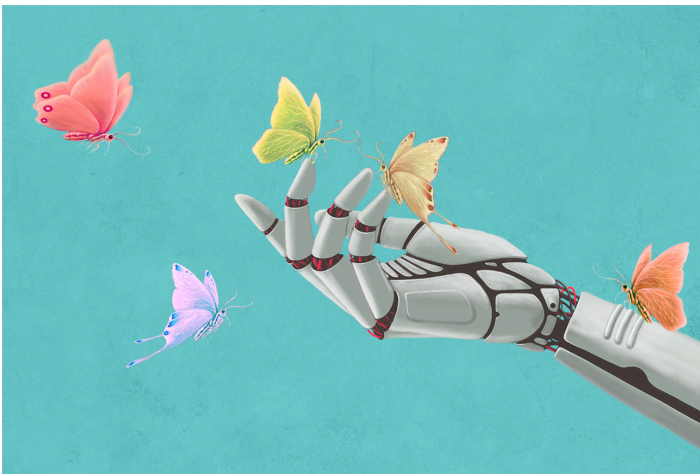


AI-at-Scale Hinges on Gaining a ‘Social License’

François Cadelon, Rodolphe Charme di Carlo, and Steven D. Mills

Businesses must make explicit efforts to build trust with all stakeholders if AI implementation is to succeed.



In January 2020, an unknown American facial recognition software company, Clearview AI, was thrust into the limelight. It had quietly flown under the radar until *The New York Times* reported that businesses, law enforcement agencies, universities, and individuals had been purchasing its sophisticated facial recognition software, whose algorithm could match human faces to a database of over 3 billion images the company had collected from the internet. The article renewed the global debate about the use of AI-based facial recognition technology by governments and law enforcement agencies.

Many people called for a ban on the use of the Clearview AI technology because the startup had created its database by mining social media websites and the internet for photographs but hadn't obtained permission to index individuals' faces. Twitter almost immediately sent the company a cease-and-delete letter, and YouTube and

Facebook followed suit. When the COVID-19 pandemic erupted in March 2020, Clearview tried to pitch its technology for use in contact tracing in an effort to regain its credibility and gain social acceptance. Although Clearview's AI technology could have helped tackle the crisis, the manner in which the company had gathered data and created its data sets created a social firestorm that discouraged its use.

In business, as in life, being responsible is necessary but far from sufficient to build trust. As exemplified by the controversies around some corporations' AI applications — such as [Amazon](#), which had to terminate its experiment with a resume-screening algorithm, and [Microsoft](#), whose AI-based chatbot was a public relations disaster — society will not agree to the use of AI applications, however responsibly they may have been developed, if they haven't a priori earned people's trust.

Rational people have a variety of concerns about AI, including the algorithmic institutionalization of income, gender, racial, and geographic prejudices; privacy concerns; and political issues. Indeed, Georgetown University's Center for Security and Emerging Technology and the Partnership on AI last year launched the [AI Incident Database](#) to record cases in which intelligent systems have caused safety, fairness, or other real-world problems; as of July, it listed 1,200 publicly reported cases of such AI failures from the past three years. That's why companies are struggling to come to terms with the gulf between what they understand

to be their legal rights to use AI and their social right, which they don't possess by default.

Why Responsible AI Isn't Enough

Adhering to the concepts of responsible AI enables companies to develop AI technology that works for the good; it forces business to go beyond algorithmic fairness and bias to identify the technology's potential effects on safety and privacy. However, following the doctrine of responsible AI has proved to be insufficient for two main reasons. First, it embodies a technology-based approach, with the focus squarely on the technical challenge of building goodness and fairness into algorithms. However, it's impossible to eliminate bias from AI unless all humans become good, fair, and unbiased — and that isn't likely to happen anytime soon.

Second, without formal regulations to follow, the articulation of each organization's principles, as well as its adherence to them, has been left to the very data scientists and software developers who write algorithms. Thus, the principles naturally vary from company to company, and, inside organizations, by business and even function, which is hardly ideal. The issue has been further confused by the variance among the multiple responsible AI guidelines and principles proposed by different organizations and companies.

