

BRIDGING THE AI DIGITAL DIVIDE **VIA STONG LEADERSHIP**



AGENDA / THESIS

In an age where artificial intelligence is reshaping our world, how do we ensure that progress aligns with the principles of equity, ethics, and social justice? Together, we will review AI's competing contingencies to determine what our institutions and students need to thrive in this shifting landscape.

View 1. Define b. Ethics 2. Determine

- 3. Deploy

Prologue: Values, Tech Philosophy, 30,000 Foot

- a. Equity
- c. Social Justice
- a. Solution
- b. Role of Admin



MY VALUES

- Trust and Transparency are necessary conditions for learning.
- Access makes learning possible; intervention makes it **real**.
- Access to knowledge is a human right.
- The CA Community College System **needs** my perspective, and it needs yours too.





MY TECH PHILOSOPHY

CURIOSITY

Can this technology effectively give us back our time? What is its potential for new positive discoveries?

HEALTHY SKEPTICISM

Does this tech actually deliver what it promises? Who benefits from our use of this technology (tech monopolies, commerce models, etc.)?





30,000 FT VIEU

SOCIETY

CULTURE

ART





HIGHER

EDUCATION

1. DEFINE

- Equity
- Ethics

1.



Social Justice



1A. EQUITY IN AI

Ensuring that the benefits of AI technologies are accessible to all, regardless of socioeconomic status, race, or gender

WORLD ECONOMIC FORUM REPORT, 2024

By 2027, 23% of all jobs will fundamentally change — primarily driven by advancements in artificial intelligence (AI) and automation. An estimated 69 million new roles are expected to be created and 83 million displaced. Communities already facing systemic inequalities are the ones most at risk of being left behind.



EQUITY IN AI







CSU UNVEILS MASSIVE VENTURE TO PROVIDE FREE AI TOOLS ACROSS ALL 23 CAMPUSES

- CSU is working to ensure that students in the nation's largest and most diverse public university system have equitable access to the technology
- Top tech companies will join academics on an advisory board to help identify AI skills needed in the California workforce.
- Industry partners will also provide internships and apprenticeships to give students real-world experience with AI on the job.

- Los Angeles Times, Feb. 4, 2025





18. ETHICS IN AI

ALGORITHM BIAS

Algorithmic bias occurs when systematic errors in machine learning algorithms produce unfair or discriminatory outcomes. It often reflects or reinforces existing socioeconomic, racial and gender biases.

Garbage in, garbage out.

DATA PRIVACY

The collection of use and personal data by AI developers and the significant risks that poses to the ownership of data.

ACADEMIC INTEGRITY

Al is not technically plagiarism; how can we ensure academic rigor?

ENVIRONMENTAL TOLL

COPYRIGHT

- The impact that AI has on the natural
- environment, such as increased energy
- consumption and water usage.

Can you copyright an Al output, US copyright office says no, and yes?





BIAS - GPT-3 DATA • ~570 GB total data





"Given both the competitive landscape and the safety implications of large-scale models like GPT-4, this report contains no further details about the architecture (including model size), hardware, training compute, dataset construction, training method, or similar." t, Open Al -GPT 4 Technical Repo



GPT-4 DATA



DATA PRIVACY

Educational market ranks the third-highest target for data hackers, trailing only the health and financial sectors.

> Emerging new privacy attacks on LLMs such as "Jailbreaking": In the context of Large Language Models (LLMs) like ChatGPT, users could potentially reverse engineer or "Jailbreak" the system to access information from previous conversations stored in its memory.

ENVIROMENTAL TOLL

 Data centers running Al applications now account for approximately 5-9% of global electricity demand and 2% of CO2 emissions
 Diverting electricity from green initiatives

AI AND COPYRIGHT

- - - feel right."
- - a. Can you copyright what AI makes?
 - - Rueters, 2023
 - 2025

1. The use of copyrighted work in AI models

a. Meta torrented at least 81.7 terabytes of data across multiple

i. In an internal message, an employee expressed concern about using Meta IP addresses "to load through torrents pirate content," stating, "torrenting from a corporate laptop doesn't

2. What are the real-world applications of citation?

i. Al-generated art cannot receive copyrights, US court s

ii. Al-generated art is public domain - Houston Law Review, Jan

iii. Generated art with human "creative arrangements or modifications" may be copyrighted.

