5º

BOOK CLUB SYNOPSIS

Power and Prediction
The Disruptive Economics of Artificial Intelligence

Ajay Agrawal, Joshua Gans, and Avi Goldfarb

Harvard Business Review Press, 2022

Listen to an audio version of this summary and past Book Club selections in the World 50 app.

Contents

Part One:
The Between Times

CHAPTER ONE

A Parable of Three Entrepreneurs

When electricity was invented, people were enthusiastic—but it still took time for it to be widely adopted. This interim period, which the authors call “The Between Times,” is similar to the moment we’re in now regarding adoption of AI. “We are now in the Between Times for AI—between the demonstration of the technology’s capability and the realization of its promise reflected in widespread adoption.”

According to the parable of the three entrepreneurs, developed by the authors, the first use of electricity was in “point solutions.” People swapped out steam for electricity at the point of entry to factories due to the promise of cost savings. This was most often seen in steam-powered mills and clothing and textile manufacturers. Later, people recognized that electrical power could be easily turned off and on, enabling users to only use it when needed.

Over time, these adaptations and developments led to more significant changes in the design of the tools used in a factory. But this, in turn, yielded another question: With the benefits and advantages of electricity now more clearly understood, how could a factory be completely redesigned to best make use of electricity?

This parable shows how different value propositions can be exploited: The first is a point solution, the second an application solution, and the final development led to a system solution. Each of these also highlights different lessons:

1. Once you understand what a new technology offers, you can increase productivity.

2. There are opportunities to completely redesign products, services, and even factories with new technology.
3. Each type of solution can help you gain a foothold in the market.

Even though the example of electricity demonstrates the potentially huge impact we might see from AI, it will still take time to see broad disruption. When it does come, however, the authors posit that this disruption will come from system change.

CHAPTER TWO

AI's System Future

While AI may not yet be as transformational as the authors expect that it will be someday, there are numerous areas in which it is already making an impact. For the most part, organizations that already did prediction work are seeing the benefits of adopting AI, yielding improved, faster, and cheaper predictions.

The authors argue that when we begin to create system solutions using AI, we will see the real benefits AI can bring. They revisit the definitions of point, application, and system solutions and then proceed to provide AI-specific definitions for each of the three. They note that "advances in modern AI are, at their essence, an improvement in prediction technology" and that predictions are valuable as "inputs into decision-making."

- AI point solution: "A prediction is valuable as a point solution if it improves an existing decision and that decision can be made independently."
- AI application solution: "A prediction is valuable as an application solution if it enables a new decision or changes how a decision is made and that decision can be made independently."
- AI system solution: "A prediction is valuable as a system solution if it improves existing decisions or enables new decisions, but only if changes to how other decisions are made are implemented."

In simpler terms, the first two types of solutions create changes within our existing system, while the third creates disruption—as well as the biggest potential return on investment.

CHAPTER THREE

AI Is Prediction Technology

When many people hear the acronym AI, robots with the capacity for thought come to mind. But the authors argue that this is not currently what the technology enables; really, it's more of "an advance in statistical techniques."

One of the most important recent developments in AI is deep learning, which is, in essence, machine prediction. This is one input into decision-making; the other two are judgment and data. Judgment is "the process of determining the reward to a particular action in a particular environment," and data, in turn, provides the information by which a prediction can be made. Whereas prediction tells you how likely it is that something will happen, judgment reflects how you value that outcome.

Yet the availability of data is not a given. At times, it is available or can be collected via experimentation, but the involvement of multiple actors can create uncertainty, rendering the acquisition of good data quite difficult. In a business setting, this might occur when a competitor or a customer has an incentive to subvert or circumvent your efforts.

When good data is available, an AI point solution works fine. One example the authors provide is how credit card companies detect fraud. They are unable to individually verify each transaction, so the companies rely on predictions. They know that in some instances, they will fall prey to declining legitimate transactions or missing instances of fraud, so they set thresholds to balance these out.