Underlying the issue is the reality that while being trusted and being responsible may be related, they are distinct from each other. Responsibility may foster trust, but it can never be a substitute for it.

Before businesses can obtain society's explicit approval to use AI applications at scale, two things need to happen. First, public institutions must enact laws and regulations, and companies must face penalties for not abiding by them. Second, businesses must adopt a human-focused approach that fosters trust among all their stakeholders — employees, executives, suppliers, shareholders, communities, civil society, and government — in the AI applications they develop. Thus, companies must gain what could be thought of as a social license for AI applications.

The Foundations of Trust

Stakeholders' trust in AI applications stem from three sources:

Social contracts. Stakeholders must accept that companies that develop AI applications can be trusted with their use as well as with the acquisition of real-time data to feed their algorithms. This especially holds true in cases where the algorithms can act independently of supervision and may have an impact on human lives. That's one reason why countries have been slow to approve the unrestricted use of self-driving automobiles.

Benefits. Businesses' perception that the advantages of using AI are greater than the costs of doing so must be widely shared by other stakeholders. It must be possible to determine the tangible and intangible trade-offs at the individual, company, and societal levels by weighing the benefits of AI-delivered outcomes — such as increased health, convenience, and comfort, in the case of health care — against the potential downsides, such as reduced security, privacy, or safety. Society's verdict may not always favor the use of an application, and businesses should be prepared for that.

Responsibility. If companies are to be answerable to society, they must be able to justify how their AI algorithms work. Businesses should be as open as they can about the algorithms they design and, whenever necessary, be able to explain the manner in which they arrive at their decisions. AI will be deemed fair only if the outcomes generated by the technology don't vary based on demographic factors or contextual changes. For example, a company that uses an AI-based recruitment system must be able to demonstrate that all of the candidates who provided the same or similar responses to a question posed by the machine on different days received the same rating or score.

Winning Trust Through Dialogue

To develop society's trust, companies must initiate discussions with all of their stakeholders to bring out the

reactions, motivations, positions, and objections of different groups of people. These conversations will allow companies to develop a shared understanding with stakeholders about each AI application and the guardrails that must fence its deployment.

Companies can work to gain stakeholders' trust in AI by taking four steps:

Identify stakeholders. The number of stakeholders whose trust will be required will vary according to the problem the AI system is trying to tackle. For instance, when German delivery company DHL wished to install an AI-based pallet-sorting system two years ago, it had to worry only about the reactions of employees, trade unions, and shareholders. In contrast, Waymo and Tesla must work with a large number of stakeholders — automobile owners and drivers (consumers); city, state, and central governments (regulators); and civil society (critics and advocates) — to obtain a social license for self-driving technologies.

Be transparent. Be it the assumptions underlying their algorithms or the impact of AI on employees, transparency alone will go a long way toward allowing businesses to proceed with the use of AI applications.

Many companies haven't developed the expertise to explain how their algorithms work without giving away proprietary information. They must partner with the major AI players to figure out how to do so. For example, scientists are creating [self-explainable AI](#) that can provide a decision as well as a supporting explanation, without sacrificing the accuracy of the former. In the same way, causal AI identifies the factors in an algorithm that lead to particular outcomes and then tests what will change them. Participating in AI ecosystems or partnering with government initiatives, such as the U.S. Defense Advanced Research Projects Agency's [Explainable Artificial Intelligence project](#), can give businesses access to tools and techniques for fostering transparency.

Companies must be open internally about the implications for their workforce when they implement AI. They must identify the [gaps as well as the surpluses in the workforce](#) that will be created by the use of AI and develop forecasts for the job functions that will be most significantly affected by an application's rollout. Businesses must upskill and reskill

affected employees, whenever possible, in order to fill new positions — while being transparent about the effects on compensation. Not all jobs created as a result of AI will necessarily pay better than those rendered obsolete by AI.

Manage risk. Companies must learn to mitigate AI-related risks by carefully mapping and evaluating their severity and probability — as well as the context, such as the regulatory environment. Managing AI-related risks doesn't differ much from tackling other kinds of risk; chief risk officers can determine the levels of legal, financial, reputational, and physical risk they're willing to take.

