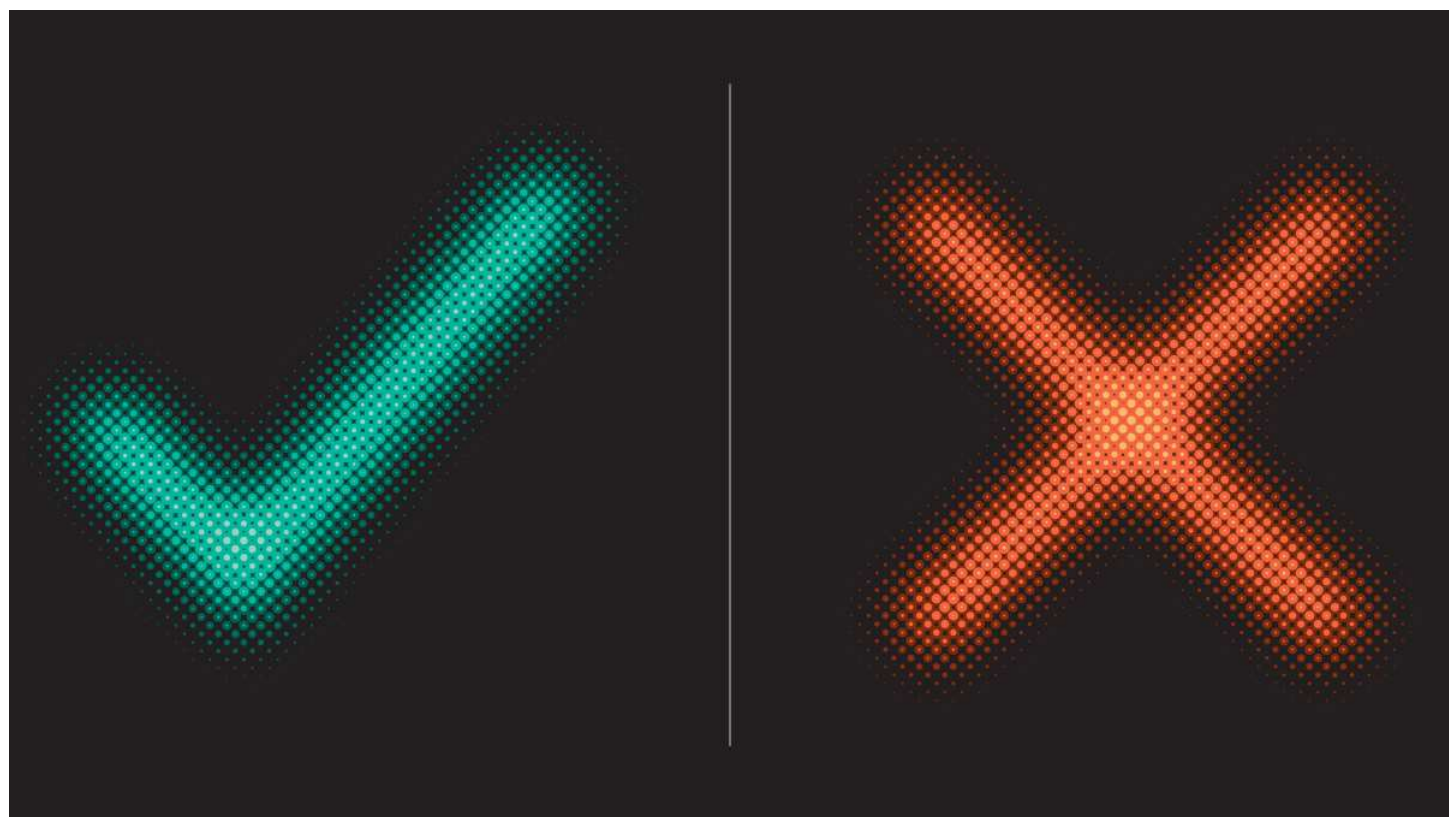


ECONOMICS & SOCIETY

Why We Need to Audit Algorithms

by James Guszczka, Iyad Rahwan, Will Bible, Manuel Cebrian, and Vic Katyal

NOVEMBER 28, 2018



THOTH_ADAN/GETTY IMAGES

Algorithmic decision-making and artificial intelligence (AI) hold enormous potential and are likely to be economic blockbusters, but we worry that the hype has led many people to overlook the serious problems of introducing algorithms into business and society. Indeed, we see many succumbing to what Microsoft's Kate Crawford calls "data fundamentalism" – the notion that massive datasets are repositories that yield reliable and objective truths, if only we can extract them using machine learning tools. A more nuanced view is needed. It is by now abundantly clear that, left unchecked, AI algorithms embedded in digital and social technologies can encode societal biases, accelerate the spread of rumors and disinformation, amplify echo chambers of public opinion, hijack our attention, and even impair our mental wellbeing.

Ensuring that societal values are reflected in algorithms and AI technologies will require no less creativity, hard work, and innovation than developing the AI technologies themselves. We have a proposal for a good place to start: auditing. Companies have long been required to issue audited financial statements for the benefit of financial markets and other stakeholders. That's because – like algorithms – companies' internal operations appear as "black boxes" to those on the outside. This gives managers an informational advantage over the investing public which could be abused by unethical actors. Requiring managers to report periodically on their operations provides a check on that advantage. To bolster the trustworthiness of these reports, independent auditors are hired to provide reasonable assurance that the reports coming from the "black box" are free of material misstatement. Should we not subject societally impactful "black box" algorithms to comparable scrutiny?