“Prediction is the process by which you convert information that you have into information that you need.” Many of the ways that AI is being used today are what the authors call “low-hanging fruit”—they’re everywhere and so integrated into our lives that we hardly notice them anymore. But the opportunities don’t end there. There are far more ways for AI to be used, and they may not be immediately obvious.

Part Two:
Rules

CHAPTER FOUR

To Decide or Not to Decide

Choosing the best possible outcome in the decisions we face each day is likely not tenable. That's why we create habits and rules—they are an attempt to make our lives easier by forgoing decision-making entirely.

Many people resort to "making do," or what economist Herbert Simon called "satisficing." They would prefer to take the "good enough" option rather than go through the effort of seeking out the best option. "Rather than continually updating their choices based on new information received, they would adopt rules, routines, and habits that would be impervious to new information and, hence, allow them to ignore information entirely."

Several factors influence whether something will fall in the default or active decision category. One dichotomy is high versus low consequences. If the consequence of the decision is low, then there is little need to spend energy on making an active decision. Another major consideration is the cost of the information needed to make a certain decision. If it is laborious to gather the information required to make the best possible choice, we'll ignore the information and create a rule for every similar instance.

With so many of us relying frequently on rules, the case for AI seems less clear. Yet the authors argue that AI will enable better decision-making by readily providing the information needed such that we will no longer have to rely on rules. In fact, they contend that "rules are not really a lack of making a decision"; rules are simply making a preemptive decision.

It can be difficult to break rules and habits once they're set. We may believe that they serve a purpose, that they are inherently better than an active decision, and that they enable reliability and decrease uncertainty. But they're not always beneficial—particularly at the organizational level, rules can make it hard to change tack when something isn't working. Rules, unlike decisions, don't allow us to easily incorporate new information that becomes available.

The authors posit that as AI improves—and therefore, so does the quality of predictions—the cost of active decision-making decreases. Making a decision rather than relying on rules will ultimately yield a higher return.

CHAPTER FIVE

Hidden Uncertainty

The idea of rules as outlined in the previous chapter returns here, exploring specifically the rules people adopt regarding air travel. People know that airports can be crowded and, therefore, create rules for when to leave for the airport. This creates a situation in which travelers tend to spend more time in airports, which benefits airport retailers. While the amenities at airports create an illusion that being there is pleasant, they really serve to obscure the fact that people are waiting there because of the uncertainty associated with travel.

AI could fix this. It could, for example, be used to predict wait times in security lines at the airport. But the incentives of the retailers and the travelers are not aligned. Increasing certainty about how early a traveler should arrive at the airport would lead to them arriving later—ultimately spending less time and not making use of the amenities contained within the airport.

The larger point in this chapter is about rules. While they are intended to decrease uncertainty, they can also have hidden costs. The Shirky Principle contends that "institutions will try to preserve the problem to which they are the solution." If an organization benefits, for example, from you needing to arrive at the airport early, then it is unlikely to allow developments that might change your approach.

The other issue is that rules may also cause unnecessary financial costs. Rules frequently lead to costly investments to protect users from risk—but these risks are often rare occurrences, which makes it hard to determine whether the investments are worthwhile. Additionally, the mitigation tactic makes it harder to measure whether it is excessive, or whether the event is actually high risk or not.

CHAPTER SIX

Rules Are Glue

Rules are a result of and a response to uncertainty. The authors illustrate this using the example of marketing. Marketers try to segment the population into groups in an effort to target the right audiences with the right products. If they actually *knew* what each person wanted, they wouldn't have to rely on rules; instead, they would be able to target individuals with the right product.

We also see tensions that arise due to rules. In education, for example, there is distinct tension between rules and flexibility. We maintain standards based on what students should be learning, but in practice, they all receive a different education. Students may also need or benefit from different teaching styles or paces. Personalized education—a flexible approach—would be best, but the rules surrounding education make it difficult to enact.

The point here is that it can be challenging to notice or uncover opportunities for change when rules are long-standing. Historically, network television stations in the U.S. followed the same rule for advertising: For every half hour of programming, eight minutes were to be used for ads. This meant that programs were written to last either 22 or 44 minutes. YouTube upended this model. Creators can make content of any length, and then the system's AI can both be used to lead users to the content they'll most enjoy and to predict what ads will resonate most with these viewers.

