Why People Resist Embracing AI
The five main obstacles—and how to overcome them

AUTHOR

Julian De Freitas
*Assistant professor,
Harvard Business School*

ARTIFICIAL INTELLIGENCE has created a striking paradox. Consider that in a 2023 Gartner survey, 79% of corporate strategists said that the use of AI, automation, and analytics would be critical to their success over the next two years. But only 20% of them reported using AI in their daily activities.

AI's success hinges not only on its capabilities, which are becoming more advanced every day, but on people's willingness to harness them. And as the Gartner findings suggest, AI is not getting great traction with users.

Unfortunately, most people are pessimistic about how it will shape the future. Seventy-seven percent of Americans are concerned that its adoption will cause job losses within the next 12 months, according to research by Forbes Advisor. Eighty percent think AI has increased the likelihood that their personal data will be used in malicious ways by criminals. And it gets worse: A poll conducted by YouGov found that nearly half of Americans believe that one day AI will attack humanity. With this much cynicism about AI, getting workers to willingly, eagerly, and thoroughly experiment with it is a daunting task.

In conducting more than a decade of research on adoption of the technology, including in-depth qualitative interviews and experiments with some 2,500 users, I have uncovered what's driving this resistance to AI: fundamental human perceptions that AI is too opaque, emotionless, rigid, and independent, and that interacting with humans is far more preferable. Understanding those drivers is critical to designing interventions that will increase AI adoption inside organizations and among consumers generally. In this article I will delve into them in detail and explain what you can do as a manager to counter them.

People Believe AI Is Too Opaque

The machine-learning algorithms underlying many AI tools are inscrutable "black boxes" to users. Their impenetrability frustrates people's basic desire for knowledge and understanding, especially when their outcomes are uncertain or unexpected. Studies find that people are willing to use opaque AI if it outperforms humans or simple, transparent AI, but they may balk at using it when its performance is more or less equivalent.

People tend to think that humans' decision-making is less of a black box than algorithms', but that belief is unfounded. Psychologists have shown that people have little insight into what other people are thinking; instead they use heuristics to interpret human behavior. In one study where people were asked to describe the process by which AI or human physicians diagnosed cancer after examining skin scans, for instance, participants realized that their grasp of the human diagnostic procedure was not as strong as they had presumed. This realization made them less biased against using medical AI.

Explanations of how AI tools work can increase their acceptance, but not all explanations are effective.

Organizations can promote the adoption of AI tools by anthropomorphizing them—for instance, by giving them a gender and a human name and voice.

Researchers have found that people prefer explanations about *why* an AI tool did something (for instance, they would rather know that an autonomous car braked because there was an obstacle ahead) to simple explanations of *what* the AI did (for instance, that it activated the vehicle's braking system and brought the car to a halt).

The style of explanations plays a crucial role too. Those that use comparative reasoning—outlining why certain alternatives weren't chosen—increase trust more than explanations that don't. For instance, one study found that an explanation that described in detail why an AI system categorized a tumor as malignant instead of as benign was seen as more credible than one that simply said the tumor resembled other tumors (even if both explanations may be true). Essentially, the most convincing explanations are those that articulate the reasons behind the decision made as well as why alternatives were dismissed.

Because the best-performing AI models are typically more complex and harder to explain than other models, managers may want to first introduce simpler models into their organizational processes—especially when getting buy-in from people is important. Consider how Miroglio Fashion, a large Italian retailer of women's apparel, approached automating the task of forecasting demand at its 1,000 stores—which at the time was performed by each store's local manager. The company developed two models. The first was simpler and understandable, leveraging a regression approach that broke down the inventory by basic clothing features like category, fabric, color, and price to predict which items would need to be allocated to each store. The second model was more complex and obtuse, using sophisticated image analysis to identify harder-to-describe visual features of clothing—such as shape, layering and draping, and combinations of materials—to make the same predictions. Even though the complex model outperformed the simpler one, the executive team first introduced the simpler model to ensure that employees developed confidence in AI and were motivated to use it. The company ran a 13-week pilot test with some stores, demonstrating that those locations were more successful than those without the model, before rolling out the simpler model to all stores. Only after about a year of using it did the retailer graduate to the more-sophisticated AI model.

