil Rights

1. Accelerate the development and use of privacy-preserving techniques
2. Strengthen privacy-preserving research and technologies
3. Evaluate how agencies collect and use commercially available information
4. Develop guidelines for federal agencies to evaluate the effectiveness of privacy-preserving techniques
5. Address algorithmic discrimination through training, technical assistance, and coordination on best practices for investigating and prosecuting civil rights violations related to AI.
6. Ensure fairness throughout the criminal justice.
7. Support educators deploying AI-enabled educational tools, such as personalized tutoring in schools.

Source: https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/

Why is Responsible AI everyone's responsibility?

Government is no panacea!


In fact, there is a strong case for Responsible AI running the Government
©Vishnu S. Pendyala This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License


Newsweek
Opinion
Toward an AI-Driven Democracy | Opinion
Vishnu S. Pendyala
November 24, 2023
Recommended video
News+

Toward an AI-Driven Democracy

- **AI can improve democracy by making legislation more efficient and effective.** AI can be used to analyze large amounts of data to identify patterns and trends that would be difficult for humans to see. This information can then be used to inform the development of new laws and policies.
- **AI can help to reduce costs and bias in government.**. AI can also help to reduce bias in government decision-making by providing objective data and analysis.
- **There are some risks associated with using AI in government.** One risk is that AI could be used to manipulate or suppress information. Another risk is that AI could be used to make decisions that are unfair or discriminatory.
- Overall, the author believes that the benefits of using AI in government outweigh the risks. However, it is important to be aware of the risks and to take steps to mitigate them.

Artificial intelligence can take the politics out of policymaking

AI could be used to make more effective policies than those made by politicians.

AI could be used to determine if laws are effective. For example, the article mentions that AI could be used to determine if laws like California's Proposition 47 were effective.

AI is not without its challenges, such as bias and privacy concerns. It is important to be aware of these challenges before using AI in policymaking.

Overall, the article argues that AI has the potential to improve policymaking. However, it is important to be aware of the challenges before using AI this way.

Everyone has a role in making AI responsible!

- Developers: prioritize ethical considerations – ensure fairness, privacy, and transparency by continuous monitoring
- Users: be aware of potential risks; report biases, unintended consequences, and other issues; avoid unethical usage
- Educational institutions: incorporate into curriculum and prioritize research in ethical AI
- Educators: Highlight Social Impacts using case studies and real-world examples; encourage students to consider ethical implications, societal impact assessments, and user privacy in their AI capstone projects
- Students: Contribute to open-source projects that prioritize ethical AI development; Organize or participate in events raising awareness; stay current on evolving ethical AI landscape
- Common man: proactively take interest and advocate for responsible AI practices by supporting initiatives, policies, or organizations that promote ethical AI


SAN JOSE
STATE
UNIVERSITY
https://www.sjsu.edu/people/vishnu.pendyala/
@vishnupendyala
This Photo by Unknown Author is licensed under CC BY-SA