AI is useful and valuable for YouTube advertising, but not television advertising, because different users can view different content on YouTube in a way that they can't on TV. Additionally, AI "*indirectly* enables flexible content lengths because the solutions for discovery and advertising solve the problem of an infinite number of combinations of content, advertising, and timings."

A similar phenomenon is at play in education. We rely on a fixed curriculum—similar to TV's fixed advertising rules—which means that personalized education couldn't work in the existing system. "In other words, the age-based curriculum rule is the glue that holds together much of the modern education system, and so an AI that personalizes learning content can only provide limited benefit in that system."

The larger lesson here is that using AI in a limited way within an existing system laden with rules will have a limited impact. "Rules glue together in a system. That's why it's hard to replace a single rule with an AI-enabled decision. Thus, it's often the case that a very powerful AI only adds marginal value because it is introduced into a system where many parts were designed to accommodate the rule and resist change. They are interdependent—glued together."

Part Three:
Systems

CHAPTER SEVEN

Glued Versus Oiled Systems

One of the greatest challenges of the COVID-19 pandemic, according to the authors, was a prediction problem: “We lacked the information to predict who was infectious and could spread the virus to others.” If we could have made more effective use of AI, we could have predicted who was infectious and dramatically changed our response, overcoming the need for costly mitigation tactics like lockdowns and social distancing.

Contact tracing and improved testing were effective for prediction, but they weren’t enough. “The system was glued together by many rules that did not lend themselves to information-based decision-making.” The global health system was not prepared to respond to information in real time, resulting in numerous other negative impacts, including on the economy and mental health. “The pandemic reminds us that we often gravitate to rules and that rules carry their own inefficiencies.”

These areas—where we rely on rules rather than information—are what the authors see as the prime targets for introducing AI. When we pay attention to the fact that a rule is in place, we can reframe the situation and consider how AI prediction might allow us to make better, active decisions instead.

CHAPTER EIGHT

The System Mindset

Cars replaced horses because they could provide the same service but better. Similarly, if machines can do something better and cheaper than a human, replacement will likely follow. This is how many people conceive of AI today: If a machine can do the same job as a human, that job will one day be completed by a machine instead.

Yet the authors argue that this line of thinking is too focused on where AI can substitute in at the task level, when the real opportunities are present at the system level. They advocate for adopting a "system mindset" regarding the adoption of AI. In contrast to a task mindset, the system mindset "sees the bigger potential of AI and recognizes that to generate real value, systems of decisions, including both machine prediction and humans, will need to be reconstituted and built."

From a sales pitch perspective, they argue that it is also more effective to focus on how AI can add value as opposed to how it can lead to cost savings. Rather than emphasizing how a machine can do something more cheaply than a human, this line of argument focuses on how AI can improve profitability by improving a service or product. Small cost savings are not what will lead to large-scale adoption of a new system—and the replacement of an existing one. "A pure cost calculus will rarely drive such replacement. … Instead, if the system does something new—that is, leads to new value creation opportunities—then that is what will drive adoption."

CHAPTER NINE

The Greatest System of All

AI has the potential to change the innovation process. “Better prediction at the hypothesis development stage could lead to an entirely new system.” Whereas in the past the scientific method relied on trial and error, AI could provide predictions surrounding new combinations to generate better hypotheses. Given innovation’s central role in society’s well-being, economic growth, and productivity, the authors believe that “the impact of AI on innovation may ultimately outweigh the impact of all other applications of AI.”

There is, nevertheless, a challenge. Even though AI, like many new technologies, offers numerous opportunities, well-established companies have little incentive to adopt it. There’s too much risk involved in transforming how their industry operates. “This is why technological change can lead to disruption,” and small firms and startups are often at the forefront of it.

Startups have more of an incentive to innovate. Over time, what they have created matures and becomes a viable alternative to what already exists, and in the process, they threaten the existing businesses.

