Artificial Intelligence and Social Justice: Critical Analysis of Current Developments

ainews.social - Generated Analysis

August 31, 2025

1 Analysis of Purpose and Intent

The explicit purpose of many authors discussing AI in social justice is to highlight the potential of AI to both advance and hinder equity and inclusion. For instance, some authors argue that AI can democratize access to education, thereby leveling the playing field for marginalized groups [13,18,19]. They justify this by pointing to initiatives like OpenAI's Learning Accelerator, which aims to provide educational resources to underrepresented communities [19]. However, the realization of these goals is contingent upon responsible implementation, which includes addressing biases inherent in AI systems [28,30,59]. The authors emphasize the need for transparency and accountability in AI development to ensure these technologies do not perpetuate existing inequalities [9,28,59].

Contrastingly, other authors express concerns about AI's potential to exacerbate social injustices, particularly through biased algorithms that can reinforce stereotypes and discrimination [5,30,69]. For example, AI hiring tools have been shown to disadvantage autistic job applicants, raising questions about the fairness and inclusivity of such technologies [5]. These authors argue that without rigorous bias mitigation strategies, AI systems may inadvertently harm the very groups they are intended to support [30,59]. This perspective reflects the needs of stakeholders who are wary of AI's impact on vulnerable populations and demand more ethical oversight [5,28,59].

The purposes expressed by these authors are realistic but face significant challenges. Achieving equitable AI deployment requires overcoming technical hurdles, such as developing bias-adjusted algorithms, and institutional challenges, such as implementing comprehensive regulatory frameworks [28,38,54]. Moreover, the diverse needs of stakeholders, including educators, policymakers, and marginalized communities, must be balanced to ensure that AI serves as a tool for social progress rather than a source of division [33,35,62]. The complexity of these challenges underscores the necessity for ongoing dialogue and collaboration among all parties involved [33,35,62].

2 Critical Questions and Inquiries

In the realm of AI and social justice, critical questions revolve around the dual potential of AI technologies to either advance or hinder equity and inclusion. One core problem being addressed is the inherent bias in AI systems, which can perpetuate existing social inequalities if not properly managed [5,30,69]. This issue is particularly pressing in areas like employment, where AI hiring tools have been shown to disadvantage marginalized groups, such as autistic job applicants, highlighting the need for rigorous bias mitigation strategies [5,30,59]. These inquiries build upon existing research by emphasizing the necessity for transparency and accountability in AI development to prevent the reinforcement of stereotypes and discrimination [9,28,59].

Methodological approaches proposed to tackle these challenges include the development of bias-adjusted algorithms and comprehensive regulatory frameworks that ensure ethical oversight [28,38,54]. Researchers suggest that integrating behavioral economics into AI design could lead to more human-like decision-making processes, potentially reducing bias [38]. These approaches reflect current challenges in balancing technological advancement with ethical considerations, as stakeholders demand more responsible AI deployment [33,35,62]. The complexity of these challenges underscores the importance of interdisciplinary collaboration

among technologists, ethicists, and policymakers to create AI systems that genuinely promote social progress [33,35,62].

The assumptions underlying these inquiries often include the belief that AI, if responsibly implemented, can democratize access to resources and opportunities, thereby leveling the playing field for marginalized communities [13,18,19]. However, this optimistic view is tempered by the recognition that achieving such outcomes requires overcoming significant technical and institutional hurdles [28,38,54]. For instance, the assumption that AI can be a neutral tool for social good is challenged by evidence of its potential to exacerbate inequalities, necessitating a more nuanced understanding of AI's role in society [5,30,69]. These critical questions and inquiries not only reflect the current state of research but also push the boundaries of how AI can be harnessed for equitable social transformation.

3 Core Assumptions and Premises

In examining the core assumptions about AI's role in social justice, a foundational belief is that AI possesses the potential to democratize access to resources and opportunities, thereby promoting equity and inclusion [13,18,19]. This assumption is grounded in initiatives like OpenAI's Learning Accelerator, which aims to provide educational resources to underrepresented communities, suggesting that AI can level the playing field if implemented responsibly [19]. However, this optimistic view is tempered by the recognition of significant technical and institutional hurdles that must be overcome to achieve such outcomes [28,38,54]. The belief that AI can be a neutral tool for social good is challenged by evidence of its potential to exacerbate inequalities, necessitating a more nuanced understanding of AI's role in society [5,30,69].