Indeed, some forward thinking regulators are beginning to explore this possibility. For example, the EU's General Data Protection Regulation (GDPR) requires that organizations be able to explain their algorithmic decisions. The city of New York recently assembled a task force to study possible biases in algorithmic decision systems. It is reasonable to anticipate that emerging regulations might be met with market pull for services involving algorithmic accountability.

So what might an algorithm auditing discipline look like? First, it should adopt a holistic perspective. Computer science and machine learning methods will be necessary, but likely not sufficient foundations for an algorithm auditing discipline. Strategic thinking, contextually informed professional judgment, communication, and the scientific method are also required.

As a result, algorithm auditing must be interdisciplinary in order for it to succeed. It should integrate professional skepticism with social science methodology and concepts from such fields as psychology, behavioral economics, human-centered design, and ethics. A social scientist asks not only, "How do I optimally model and

use the patterns in this data?” but further asks, “Is this sample of data suitably representative of the underlying reality?” An ethicist might go further to ask a question such as: “Is the distribution based on today’s reality the appropriate one to use?” Suppose for example that today’s distribution of successful upper-level employees in an organization is disproportionately male. Naively training a hiring algorithm on data representing this population might exacerbate, rather than ameliorate, the problem.

An auditor should ask other questions, too: Is the algorithm suitably transparent to end-users? Is it likely to be used in a socially acceptable way? Might it produce undesirable psychological effects or inadvertently exploit natural human frailties? Is the algorithm being used for a deceptive purpose? Is there evidence of internal bias or incompetence in its design? Is it adequately reporting how it arrives at its recommendations and indicating its level of confidence?

Even if thoughtfully performed, algorithm auditing will still raise difficult questions that only society – through their elected representatives and regulators – can answer. For instance, take the example of ProPublica’s investigation into an algorithm used to decide whether a person charged with a crime should be released from jail prior to their trial. The ProPublica journalists found that the blacks who did not go on to reoffend were assigned medium or high risk scores more often than whites who did not go on to reoffend. Intuitively, the different false positive rates suggest a clear-cut case of algorithmic racial bias. But it turned out that the algorithm actually *did* satisfy another important conception of “fairness”: a high score means approximately the same probability of reoffending, regardless of race. Subsequent academic research established that it is generally impossible to simultaneously satisfy both fairness criteria. As this episode illustrates, journalists and activists play a vital role in informing academics, citizens, and policymakers as they investigate and evaluate such tradeoffs. But algorithm auditing should be kept distinct from these (essential) activities.

Indeed, the auditor’s task should be the more routine one of ensuring that AI systems conform to the conventions deliberated and established at the societal and governmental level. For this reason, algorithm auditing should ultimately become the purview of a learned (data science) profession with proper credentialing, standards of practice, disciplinary procedures, ties to academia, continuing education, and training in ethics, regulation, and professionalism. Economically independent bodies could be formed to deliberate and issue standards of design, reporting and conduct. Such a scientifically grounded and ethically informed approach to algorithm auditing is an important part of the broader challenge of establishing reliable systems of AI governance, auditing, risk management, and control.

As AI moves from research environments to real-world decision environments, it goes from being a computer science challenge to becoming a business and societal challenge as well. Decades ago, adopting systems of governance and auditing helped ensure that businesses broadly reflected societal values. Let’s try replicate this success for AI.

James GUSZCZA is the US Chief Data Scientist for Deloitte Consulting LLP.


Iyad Rahwan is an Associate Professor of Media Arts & Sciences at the MIT Media Lab.

Will Bible is the chief digital officer and a partner of Deloitte & Touche LLP

Manuel Cebrian is a research scientist manager at MIT Media Lab

Vic Katyal is a principle and the global data risk and analytics leader at Deloitte and Touche LLP

This article is about ECONOMICS & SOCIETY

 FOLLOW THIS TOPIC

Related Topics: [AUDITING](#) | [DATA](#)

Comments

Leave a Comment

POST

6 COMMENTS

Ali-Abbas Ali 3 months ago

Unfortunately, you start from the premise that auditing is a value-adding, well functioning process. Successive business failures have proven that the asymmetries of information (businesses know their business better), and of skill (businesses pay better than auditors and so attract better calibre people), and poorly aligned incentives override any conceptual benefit from auditing. None of these issues will be resolved in the model you set out. Indeed, in putting together the multi-functional teams you propose, replicating the intellectual effort of the business, you are designing in a high cost, essentially a tax.

The asymmetries of information and skill are multiplicative through the policy chain. Auditors know less than the business, policy professionals setting the rules for auditors know even less, and politicians less still.

It would be better to measure outcomes - a more deliverable method, which will gain public traction for something new and unknown like AI is to say "I don't care how you have written your algorithm, but its outputs should be ones that are measurably fair" the onus is on the company to comply, and the penalties for not doing so, whether by design or incompetence, should be harsh.

In that model the challenge is to get politicians and policymakers to enact and enforce rules on an area of activity that is seen as a driver of international competitive advantage, efficiency and wealth creation. I concede this is also difficult.

My broader thoughts on the audit "profession" are covered in my blog post "The Eighth Year and the Myth of Audit"

REPLY

1  0 

 [JOIN THE CONVERSATION](#)

POSTING GUIDELINES

We hope the conversations that take place on HBR.org will be energetic, constructive, and thought-provoking. To comment, readers must sign in or register. And to ensure the quality of the discussion, our moderating team will review all comments and may edit them for clarity, length, and relevance. Comments that are overly promotional, mean-spirited, or off-topic may be deleted per the moderators' judgment. All postings become the property of Harvard Business Publishing.