Part Four:
Power

CHAPTER TEN

Disruption and Power

The first three sections of the book examined point, application, and system solutions and outlined how the adoption of the first two typically comes first; however, system solutions are where the real opportunities lie. Nevertheless, it can be difficult to create system solutions, as they require overhauling the rules that have long kept the existing system stable and in place.

The rest of the book explores what AI innovations at the system level may look like and what challenges may arise. First and foremost, it's necessary to understand that such changes will be disruptive. Although this term is too often used as a buzzword, the authors emphasize that there are three main reasons why systemwide AI adoption will be truly disruptive:

1. Existing industries tend to have blind spots regarding how AI can be used.
2. Big technological changes require replacing existing systems with new ones.
3. These new systems typically lead to a power shift—specifically, a shift in economic power.

Established companies are the most likely to suffer the consequences of such a disruption. This is because they are less likely to willingly adopt this new technology, which provides gaps in the market for new entrants. However, the reason for this is somewhat ironic, as explained by a quote from Jill Lepore discussing the work of Clayton Christensen: "[It] isn't because their executives made bad decisions but because they made good decisions, the same kind of good decisions that had made those companies successful for decades."

This is nicely illustrated by the downfall of BlackBerry. When the iPhone was first released, the makers of BlackBerry didn't think that the new phone would be a real competitor and therefore didn't take it seriously. The new product didn't appear to offer anything that BlackBerry's consumers wanted, so they disregarded the potential it held. It was only once it was too late that they were able to see what they had initially missed.

It's not just that companies are often unable to see opportunities that new technology can afford. There are also those who can see what is happening or what is on the horizon—yet they resist the change. They neglect to consider that while the change may be hard, the payoff could be even bigger.

CHAPTER ELEVEN

Do Machines Have Power?

"Machines don't have power, but when deployed, they can change who does." Machines follow instructions, doing what we tell them to. The reason it appears that machines are making decisions is because of automation, which codifies judgment into the algorithm.

There are limitations to the power or potential of such automation, though. For it to work, the pre-programmed judgment must be applicable in a large range of scenarios. Yet we live in a complex world, and predictions are imperfect. As such, the authors argue that automation is best in instances with narrow contexts and goals; in many instances, we will need to lean on human involvement, as well. They propose "human-machine collaboration."

Even though machines don't in themselves have power, they *can* create or reallocate it. Instead of a decision being made by a human in the moment, a machine can be coded by a human before the moment of decision-making and used at scale.

This distinction is important. Regarding machines used in hiring and firing, it's important to realize that "nobody ever lost a job to a robot. They lost their job because of the way someone decided to program a robot." The robot appears to be culpable in the case of a firing—or at least it's easier to blame the robot—because the human likely doesn't know who was the person responsible for deciding how the machine should be coded. "The problem with machine automation is that it obscures the ultimate person who is responsible for a decision."

CHAPTER TWELVE

Accumulating Power

In AI, the advantage often goes to first movers. One reason for this is because the more you use AI, the better it gets.

For AI to work well, you need to have good algorithms, but these algorithms improve when you have access to good input data and good feedback data. Those who move first are more likely to attract more users. More users will generate more feedback data, which will further improve predictions, creating a positive feedback loop and making it harder for competitors to catch up.

Fast feedback loops often lead to racing. Given the major advantages earned by first movers, many companies invest heavily to start turning the flywheel early on—even if the AI doesn't seem ready yet. The AI with the best predictions will be the one that is most widely used. "Once the flywheel starts to spin, the AI that had only a small advantage at the beginning can develop a large advantage over time."

Part Five:
How AI Disrupts

CHAPTER THIRTEEN

A Great Decoupling

The value of AI lies in its ability to improve decision-making. As previously explored, the primary inputs of decisions are prediction and judgment. The latter is about how much something matters, yet these judgments are often made on implicit, rather than explicit, predictions.

As humans, when we make decisions, we often fail to recognize that predictions and judgments are two separate inputs, as they both exist within our minds. AI offers the benefit of allowing us to decouple prediction from judgment, as the prediction is done by a machine. "Decoupling prediction and judgment creates opportunity. It means that who makes the decision is driven not by who does prediction and judgment best as a bundle, but who is best to provide judgment utilizing AI prediction."