These assumptions shape analysis approaches by emphasizing the need for transparency and accountability in AI development to prevent the reinforcement of stereotypes and discrimination [9,28,59]. Researchers propose methodological approaches such as the development of bias-adjusted algorithms and comprehensive regulatory frameworks to ensure ethical oversight [28,38,54]. The integration of behavioral economics into AI design is suggested as a means to achieve more human-like decision-making processes, potentially reducing bias [38]. These approaches reflect the current challenges in balancing technological advancement with ethical considerations, as stakeholders demand more responsible AI deployment [33,35,62].

Evidence supporting these assumptions includes the potential for AI to democratize education and other resources, as seen in various educational initiatives [13,18,19]. However, contrasting evidence highlights the risks of biased algorithms, particularly in employment contexts where AI hiring tools have disadvantaged marginalized groups such as autistic job applicants [5,30,59]. This duality underscores the complexity of AI's impact on social justice, where the same technology can both advance and hinder equity depending on its implementation and oversight [5,30,69].

Assumptions about AI's role in social justice vary across perspectives, with some stakeholders viewing AI as a transformative tool for social progress, while others express caution about its potential to perpetuate existing inequalities [33,35,62]. These differing viewpoints highlight potential biases in how AI is perceived and implemented, emphasizing the need for interdisciplinary collaboration to address these challenges [33,35,62]. By recognizing these biases, stakeholders can work towards creating AI systems that genuinely promote social progress and do not serve as sources of division [33,35,62].

4 Key Concepts and Theoretical Framework

In the exploration of AI within the context of social justice, several key concepts and theoretical frameworks emerge, each contributing to a nuanced understanding of AI's potential impacts. Central to these discussions is the concept of algorithmic bias, which underscores the risk of AI systems perpetuating existing social inequalities if not properly managed [5,30,69]. This concept is foundational in understanding how AI technologies can either advance or hinder equity and inclusion, particularly in sensitive areas such as employment and education [5,30,59]. Over time, the development of bias-adjusted algorithms and comprehensive regulatory frameworks has been proposed as a means to mitigate these risks, reflecting a growing recognition of the need for ethical oversight in AI deployment [28,38,54].

Another pivotal concept is the democratization of access to resources and opportunities through AI, which posits that AI can level the playing field for marginalized communities if implemented responsibly

[13,18,19]. This idea is exemplified by initiatives like OpenAI's Learning Accelerator, which aims to provide educational resources to underrepresented groups, suggesting that AI can serve as a transformative tool for social progress [19]. However, this optimistic view is tempered by evidence of AI's potential to exacerbate inequalities, necessitating a more nuanced understanding of AI's role in society [5,30,69]. The interplay between these concepts highlights the dual potential of AI technologies, where the same tools can both democratize and discriminate, depending on their design and implementation [5,30,69].

Theoretical frameworks addressing AI in social justice often integrate interdisciplinary perspectives, combining insights from technology, ethics, and policy to address complex challenges [33,35,62]. These frameworks emphasize the importance of transparency and accountability in AI development, advocating for the integration of behavioral economics to achieve more human-like decision-making processes [9,28,38]. By incorporating diverse disciplinary insights, these frameworks aim to balance technological advancement with ethical considerations, ensuring that AI systems genuinely promote social progress rather than reinforce existing disparities [33,35,62].

Despite their strengths, these frameworks face limitations, particularly in addressing the rapid pace of AI development and the evolving nature of ethical and legal standards [4,36,62]. The dynamic landscape of AI technologies necessitates continuous adaptation and refinement of theoretical models to remain relevant and effective [4,36,62]. Furthermore, the assumption that AI can be a neutral tool for social good is challenged by evidence of its potential to perpetuate stereotypes and discrimination, highlighting the need for ongoing critical evaluation and revision of existing frameworks [5,30,69].

In practice, these theoretical frameworks reflect the complexities of implementing AI in a socially just manner. They underscore the necessity for interdisciplinary collaboration among technologists, ethicists, and policymakers to create AI systems that align with ethical principles and promote equitable outcomes [33,35,62]. By advancing these frameworks, stakeholders can work towards harnessing AI's potential for social transformation while mitigating its risks, ultimately contributing to a more just and inclusive society [33,35,62].