The complexity of an explanation relative to the task's also matters. Research has shown that if an explanation indicates that the algorithm seems too simplistic for a task—for instance, if the AI compared a scan against just one image of a tumor to diagnose cancer—users may be less likely to follow the AI's guidance. However, an explanation that suggests that the algorithm is complicated—perhaps saying that the AI compared a scan against thousands of examples of how a malignant tumor looks and how a benign tumor looks and consulted medical research to back up its assessment—does not reduce adherence. This means that you should understand users' perceptions of the task before crafting an explanation.

You should also avoid suggesting that your AI is too simple for it; otherwise, it might be better to provide no explanation at all.

People Believe AI Is Emotionless

Though consumers tend to ascribe some human capabilities to AI tools, they don't think that machines can experience emotions and therefore are skeptical that AI can accomplish subjective tasks that seem to require emotional capabilities. That skepticism hinders the acceptance of AI systems that already can perform subjective tasks at the same skill level as humans can—such as recognizing emotions in faces and producing still images and video. For example, people are just as open to financial recommendations from AI as they are from humans, because that task is viewed as objective. However, when it comes to something like dating advice, which is seen as highly subjective, they have a clear preference for human input.

Organizations can address this hurdle by framing tasks in objective terms—by focusing on their quantifiable and measurable aspects. For example, with AI-generated dating advice, you could highlight the benefits of relying on quantifiable outcomes from personality assessments to guide the matchmaking process. The online dating service OkCupid complements its algorithms with personality assessments and extensive user-data analysis; it also emphasizes how the algorithms filter and rank potential matches to find the person who perfectly fits a user's preferences.

Organizations can also promote the adoption of AI tools by anthropomorphizing them—for instance, by giving them a gender and a human name and voice. In one study using an autonomous vehicle simulation, participants expressed greater trust and comfort when the vehicle's AI had features like a human voice and a human avatar. Another example is Amazon's Alexa, which has a female gender and some humanlike traits, including a name and a voice. These features create a familiar personality that helps users relate to AI better and feel more comfortable interacting with it.

Other researchers have found that individuals who have a lower tendency to anthropomorphize AI also have less trust in AI's abilities, leading them to resist using it. For instance, people who are less inclined to humanize a telemarketing chatbot tend to end calls quicker than they would with a human telemarketer.

While anthropomorphizing AI can often increase adoption, sometimes it can be counterproductive, such as in sensitive or embarrassing contexts like obtaining medicine for sexually transmitted diseases. In those situations, consumers often prefer AI without human traits because they believe it will be less judgmental.

People Believe AI Is Too Inflexible

People generally hold the view that mistakes help humans learn and grow, instead of interpreting errors as a sign of unchangeable defects. But they frequently think AI tools are rigid and not adept at adjusting and evolving—a belief that may stem from past experiences of machines as static devices that carry out limited functions.

Perceptions like that can diminish trust in the technology and create concerns about its efficacy in new scenarios. Studies have indicated, however, that consumer use of AI output rises when people are told that AI has the capacity for adaptive learning. Even nominal cues that imply learning potential, such as branding AI as "machine learning" instead of merely an "algorithm," have boosted engagement. Netflix frequently publicizes how its content recommendation algorithm continuously improves its selections as it collects more data on users' viewing habits. It reinforces that message by putting labels like "for you" on its recommendations and explaining that they were made "because you watched x," further reassuring users that the algorithm is considering their evolving preferences.

People who think that AI is inflexible may believe that it will treat every person identically, rigidly applying a one-size-fits-all approach that ignores an individual's unique traits. Indeed, the more distinctive consumers perceive themselves to be, the less likely they are to use AI. In one study, for instance, the more exceptional participants thought that their own ethical characteristics were, the more resistant they were to an AI system that assessed moral qualities.

At the same time, there's a delicate balance between flexibility and predictability. Even though adoption often increases when companies highlight AI's ability to learn and evolve, if users feel that the outputs of the system are too unpredictable, the intervention could backfire.

A more adaptable AI system is also riskier since it allows a greater spectrum of user interactions, some of which may not be captured in the data used to train the AI. When AI is more flexible, it increases the possibility that people will use it in inappropriate ways and that in those cases the algorithms might provide undesirable responses, creating new risks for users and companies alike.