The authors illustrate what this looks like via a decision tree. Each choice will yield a different outcome, and the prediction is the probability associated with each outcome occurring. But judgment is required for determining which choice to make, as it generates the payoffs associated with each outcome. A decision tree helps us more easily see how to separate predictions and judgments as two distinct inputs that factor into a decision.

We can see what this looks like in practice when we consider the real example of Michael Jordan at the height of his career in the NBA when he was recovering from a broken foot. Jordan and his coach were faced with the question of whether he should return to play before he had entirely recovered. If he did play, doctors believed there was a 10% chance of a career-ending injury. While a 90% chance that he would be OK seems high, it's not only the prediction that mattered—the payoff (or judgment) mattered, too.

Jordan had two choices with three possible outcomes: He could not play, play and be fine, or play and face a serious injury. When each decision is weighted, it is easier to see why the rewards associated with playing and potentially facing an injury outweigh the cost of not playing at all.

CHAPTER FOURTEEN

Thinking Probabilistically

Humans tend to think deterministically. Machines function probabilistically. "Thinking in bets requires you to recognize that predictions are uncertain, and to understand that the outcomes you experience are partly determined by luck."

We can see this by considering human-driven cars versus self-driving cars. When a human is behind the wheel, they are responsible for both the prediction and judgment. If an accident occurs, it's hard to know whether the former or the latter is at fault, but we tend to assume it's a prediction error.

Self-driving cars allow us to more easily measure prediction error; however, this requires quantifying judgment. Are we willing to program a car to continue driving if there's a 0.01% chance that it senses a human in its path? Or if there's a 0.001% or 0.0001% chance? Although this tradeoff always exists, it is made explicit when programming a self-driving car.

We must consider the likelihood of something occurring and make a choice accordingly. If the prediction is fairly precise, the decision is not that difficult. But we don't always have enough data to predict the likelihood at a high level of certainty. "This means that prediction and judgment are potentially subject to a chicken-and-egg problem, which in turn creates a barrier to adopting prediction machines and building new AI systems."

There are two primary ways that humans build judgment: Either we learn by researching and watching others, or we learn from firsthand experience. In the former, we anticipate possible outcomes based on what we've observed before. In the latter, we make a choice and see what happens. Both inform how we'll behave when faced with similar situations in the future.

Experience provides judgment. Thus, it can also lead to better decisions. But our choices determine what kind of experiences we have. We can think of the choices as existing in two camps. The first is the status quo action. This is doing what we've done before, which will lead to a known outcome. The second is the risky action—we've never done it before, so we don't know what will happen.

This creates a challenge to building AI systems. Even if AI can provide us with predictions, if we don't know what we'll do with that information, we may still end up going for the status quo action. "If the payoff of an AI system solution is high enough, [however,] it will be worth investing in building judgment, so this chicken-and-egg situation may not be hopeless."

There's another way we can encourage people to feel more comfortable existing in this probabilistic space: We can introduce regulatory processes—similar to the role the FDA plays in drug approvals—to help humans navigate it. They can serve "to reassure citizens that, despite some risks, the benefits [are] overall net positive."

CHAPTER FIFTEEN

The New Judges

Prediction machines can be useful in instances of uncertainty. An example of this is when the city of Flint, Michigan, was grappling with the discovery of lead in its water pipes. Researchers developed an AI tool that predicted where the lead would be; the challenge, however, was in the judgment, as that was political. Local politicians didn't want to use the prediction machine.

Decision-makers are those who provide the judgment. Because AI can influence who provides the judgment, that person may change with the adoption of AI. But those who hold decision-making power today are likely to resist having that power taken from them.

A number of factors determine who should or will be a decision-maker in business: efficiency, information, skill level, incentives, and impacts. In other words, who can make the best decision at the lowest cost (efficiency), who has access to the best information (information), or who has the skills required to make the best decision (skill level)? It could be that the person has interests that are best aligned with the organization (incentives), or that the person best understands the impacts associated with each option (impacts).