5 Implications and Future Directions

The implications of AI in social justice are multifaceted, with potential for both significant advancements and challenges. A major claim is that AI can drive concrete changes in democratizing access to education and resources, as evidenced by initiatives like OpenAI's Learning Accelerator, which aims to provide educational resources to underrepresented communities [13,18,19]. This initiative exemplifies how AI can serve as a transformative tool for social progress, potentially leveling the playing field for marginalized groups [19]. However, the realization of these benefits is contingent upon addressing the inherent biases in AI systems, which, if left unchecked, can exacerbate existing inequalities [5,30,69].

Evidence supporting the democratization potential of AI includes its role in reshaping educational landscapes, where AI tools are increasingly integrated into learning environments to enhance accessibility and personalization [12,18,19]. Nonetheless, contrasting evidence highlights the risks associated with AI's implementation, particularly in employment contexts where AI hiring tools have been shown to disadvantage marginalized groups, such as autistic job applicants [5,30,59]. This duality underscores the complexity of AI's impact on social justice, where the same technology can both advance and hinder equity depending on its design and oversight [5,30,69].

Stakeholders differ in their perspectives on AI's role in social justice, with some viewing it as a catalyst for social progress, while others express caution about its potential to perpetuate existing inequalities [33,35,62]. These differing viewpoints highlight potential biases in how AI is perceived and implemented, emphasizing the need for interdisciplinary collaboration to address these challenges [33,35,62]. By recognizing these biases, stakeholders can work towards creating AI systems that genuinely promote social progress and do not serve as sources of division [33,35,62].

The mechanisms driving change in AI's role in social justice include the development of bias-adjusted algorithms and comprehensive regulatory frameworks to ensure ethical oversight [28,38,54]. These mechanisms are crucial in mitigating the risks of algorithmic bias and ensuring that AI systems are transparent and accountable [9,28,59]. The integration of behavioral economics into AI design is suggested as a means to achieve more human-like decision-making processes, potentially reducing bias and enhancing fairness

[38]. These approaches reflect the current challenges in balancing technological advancement with ethical considerations, as stakeholders demand more responsible AI deployment [33,35,62].

Implementation challenges persist, particularly in addressing the rapid pace of AI development and the evolving nature of ethical and legal standards [4,36,62]. The dynamic landscape of AI technologies necessitates continuous adaptation and refinement of theoretical models to remain relevant and effective [4,36,62]. Furthermore, the assumption that AI can be a neutral tool for social good is challenged by evidence of its potential to perpetuate stereotypes and discrimination, highlighting the need for ongoing critical evaluation and revision of existing frameworks [5,30,69]. By advancing these frameworks, stakeholders can work towards harnessing AI's potential for social transformation while mitigating its risks, ultimately contributing to a more just and inclusive society [33,35,62].

6 Interpretative Analysis and Synthesis

In the Interpretative Analysis and Synthesis section, we delve deeper into the nuanced interpretations of AI's role in social justice, synthesizing evidence from diverse sources to uncover underlying reasoning patterns and methodological considerations. A recurring theme is the dual potential of AI technologies to either democratize access to resources or exacerbate existing inequalities, contingent upon their design and implementation [5,30,69]. This duality is evident in the educational sector, where AI initiatives like OpenAI's Learning Accelerator aim to provide equitable learning opportunities, yet face challenges in addressing algorithmic biases that could undermine these efforts [19,30,59].

The reasoning patterns emerging from the literature highlight a critical tension between the optimistic view of AI as a tool for social progress and the cautionary perspective that emphasizes its potential to perpetuate systemic biases [33,35,62]. This tension is particularly pronounced in employment contexts, where AI hiring tools have been shown to disadvantage marginalized groups, such as autistic job applicants, thereby reinforcing existing disparities [5,30,59]. The evidence suggests that while AI can facilitate more efficient decision-making processes, it requires robust bias mitigation strategies to ensure fairness and inclusivity [28,38,54].

Alternative views on AI's role in social justice often stem from differing assumptions about the neutrality of technology. Some scholars argue that AI systems, if designed with transparency and accountability, can be neutral tools for social good [9,28,38]. However, others contend that AI inherently reflects the biases of its creators and the data it is trained on, necessitating continuous critical evaluation and revision of existing frameworks to prevent the entrenchment of stereotypes and discrimination [5,30,69]. This debate underscores the importance of interdisciplinary collaboration in developing AI systems that align with ethical principles and promote equitable outcomes [33,35,62].