AI AND ACADEMIC INTEGRITY

Al is not technically plagiarism; however, it can fall under the umbrella of Academic fraud/dishonestly. • "In this paper, we show both theoretically and empirically, that these state-of-the-art detectors cannot reliably detect LLM outputs in practical scenarios. Our experiments show that paraphrasing the LLM outputs helps evade these detectors effectively. Moreover, our theory demonstrates that for a sufficiently advanced language model, even the best detector can only perform marginally better than a random classifier." - "Can Al Generated text be Reliably

- Detected?"
- accuracy. -

OpenAI decommissions its AI Classifier due to poor

1C. SOCIAL JUSTICE

"While artificial intelligence (AI) systems have the potential to dramatically affect society, the people building AI systems are not representative of the people those systems are meant to serve. The AI workforce remains predominantly male and lacking in diversity in both academia and the industry, despite many years highlighting the disadvantages and risks this engenders. The lack of diversity in race and ethnicity, gender identity, and sexual orientation not only risks creating an uneven distribution of power in the workforce, but also, equally important, reinforces existing inequalities generated by AI systems, reduces the scope of individuals and organizations for whom these systems work, and contributes to unjust outcomes."

-Artificial Intelligence Index Report, Stanford University

$\Lambda \Lambda \Lambda \Lambda \Lambda \Lambda \Lambda$

ARTIFICIAL INTELLIGENCE INDEX REPORT, 2024

Gender of New CS graduates in the United States and Canada, 2022:

- Bachelors: 77.7% Male | 22.2% Female | 0.10% Nonbinary/Other
- Masters: 73.65% Male | 26.26% Female | 0.08% Nonbinary/Other

Source: CRA Taulbee Survey, 2023 | Chart: 2024 Al Index report

2022:

Ethnicity of New CS graduates in the United States and Canada,

• Bachelors: 44.58%, White | 35.74%, Asian | 11%, Hispanic (any race) | 4.36%, Multiracial (not Hispanic) | 4.08%, Black or African American | 0.13%, American Indian or Alaska Native 0.11%, Native Hawaiian or Pacific Islander

• Masters: 47.87%, White | 35.76%, Asian | 8.19%, Hispanic (any race) | 4.22%, Black or African American | 3.48%, Multiracial (not Hispanic) | 0.35%, American Indian or Alaska Native | 0.13%, Native Hawaiian or Pacific Islander

MILLIONS AFFECTED BY RACIAL BIAS IN HEALTHCARE ALGORITHM

An algorithm used in U.S. hospitals was supposed to direct patients to healthcare programs, but a 2019 study found it was less likely to refer black patients to the care they need than equally sick white patients. Because the algorithm assigned people to high-risk categories on the basis of costs (care provided to black people costs an average of \$1,800 less per year than the care given to white people with the same number of chronic health problems), those biases were passed on in its results: black people had to be sicker than white people before being referred for additional help. Only **17.7%** of patients that the algorithm assigned to receive extra care were black. The researchers calculate that the proportion would have been 46.5% if the algorithm was unbiased.

EQUITY

ETHICS

2. DETERMINE - SOLUTION = AI LITERACY

Al literacy is the ability to technically understand, use, and evaluate artificial intelligence (Al) systems and their outputs.

1.How these systems work

What they are and are not capable of
2.The biases and ethical issues inherent in the structure and use of these systems
3.How to use with these systems effectively and efficiently

• Prompt engineering

2. DETERMINE - SOLUTION

- **UNESCO** AI Competency Frameworks Crucial \bullet for the Future Economic Success of Students
- EU AI Act's AI literacy mandate spurs 866% surge in enterprise upskilling
- **IBM** AI Literacy: Closing the artificial intelligence skills gap
- Organization for Economic Cooperation and **Development (OECD)** - Al and the Future of <u>Skills</u>
- MIT Al Literacy Raise Initiative

BO,OO FT VIEU

CULTURE

SOCIETY

— Systems

BRT BO,DOG FT VIEW

2. DETERMINE -ROLE OF ADMIN

We've defined the problem and identified the solution. One of the final steps is to determine the role of administrators in this context.

Provide strong high quality leadership in the deployment of Al literacy initiatives that will appropriately prepare our institutions and students to thrive in this shifting landscape.

3. DEPLOY - AI LITERACY

- 7. Professional development / AI literacy for administrators
- 2. Professional development / literacy initiatives for faculty and staff 3. Institutionalized AI literacy program
- for students
 - a. Couched in already existing information literacy courses.