AI has a direct influence on this. Because AI creates a decoupling of prediction and judgment, and the machine's prediction is objective, judgment becomes the key determinant of whether someone should be the decision-maker. In some cases, this could lead to a reinforcement of the existing decision-maker; in others, it could lead to someone else becoming the decision-maker.

AI can also have an impact on the concentration of power. Because AI prediction can be used across a wide set of decisions, it enables the scaling of judgment. "This scale drives potential efficiencies from formulating judgment and coding it into automated processes." Judgment may be provided on a decision-by-decision basis, or it can be scaled for numerous decisions. In the latter case, the person who provides the judgment wields considerably more power.

Part Six:
Envisaging New Systems

CHAPTER SIXTEEN

Designing Reliable Systems

Inserting AI into a system has an impact on how we coordinate decisions. Organizations rely on rules to create reliability. Each person has an expectation of what others are doing and aligns their actions—or decisions—accordingly. Without this reliability, coordination across an organization would be much more challenging.

The authors illustrate this in more detail with the example of a restaurant. A menu creates reliability: It ensures that the cooks can provide what the patrons order. If the restaurant were to adopt AI for demand forecasting, it could have net positive effects for the restaurant but would also introduce unreliability for the supplier. These effects would cascade all the way up to growers.

"We call this effect—when implementing an AI improves the quality of one decision but harms others in the system by lowering reliability for other decisions—the AI bullwhip. Like a bullwhip, a small change in one place can lead to a big crack elsewhere."

There are two ways to address this challenge: We can lean on coordination or modularity. Coordination makes use of communication and synchronization among actors to decrease unreliability, while modularity is useful in instances where such alignment isn't possible. In this case, it focuses on insulating or compartmentalizing a decision so that its impacts on other actors are limited.

"Modularity insulates decisions from the variability that comes from AI. It reduces the importance of reliability. Coordination, in contrast, creates reliability directly. Successful AI systems enable coordination where possible and modularity where necessary."

The latter option offers the benefit of scalability and can be more resilient in complex situations. However, it is still important to consider whether the benefits of adopting AI in this siloed space are worth more than the potential costs in other areas.

But there's another way to address the challenges associated with a complex, interdependent system. You can introduce a "'digital twin,' or a virtual representation of a physical object or system." Digital twins allow you to simulate a variety of options and test them. They also allow you to simulate the impact of the idea by using AI to predict outcomes. Using a simulation in this way allows you to explore innovations at a considerably lower cost than actual deployment.

CHAPTER SEVENTEEN

The Blank Slate

The authors argue that industries like insurance don't fully appreciate the huge opportunity that AI could provide because it's outside their normal business models. Instead of using AI to help clients better understand their risks and to enable them to address these risks, many insurance companies largely focus on using AI to improve their underwriting predictions.

Because many companies simply haven't considered this opportunity, the authors suggest using the blank slate approach, or a tool that they call the "AI Systems Discovery Canvas." The tool is intended to help users uncover opportunities for AI that they may not otherwise consider. It can help expose rules that hide uncertainty, allowing the user to convert those rules into decisions. It can then help the user more easily consider what impact an AI solution could have on systems.

The canvas contains several components: mission, decision, prediction, and judgment. First, the user must identify the mission of their organization, then they must identify the key decisions needed to achieve it. Then, they must go a level deeper and consider what information is required to make a decision. This includes both information that is already available, as well as anything that could be important. "Predictions are what AI can potentially supply, so this exercise links predictions to the core decisions in your organization." For each decision, the user must outline the tradeoffs involved or possible errors and then rank them. "Your judgment is how you rank those errors."

The authors explore what applying the canvas might look like for a business in the home insurance industry. The mission might be to "provide homeowners with peace of mind against catastrophic loss of what is for many the most valuable asset they own." Decisions could be categorized in three areas: marketing, underwriting, and claims.