Uncertainties in the field are primarily related to the rapid pace of AI development and the evolving nature of ethical and legal standards [4,36,62]. These uncertainties pose significant challenges for stakeholders attempting to balance technological advancement with ethical considerations. The dynamic landscape of AI technologies necessitates continuous adaptation and refinement of theoretical models to remain relevant and effective [4,36,62]. Moreover, the assumption that AI can be a neutral tool for social good is increasingly challenged by evidence of its potential to perpetuate stereotypes and discrimination, highlighting the need for ongoing critical evaluation and revision of existing frameworks [5,30,69].

Methodological issues arise in the evaluation of AI's impact on social justice, particularly in measuring the effectiveness of bias mitigation strategies and the transparency of AI systems [28,38,54]. The integration of behavioral economics into AI design is suggested as a means to achieve more human-like decision-making processes, potentially reducing bias and enhancing fairness [38]. However, the effectiveness of these approaches remains an area of active research, with stakeholders demanding more responsible AI deployment to ensure that AI systems genuinely promote social progress rather than reinforce existing disparities [33,35,62]. By advancing these frameworks, stakeholders can work towards harnessing AI's potential for social transformation while mitigating its risks, ultimately contributing to a more just and inclusive society [33,35,62].

7 References

1. AI Ethics (2025).
   https://www.currentaffairs.org/news/ai-ethics-discourse-ignores-its-deadliest-use-war

2. AI Ethics and Regulatory Risk (2025).
   https://www.ainvest.com/news/ai-ethics-regulatory-risk-legal-reputational-crises-reshape-tech-sector-valuations-2508

3. AI Ethics and Regulatory Risk (2025).
   https://www.ainvest.com/news/ai-ethics-regulatory-risk-legal-reputational-crises-reshape-tech-sector-valuations-2508/

4. AI Ethics and Regulatory Risk: How Legal and Reputational Crises Reshape Tech Sector Valuations (2025).
   https://www.ainvest.com/news/ai-ethics-regulatory-risk-legal-reputational-crises-reshape-tech-sector-valuations-2508

5. AI Hiring Tools Elevate Bias Danger for Autistic Job Applicants (2025).
   https://news.bloomberglaw.com/daily-labor-report/ai-hiring-tools-elevate-bias-danger-for-autistic-job-applicants

6. AI Literacy and Critical Thinking (2025).
   https://libguides.macalester.edu/ailiteracy

7. AI Policy Considerations (2025).
   https://library.buffalostate.edu/c.php?g=1372906&p=10459576

8. AI Transparency (2025).
   https://www.plainconcepts.com/es/ai-transparency/

9. AI Transparency: Pilar fundamental para una IA etica y segura (2025).
   https://www.plainconcepts.com/es/ai-transparency/

10. AI as a Child in a Cage (2025).
    https://philpapers.org/rec/SAFAAA-4

11. AI as a Child in a Cage: On Mirrors, Obedience, and the Illusion of Intelligence (2025).
    https://philpapers.org/rec/SAFAAA-4

12. AI changes education landscape (2025).
    https://komonews.com/news/back-to-school/ai-changes-education-landscape

13. AI could democratise higher education if implemented responsibly (2025).
    https://dig.watch/updates/ai-could-democratise-higher-education-if-implemented-responsibly

14. AI for Learning, Design, and Technology (2025).
    https://www.cmich.edu/program/AI-for-Learning-Design-and-Technology-(Graduate-Certificate

15. AI fuels demand for continuing education in the workplace (2025).
    https://valorinternational.globo.com/business/news/2025/08/26/ai-fuels-demand-for-continuing-education-in-the-workplace.ghtml

16. AI in Classrooms Needs Balance (2025).
https://www.bweducation.com/article/ai-in-classrooms-needs-balance-569496

17. AI in Education (2025).
https://www.editorji.com/world-news/ai-in-education-innovation-partner-1756189371100

18. AI in Education: From Task Tool to Innovation Partner (2025).
https://www.editorji.com/world-news/ai-in-education-innovation-partner-1756189371100

19. AI in Education: OpenAI Launches Learning Accelerator with IIT Madras (2025).
https://www.deccanherald.com/india/openai-unveils-learning-accelerator-india-first-initiative-partners-with-iit-madras-others-3696176

20. AI in Healthcare (2025).
https://www.fiercehealthcare.com/health-tech/how-nvidia-backed-artisight-taking-smart-hospitals-next-level-ai-and-computer-vision

