Book Review

The AI Con: How to Fight Big Tech's Hype and Create the Future We Want

Beyond the Hype: AI, Power, Labor, and Institutional Accountability

Karnav Rana

Student, Golden Gate University

Kshitij Taluja

Undergraduate Researcher, Bachelor of Computer Science, Golden Gate University

Dr. Smrite Goudhaman

Global Professor of Practice, Golden Gate University

Author(s): Emily M. Bender and Alex Hanna

Publisher: Harper, an imprint of HarperCollins Publishers

Publication year: 2025

ISBN: 978-0-06-341856-1

Pages: 288

Abstract

The AI Con: How to Fight Big Tech's Hype and Create the Future We Want by Emily M. Bender & Alex Hanna examines artificial intelligence (AI) as a social, economic, and political phenomenon rather than merely as a technology. Rather than accepting the narrative currently being told about AI by its most prominent proponents, Bender and Hanna ask what it means to consider AI from the perspectives of power, labor, data, institutional responsibility, and the distribution of benefits from new technologies. One of the book's greatest contributions is demonstrating how organizations may make decisions based on hype about AI before sufficient evidence has

developed about whether those capabilities exist. At times, however, Bender and Hanna's strong critique of the current state of AI hype can obscure significant variations among different types of AI applications and therefore requires additional empirical study to test their recommendations across institutions.

Keywords: artificial intelligence; AI hype; technology governance; labor; institutional accountability; higher education

Bender and Hanna argue that competing narratives about AI can create a false urgency. Therefore, an important contribution of the book is not simply its criticism of individual AI applications, but its challenge to the way we think about and talk about those applications.

Bender and Hanna differentiate between what they call "intelligence" (claims made about some type of cognition) and what they refer to as "operation" (the actual process of how computer programs operate). Specifically, they focus on the generation of output by computational language models based on statistical patterns in large amounts of training data. They write that using terms like "smart" or "intelligent" may lead people to attribute understanding, intent, agency, or reliability to an automated system that has none of those characteristics in the same way that humans do.

This argument also connects directly with Bender et al. (2021), whose earlier critique of large language models similarly cautions against treating fluent system outputs as evidence of human-like understanding. The AI Con extends that concern beyond model behavior to the institutional and economic decisions made around such systems.

Perhaps one of the strongest conceptual moves of the book is treating AI hype as an institutional issue rather than simply as a series of exaggerated predictions about the future. By rejecting that common premise, Bender and Hanna shift attention away from speculative ideas

about how AI might evolve in the future, and instead focus on decision-making processes that can be assessed today (e.g., control over data, who pays for deployments, what kind of work is included in datasets, who profits when organizations implement AI-based systems). This focus on assessment of existing institutions makes the book particularly relevant to universities, employers, policymakers and other institutions considering implementing new uses for AI.

While the book offers vital insights, one of its shortcomings is that its forceful critique of AI hype can sometimes flatten meaningful differences between various AI applications. For example, a generative AI system designed to write internal corporate meeting summaries will likely result in less serious consequences than a similar AI system used in college or university admissions, determining individuals' eligibility for government assistance, or assisting with medical treatment decisions. AI systems used for internal summary writing will most likely raise issues related to accuracy, privacy, and copyright. In contrast, AI systems that assist with admission recommendations, government benefit determinations, or clinical decision-making will present much more critical risks regarding fairness, transparency, and accountability, as well as the potential harm caused by errors. Although the general critique is applicable to all of these examples, the regulatory framework used to assess them would need to differ substantially.

This book serves as a critical conceptual intervention rather than an empirical analysis of all the claims raised within it. While there is nothing inherently problematic about this approach, it does mean that many of the more general prescriptive ideas raised by Bender and Hanna will need to be tested in various institutional settings and national contexts. If future research were able to test whether organizations can operationalize Bender and Hanna's questions regarding

accountability, working conditions, public services, or informational quality, the argument presented in this book will be strengthened.

Applying this perspective to higher education, the reviewers see an opportunity to translate the book’s concerns into a practical process of institutional reflection. The sequence below is therefore a synthesis derived from the book rather than a framework explicitly presented by Bender and Hanna.

In particular, the reviewers see The AI Con as providing a suitable basis for institutional reflection in higher education. As universities determine how best to use generative AI in instruction, assessment, research, administrative functions, and student services, Bender and Hanna's methodological approach directs institutions toward asking the right questions. Instead of focusing on the simple question, 'Can an AI system do this?' institutions should examine the specific problem being addressed through the use of AI. They should then assess what evidence exists to support the adoption of a given AI system. Furthermore, they should examine what types of human labor are either being replaced or modified by the implementation of the AI system. Additionally, they should identify any data sources or intellectual property used by the AI system. Finally, institutions should evaluate who will remain accountable for the results produced by the AI system. These questions direct institutional decision-making away from enthusiasm for technology toward more responsible decision-making.

The book is most beneficial to readers who want to question the assumptions behind many of today's AI applications rather than to readers primarily interested in how they work. The authors have made an effort to write the book so that those with little or no background in computer science can understand it as well. However, they also raise questions regarding the use of AI and its broader social implications that researchers, policymakers, university

administrators, and business leaders can ask themselves about their own operations. As such, the book's most practical application is to provide a means by which individuals and organizations can think critically about AI-related decision-making prior to committing time, money, or authority to a given system.

Perhaps its most significant strength is its emphasis on viewing AI as a social and institutional phenomenon instead of accepting AI as a force acting independently and having a predetermined trajectory. Ultimately, Bender and Hanna's book represents a viable and accessible starting point for readers looking for a critical foundation for understanding AI outside of boosterist and catastrophist narratives.

References

Bender, E. M., & Hanna, A. (2025). The AI Con: How to Fight Big Tech's Hype and Create the Future We Want. Harper.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610-623. https://doi.org/10.1145/3442188.3445922