A study my coresearchers and I did reveals how. First we analyzed more than 20,000 human–AI conversations on five AI-based companion apps and found that about 5% of the users were discussing serious mental health crises with them. In essence they were using the apps as therapists rather than companions. Next we sent more than 1,000 crisis messages to the apps and asked trained clinical experts to classify the responses. The experts and I determined that 25% of the AI-generated responses were problematic because they increased the users' likelihood of harming themselves. Then we asked a separate group of people to consider how each app had responded to the crises. Most of them gave the apps low ratings, indicated that they would stop using the apps, and said the app companies would be liable if the users ended up hurting themselves.

Therefore, AI systems must balance flexibility against predictability and safety. To do that they can incorporate user feedback and include safeguards for handling unexpected input appropriately.

People Believe AI Is Too Autonomous

AI tools that can perform tasks without active human input often feel threatening to people. From early on in life humans strive to manage their surroundings to achieve their goals. So they're naturally reluctant to adopt innovations that seem to reduce their control over a situation.

AI endows algorithms with a high degree of independence, allowing them to formulate strategies, take action, and keep refining their capabilities, all while adjusting to new situations without needing direct human guidance. The possibility that AI tools might completely take over tasks previously handled by humans, rather than just assist with them, stirs up deep concerns and worries. A significant majority of Americans (76%) are apprehensive about being passengers in self-driving vehicles, for instance. Similarly, people are afraid that smart home gadgets might invade their privacy by surreptitiously gathering their personal data and using it in unforeseen ways.

People also resist surrendering tasks to AI because they believe their personal performance is superior to the technology's. Interestingly, in experiments with more than 1,600 nationally representative U.S. participants ranging in age from 18 to 86, I found that people chose higher levels of vehicle automation for others than they did for themselves. The reason? They believed that they were better drivers than the automated vehicles were but that other people were not.

To increase utilization of AI systems, companies can restore consumers' sense of agency by having people provide input to the systems (thereby creating what are known as "human-in-the-loop systems"). Consider Nest, a smart home product that allows users to customize it, such as by manually adjusting thermostats or setting specific schedules. Its users can choose between automated learning and manual input. A sense of control can also be heightened by tweaking design elements of the product. For example, iRobot programs the Roomba vacuum to move in predictable paths rather than unpredictable ones that may make the vacuum appear "alive."

Allowing people to have too much control over AI systems can potentially diminish the quality of their output and their effectiveness, however. Fortunately, studies find that consumers need to retain only a small amount of input to feel comfortable. Marketers can thus calibrate AI systems so that there is an optimal balance between perceived human control and the systems' accuracy.

People Would Rather Have Human Interaction

In one of my studies I examined whether people preferred being served by human salespeople to being served by hypothetical AI-enabled robots whose appearance and physical and mental capabilities were described as being indistinguishable from those of humans. On a range of measures, including anticipated comfort interacting with human or robot salespeople, willingness to visit stores where they worked, and anticipated level of customer service, people consistently preferred humans. This stemmed from the belief that robots didn't have humanlike awareness and lacked the capacity for understanding meaning. In addition, the more different from humans that people felt that robots were (as measured by asking them to rate their agreement with statements like "Morally, robots will always count less than humans"), the more strongly they exhibited this preference.

Cultural context is most likely an important factor in anti-AI tendencies. In Japan, for example, the belief that even inanimate objects have souls or spirits is more widespread than in other countries, which may lead to greater acceptance of AI that highly resembles humans.

NO MATTER HOW much money your business invests in artificial intelligence, your leadership team must consider the psychological barriers to its adoption. And with each of the five barriers I've described, you must realize that interventions meant to increase acceptance can inadvertently increase resistance to AI.

Rather than leaping straight into solution mode, tread carefully. Every AI system, use case, pilot, and full-scale deployment will encounter different barriers. It's your job as a leader to recognize them and help your customers and employees overcome them. 🛡

HBR Reprint S25013
Spotlight Package Reprint R2501B

JULIAN DE FREITAS *is an assistant professor in the marketing unit at Harvard Business School.*