21. AI in Healthcare & Cybersecurity (2025).
https://nwnews.com/ai-healthcare-cybersecurity-spokane-falls-program/

22. AI in Healthcare & Cybersecurity: Spokane Falls Community College Bets Big on the Future of Tech (2025).
https://nwnews.com/ai-healthcare-cybersecurity-spokane-falls-program/

23. AI really is cutting out entry-level jobs for human workers, study claims (2025).
https://www.techradar.com/pro/ai-really-is-cutting-out-entry-level-jobs-for-human-workers-study-claims

24. AI's future can't be 'Made in America' alone (2025).
https://www.pwmnet.com/content/d33fe15d-e504-4a09-a2e1-491ea12591d7

25. AI, Wealth Inequality, and the Future of Retail Investing (2025).
https://www.ainvest.com/news/ai-wealth-inequality-future-retail-investing-algorithmic-capital-shifts-redefining-financial-inclusion-2508

26. AI, Wealth Inequality, and the Future of Retail Investing (2025).
https://www.ainvest.com/news/ai-wealth-inequality-future-retail-investing-algorithmic-capital-shifts-redefining-financial-inclusion-2508/

27. AI-driven feedback for improving teamwork (2025).
https://www.cambridge.org/core/journals/proceedings-of-the-design-society/article/aidriven-feedback-for-improving-teamwork-and-learning-in-collaborative-engineering-design/D85A46967E8501700ADDB030EDBEA722

28. Addressing Compliance, Bias, and Transparency in AI Systems (2025).
https://link.springer.com/chapter/10.1007/978-3-031-92916-8_15

29. Algorithmic Bias in Early Childhood AI Tools (2025).
https://www.researchgate.net/profile/Jordan-Nelson-15/publication/394826217_Algorithmic_Bias_in_Early_Childhood_AI_Tools_Implications_for_Equity_and_Inclusion/links/68a7bd157984e374acea7a86/Algorithmic-Bias-in-Early-Childhood-AI-Tools-Implications-for-Equity-and-Inclusion.pdf

30. Algorithmic Bias in Early Childhood AI Tools: Implications for Equity and Inclusion (2025). https://www.researchgate.net/profile/Jordan-Nelson-15/publication/394826217_Algorithmic_Bias_in_Early_Childhood_AI_Tools_Implications_for_Equity_and_Inclusion/links/68a7bd157984e374acea7a86/Algorithmic-Bias-in-Early-Childhood-AI-Tools-Implications-for-Equity-and-Inclusion.pdf

31. Artificial Intelligence and the Future of Teaching and Learning (2025). https://digital.library.unt.edu/ark:/67531/metadc2114121/

32. Artificial Intelligence as a Factor of Social Progress (2025).

33. Artificial Intelligence as a Factor of Social Progress: Opportunities and Challenges (2025).

34. Artificial Intelligence in Social Sciences (2025). https://academia.edu.pk/index.php/AIRJ/article/view/671

35. Artificial Intelligence in Social Sciences: Ethics, Economy, and Governance (2025). https://academia.edu.pk/index.php/AIRJ/article/view/671

36. Artificial Intelligence, Ethics, and the Legal Responsibility for Algorithmic Decisions (2025). http://www.rc-archive.com/index.php/Journal/article/view/246

37. Beyond Information Literacy: Exploring AI's Impact on Labor in Academic Libraries (2025). https://muse.jhu.edu/pub/1/article/968498/summary

38. Bias-Adjusted LLM Agents for Human-Like Decision-Making via Behavioral Economics (2025). https://arxiv.org/abs/2508.18600

39. Digital Transformation in Education (2025). https://journals.ncert.gov.in/IJET/article/view/1400

40. Ethical Challenges in Financial Data Science (2025). https://journals.ai-mrc.com/jmri/article/view/566

41. Faculty Resources - Northwestern AI in Education (2025). https://library.ctstate.edu/c.php?g=1451539&p=10789604

42. Generative AI in Academia (2025). https://journals.ncert.gov.in/IJET/article/download/1406/369

43. Harnessing Generative AI for Assessment Item Development (2025). https://onlinelibrary.wiley.com/doi/abs/10.1111/ijsa.70021

