

Historical Perspectives on Artificial Intelligence

October 20 at 9:00 a.m. ET
Dirksen Senate Office Building G11

Exploring AI's Dual-Decontextualization and Why Metaphors Matter

What's in a name?

- Artificial intelligence (AI) is rarely artificial and never intelligent. It is math—math dressed in abstraction and modeling for automation. It requires human labor to train and guide it.
- Dual-decontextualization is a problem of lost *historical contexts* and lost *data contexts*. AI emerged in a particular scientific and Cold War moment at the 1956 Summer AI Conference at Dartmouth. Elite, military-connected scientists marketed AI to the government.
- Some AI, with transparency and limited datasets, has been useful. But AI based on large language models (LLM) is context free, data appropriating, hallucination spawning, and risky. While existential threats are unfounded science fiction, generative AI involves major risks and realities of extending power imbalances, and amplifying racial, gender, and ableist biases. It can also proliferate misinformation and environmental harms.

What kinds of intelligence has AI been expected to have—and what kinds are left out?

Notions of intelligence in AI history

- The Turing Test, in which a computer is considered intelligent if it can fool a person into thinking it is a human being, equates intelligence with the ability to tell a successful lie, anticipating current AI's tendency to produce made-up results that merely look plausible.
- Chess: In the 1960s, the intelligence of AI systems was often measured by their ability to play and win games of chess. The choice of chess reflected cultural stereotypes of white male researchers. Benchmarks of intelligence always encode values.
- Expert Systems: Incorporating the knowledge base of human experts to produce correct results in specialized domains (i.e., diagnosing disease).

What is left out of this view of intelligence?

- Moral reasoning—we cannot automate moral reasoning or an ethic of care—and social intelligence, which stems from living in a society with other people.

What does it mean for AI to replace or be equivalent to a human being?

- AI works best at producing an expected or stereotyped version of a human being, e.g., customer service chatbots that follow scripts. Part of the “success” of AI in imitating a human is that we respond to conversational cues with our social instincts: we fill in the blanks.
- Human intelligence and human labor have always been part of the systems that make up AI.

What does it mean for AI to “solve” a problem?

- It is important to define the criteria for success for AI “solving” a problem. Is there an objective correct answer? Are there constraints on acceptable answers (e.g., avoiding gender bias in hiring decisions)? If AI needs to be trained, who decides what the parameters are?
- It is also important to define criteria for failure. Many generative AI programs are blocked from creating hate speech, even if the user asks for it. In that case, the user’s criterion for success conflicts with the public-interest criterion for failure.

AI and Declassification

- Whatever historians think of AI, it will be increasingly difficult, if not impossible, to preserve historical records and make them available to the public without better information technology.
- Recent years have witnessed a dramatic slowdown in declassification, which still depends on page-by-page review. This has stunted research on more recent military, intelligence, and diplomatic history, jeopardizing our ability to inform the public about critical foreign policy choices.
- At the same time, the executive branch has ignored legal requirements to report on the ever-expanding scope of classification activity and produce an official record of US foreign relations.
- For many years, the Public Interest Declassification Board has urged the use of AI as part of a more rational, risk management approach, and this idea is finally receiving high-level attention. Both the CIA and the State Department have been experimenting with AI for declassification, and the 2024 National Defense Authorization Act requires that the executive branch develop a comprehensive plan.
- But almost no information has been made public about this research, and the NDAA does not authorize any new resources for declassification. If nothing is done, the ever-growing volume and variety of classified information will continue to grow, undermining trust in government while making it harder to protect the relatively small amount of information that really could kill people.

Participant Biographies

Matthew L. Jones is the Smith Family Professor of History at Princeton University. In 2023, Norton published his *How Data Happened: A History from the Age of Reason to the Age of Algorithms*, written with Chris Wiggins. He is completing a book, *Great Exploitations*, on state surveillance of communications and information warfare. His previous books include *Reckoning with Matter: Calculating Machines, Innovation, and Thinking about Thinking from Pascal to Babbage* and *The Good Life in the Scientific Revolution: Descartes, Pascal, Leibniz and the Cultivation of Virtue*.

Janet Abbate is professor of science, technology and society at Virginia Tech. She is the author of two award-winning books: *Inventing the Internet*, the first scholarly history of the Internet, and *Recoding Gender: Women's Changing Participation in Computing*. Her most recent book is *Abstractions and Embodiments: New Histories of Computing and Society* (co-edited with Stephanie Dick). She is currently writing a history of computer science as an intellectual discipline.

Matthew Connelly is a professor of international and global history at Columbia University and director of the Centre for the Study of Existential Risk at the University of Cambridge. He is also the principal investigator of History Lab, an NSF- and NEH-funded project that uses data science to analyze state secrecy. His publications include *A Diplomatic Revolution: Algeria's Fight for Independence and the Origins of the Post-Cold War Era*, *Fatal Misconception: The Struggle to Control World Population*, and *The Declassification Engine: What History Reveals about America's Top Secrets*.

Jeffrey R. Yost is Research Professor, History of Science and Technology, and director of the Charles Babbage Institute for Computing, Information, and Culture at the University of Minnesota. The most recent of his seven books are *Computer: A History of the Information Machine*, 4th ed., coauthored, and *Making IT Work: A History of the Computer Services Industry*. He co-edits Johns Hopkins University Press' Computing and Culture book series, co-edits the journal *Interfaces*, and founded the *Blockchain and Society* blog.

About the American Historical Association

The American Historical Association's Congressional Briefings series seeks to provide congressional staff members, journalists, and other interested parties with the historical background to topics of current concern. The sessions are intended to be strictly nonpartisan and to avoid advancing particular policy prescriptions or legislative agendas. The AHA is grateful to the Mellon Foundation for its generous support of the Congressional Briefings series.

The AHA promotes historical work and the importance of historical thinking in public life. Incorporated by Congress in 1889, its mission to enhance the work of historians also encompasses professional standards and ethics, innovative scholarship and teaching, academic freedom, and international collaboration. As the largest membership association of professional historians in the world (over 11,000 members), the AHA serves historians in a wide variety of professions, and represents every historical era and geographical area.

For more, visit our website at www.historians.org.