AI LITERACY INITIATIVE FOR ADMINISTRATORS

bit.ly/Al4CCCAdmins

References

- D. Zhang, S. Mishra, E. Brynjolfsson, J. Etchemendy, D. Ganguli, B., Grosz, T. Lyons, J. Manyika, J. Carlos Niebles, M. Sellitto, Y. Shoham, J. Clark, and R. Perrault, "The AI Index 2021 Annual Report," AI Index Steering Committee, Human-Centered AI Institute, Stanford University, Stanford, CA, March 2021.
- A. Goldstein, J., Sastry, G., Musser, M., DiResta, R., Gentzel, M., & Sedova, K. (2023). Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations. Georgetown University's Center for Security and Emerging Technology, OpenAI, Stanford Internet Observatory. chromeextension://efaidnbmnnnibpcajpcglclefindmkaj/https://cdn.openai.com/papers/forecasting-misuse.pdf
- Angwin (ProPublica), J. (2015, March 19). Practitioners-Guide-to-COMPAS-Core. Www.documentcloud.org. https://www.documentcloud.org/documents/2840784-Practitioner-s-Guide-to-COMPAS-Core.html#document/p30/a296482
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine Bias. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminalsentencing
- Bakker, M. A., Chadwick, M. J., Sheahan, H. R., Tessler, M. H., Campbell-Gillingham, L., Balaguer, J., McAleese, N., Glaese, A., Aslanides, J., Botvinick, M. M., & Summerfield, C. (2022). Fine-tuning language models to find agreement among humans with diverse preferences. ArXiv:2211.15006 [Cs], 1. https://arxiv.org/abs/2211.15006
- Douglas, L. (2017, December 5). Al is not just learning our biases; it is amplifying them. Medium; Medium. https://medium.com/@laurahelendouglas/ai-is-not-justlearning-our-biases-it-is-amplifying-them-4d0dee75931d
- Feathers, T. (2019, August 20). Flawed Algorithms Are Grading Millions of Students' Essays. Www.vice.com. https://www.vice.com/en/article/pa7dj9/flawedalgorithms-are-grading-millions-of-students-essays
- Glaser, A. (2017, October 24). Who Trained Your A.I.? Slate. https://slate.com/technology/2017/10/what-happens-when-the-data-used-to-train-a-i-is-biased-andold.html
- Hao, K. (2019, April 17). Al's white guy problem isn't going away. MIT Technology Review. https://www.technologyreview.com/2019/04/17/136072/ais-white-guyproblem-isnt-going-away/
- Isbell, A. H. and C. (2020, September 21). Diversity in AI: The Invisible Men and Women. MIT Sloan Management Review. https://sloanreview.mit.edu/article/diversity-inai-the-invisible-men-and-women/
- Kamath, B. (2022, October 28). Generative AI Is Biased. But Researchers Are Trying to Fix It. Analytics India Magazine. https://analyticsindiamag.com/generative-ai-isbiased-but-researchers-are-trying-to-fix-it/
- Larson, J., Mattu, S., Kirchner, L., & Angwin, J. (2016, May 23). How We Analyzed the COMPAS Recidivism Algorithm. ProPublica; ProPublica. https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm
- Ryder, C. (2017, October 18). Professor Brain Discusses Machine Learning and Bias: The Story of the Enron Email Archive | NYU Tandon School of Engineering. Www.nyu.engineering. http://www.nyu.engineering/news/professor-brain-discusses-machine-learning-and-bias-story-enron-email-archive
- SIMONITE, T. (2020, July 10). Meet the Secret Algorithm That's Keeping Students Out of College. Wired. https://www.wired.com/story/algorithm-set-students-grades-

References Cont.

- Spielkamp, M. (2017, June 15). Inspecting Algorithms for Bias. MIT Technology Review. https://www.technologyreview.com/2017/06/12/105804/inspectingalgorithms-for-bias/
- David, E. (2023, August 29). US Copyright Office wants to hear what people think about AI and copyright. The Verge. https://www.theverge.com/2023/8/29/23851126/us-copyright-office-ai-public-comments
- Dastin, J. (2018, October 11). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. https://www.reuters.com/article/us-amazon-comjobs-automation-insight-idUSKCN1MK08G
- Bashir, N., Donti, P., Cuff, J., Sroka, S., Ilic, M., Sze, V., Delimitrou, C., & Olivetti, E. (2024). The Climate and Sustainability Implications of Generative AI. An MIT Exploration of Generative AI. https://doi.org/10.21428/e4baedd9.9070dfe7
- BloombergNEF. "New Energy Outlook 2022 | Bloomberg NEF." Bloomberg NEF, 2022, about.bnef.com/new-energy-outlook/.
- Orland, Kyle. "Taking a Closer Look at AI's Supposed Energy Apocalypse." Ars Technica, 25 June 2024, arstechnica.com/ai/2024/06/is-generative-ai-really-going-towreak-havoc-on-the-power-grid/.
- Ledford H. (2019). Millions of black people affected by racial bias in health-care algorithms. Nature, 574(7780), 608–609. https://doi.org/10.1038/d41586-019-03228-6
- Obermeyer, Z et al. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science 366, 447-453. DOI: 1