44. Impacts of ChatGPT on Critical Thinking (2025). https://www.jet.or.id/index.php/jet/article/view/568

45. Indautor y SCJN (2025). https://www.elfinanciero.com.mx/nacional/2025/08/30/indautor-y-scjn-rechazan-registro-de-obras-generadas-con-ia-la-ley-exige-que-sean-de-creacion-humana/

46. Indautor y SCJN rechazan registro de obras generadas con IA (2025). https://www.elfinanciero.com.mx/nacional/2025/08/30/indautor-y-scjn-rechazan-registro-de-obras-generadas-con-ia-la-ley-exige-que-sean-de-creacion-humana/

47. Jensen Huang (2025). https://fr.benzinga.com/latest-news/jensen-huang-de-nvidia-declare-que-lia-pourrait-declencher-des-semaines-de-travail-de-4-jours-mais-avertit-que-nous-serons-plus-occupes-que-jamais/

48. Jensen Huang de Nvidia declare que l'IA pourrait declencher des semaines de travail de 4 jours (2025). https://fr.benzinga.com/latest-news/jensen-huang-de-nvidia-declare-que-lia-pourrait-declencher-des-semaines-de-travail-de-4-jours-mais-avertit-que-nous-serons-plus-occupes-que-jamais/

49. LG Uplus launches AI Universe to improve public accessibility to AI in Korea (2025). https://biz.chosun.com/en/en-it/2025/08/31/ZSGHOF7NPZGALPOXC2JJSRTEBY/

50. Les modeles simples sont parfois meilleurs que les modeles avances (2025). https://actu.epfl.ch/news/les-modeles-simples-sont-parfois-meilleurs-que-l-3

51. Microsoft Copilot (2025). https://www.hp.com/co-es/shop/tech-takes/microsoft-copilot-sector-salud

52. Microsoft Copilot transforma la atencion medica con IA (2025). https://www.hp.com/co-es/shop/tech-takes/microsoft-copilot-sector-salud

53. PHI 295 - Artificial Intelligence (AI) Ethics (2025). https://www.nvcc.edu/courses/phi/PHI295.html

54. Responsible AI (2025). https://www.cmu.edu/block-center/responsible-ai

55. Role of Artificial Intelligence in Identifying Health Inequalities: A Scoping Review (2025).

56. Soc(AI)ety Seminars Series (2025).

57. Sur internet (2025). https://www.tvanouvelles.ca/2025/08/29/sur-internet-la-vague-des-fausses-videos-generees-par-lia-alimente-le-sexisme

58. Sur internet, la vague de fausses videos generees par l'IA alimente le sexisme (2025). https://www.tvanouvelles.ca/2025/08/29/sur-internet-la-vague-des-fausses-videos-generees-par-lia-alimente-le-sexisme

59. Systematic literature review on bias mitigation in generative AI (2025). https://link.springer.com/article/10.1007/s43681-025-00721-9

60. Talking AI and the future of education with the University of Minnesota (2025). https://twin-cities.umn.edu/news-events/talking-ai-and-future-education-university-minnesota

61. The Bias Is Inside Us (2025). https://muse.jhu.edu/pub/1/article/968492/summary

62. The Development of Artificial Intelligence and Its Impact on Human Rights Protection from An Ethical and Legal Perspective (2025).
https://research.e-greenation.org/GIJLSS/article/view/499

63. The Impact of Generative AI Tools on Academic Work-load, Teaching, Learning, Integrity, and Engagement (2025).
https://www.learntechlib.org/primary/p/225321/

64. The Impact of Generative AI on the LIS Ecosystem (2025).
https://muse.jhu.edu/pub/1/article/968527/summary

65. The Impact of Generative AI on the LIS Ecosystem: Threats and Opportunities (2025).
https://muse.jhu.edu/pub/1/article/968527/summary

66. UH launches systemwide approach to artificial intelligence (2025).
https://www.hawaii.edu/news/2025/08/26/uh-launches-systemwide-ai-approach/

67. University of Toronto Artificial Intelligence Club (2025).
https://sop.utoronto.ca/group/university-of-toronto-artificial-intelligence-club-uoft-ai/

68. Values in the Age of AI (2025).
https://www.scu.edu/execed/news/top-stories/values-in-the-age-of-ai-preparing-tomorrows-ethical-executives.html

69. Your AI Bosses Are Still Prejudiced: The Emergence of Stereotypes in LLM-Based Multi-Agent Systems (2025).
https://arxiv.org/abs/2508.19919