However systematic risks are assessed, companies must institutionalize human oversight over AI. They must ensure that the AI system learns to flag exceptions that it cannot process and recommend human over-the-loop interventions to deal with them. In March 2020, for example, the AI-based cybersecurity mechanisms of one of Britain's premier online grocery stores suddenly shut down its website because it had confused the COVID-19-related surge in demand with a distributed denial-of-service attack. If the retailer had created a way for human decision makers to monitor and override the system's knee-jerk response, it would have generated more revenue rather than antagonizing customers that day.

Communicate and educate. Companies have to explain all the possible benefits of their AI applications as well as all the potential downsides. They shouldn't shy away from describing the risks while explaining how it will tackle them. Google, for example, shares an organizationwide commitment to ensuring the safe, fair, and unbiased use of AI by publishing and periodically updating its [principles for responsible AI](#). It serves as a public reference for anyone who wants to hold the company accountable for complying with its own rules.

Business would do well to help governments at all levels understand the technology and its impact. Regulators find it difficult to keep pace with technology and should welcome the help of business leaders in developing AI-related policies. Close engagement with governments will help companies proactively develop AI technology that aligns with the direction of future laws and regulations; meanwhile, regulators could help accelerate the pace at which AI earns

social approval.

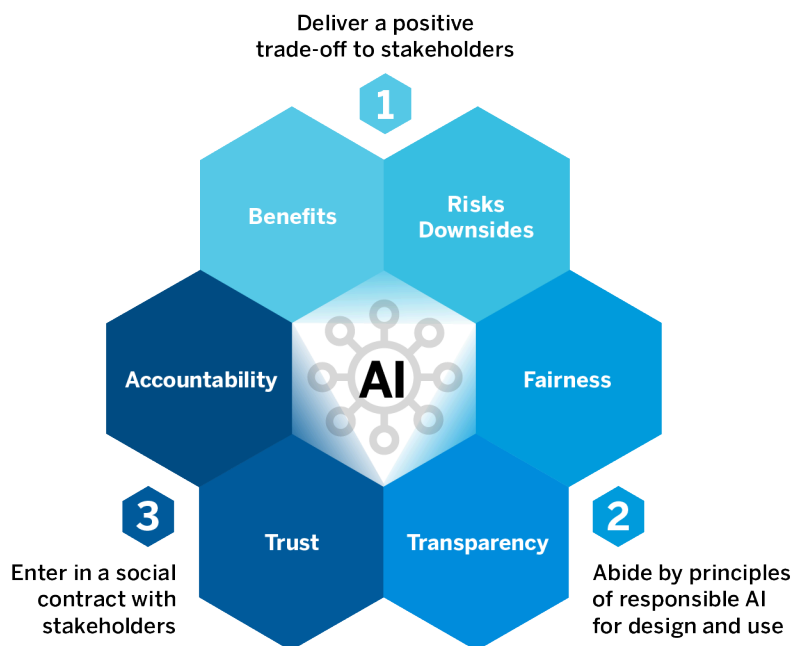
Currently, businesses planning to deploy AI are most concerned with obtaining an economic license to do so — that is, gaining support from shareholders and executives — and a legal license, in the form of regulatory permits and compliance with statutory obligations. But only when companies are able to earn a social license, winning the trust of employees, customers, and society at large, will they have what is required for the sustained use of AI at scale.

About the Authors

François Cadelon is a managing director and senior partner at Boston Consulting Group and the global director of the BCG Henderson Institute. Rodolphe Charme di Carlo is a partner at Boston Consulting Group and an ambassador at the BCG Henderson Institute. Steven D. Mills is a managing director, partner, and chief AI ethics officer at Boston Consulting Group.

The Three Pillars of the Social License for AI

Winning a social license to deploy AI will require organizations to adhere to the principles of responsible AI design; ensure that all stakeholders perceive the benefits of using AI as greater than the costs; and demonstrate that they can be trusted and will be accountable.



Reproduced with permission of copyright owner. Further reproduction
prohibited without permission.