Take marketing, as an example. The decision is who to target with marketing, and the critical prediction for that decision is identifying the willingness-to-pay for potential customers. The related judgment presents the consequences of possible errors: "the cost of targeting someone who doesn't purchase vs. not targeting someone who would have purchased." For the underwriting decision, the prediction is identifying which customers are least likely to file a claim, and the judgment concerns the tradeoffs of setting prices according to the overarching strategy. For claims, the critical prediction is whether a claim is valid, and the judgment considers the risk of not paying someone with a legitimate claim versus the cost of covering an illegitimate one. By looking at the industry holistically, it may become easier to understand where there are possibilities for the application of AI.

"Most companies have created systems comprising so many interdependent rules, along with so much associated scaffolding to manage uncertainty, that it's difficult to think about how to undo parts of it and contemplate the new system design possibilities AI predictions afford." The canvas is one way to unlock new thinking and uncover new opportunities.

CHAPTER EIGHTEEN

Anticipating System Change

This chapter explores what system change could look like through the example of a hospital and the decisions surrounding a patient possibly suffering a heart attack.

Due to the life-or-death nature of a heart attack, a doctor's judgment is crucial: It informs what decision they will make regarding interventions, which can have serious outcomes for the patient. In some cases, the fear of liability or the incentives of being paid more may interfere, leading doctors to overtest.

First, the authors explore what an AI point solution might look like as an emergency medicine intervention. If AI can support diagnosis, the doctor may be better able to do their job. Yet the question remains whether the cost of implementing such a tool is worth the benefits it provides.

The authors then consider whether an application solution could better suit the doctor's needs. Instead of providing a prediction about whether the patient is having a heart attack based on the results of a test, perhaps the AI could help the doctor determine whether to test. But there is still another challenge: It's typically the administration that is responsible for allocating resources and purchasing new devices.

This leads the authors to suggest that there may be scope for a system change. They apply the AI Systems Discovery Canvas to emergency medicine to explore what this might look like. While the possibilities they uncover may not yet be possible, the exercise is useful for understanding what system change could look like.

“There are two broad classes of choices that define what we call a ‘system.’ The first is who sees what. The second is who decides what.” Who sees what is about information filtering—sometimes everyone has access to that information, and sometimes it stays within a single division of an organization.

Who decides what is about decision authority, but no single person can make every decision. Typically, there will be decision authorities in each division, meaning that there is not a single person who has information access and decision authority over the entire organization. Organizing it in this way creates efficiencies but also limitations for complex, interdependent organizations.

AI point and application solutions can work well within a division in an organization if the divisions are clearly delineated. But AI adoption is more challenging for organizations where this isn’t the case, as illustrated by the emergency medicine example. Making it work requires significantly changing the way the organization functions, and tradeoffs must be made—or a system-level change.

While the authors are convinced of the potential of AI, they conclude that its widespread adoption will likely require system changes, which will be disruptive and take time.

EPILOGUE

AI Bias and Systems

For AI to make predictions, it needs data. When that data comes from imperfect people—who are innately imperfect—the AI will inherit their imperfections.

This can be seen via two experiments conducted by Sendhil Mullainathan. In the first, he and his colleagues sent out fake résumés with different names intended to suggest applicants of different racial backgrounds. The researchers found that the résumés with white-sounding names received more callbacks than those with Black-sounding names.

In the second experiment, Mullainathan and his colleagues found that an algorithm used to identify health needs of patients was racially biased. Both humans and the algorithm exhibited discrimination, yet he argued that identifying and rectifying the bias in the algorithm is much more straightforward.

The algorithm may have been programmed by a human who intentionally or inadvertently created a system that perpetuated bias; however, it is easier to adjust the software than it is to reduce discrimination in a human's mind. In this way, the AI could actually serve to *reduce* discrimination rather than serving as a source of discrimination.

There is, of course, a caveat: This can only happen if the people who program and manage the AI want to fix it. And changing the system in this way can also change who holds decision-making power, which may lead to resistance.

Because of the way AI is discussed today, many of the people who oppose the adoption of AI tend to be those who are most concerned about discrimination. Over time, however, the authors believe it will be those who benefit from the perpetuation of bias who will be most resistant to AI adoption. As outlined in this epilogue, AI can be designed to reduce discrimination and continuously monitored to ensure that it continues to serve this function. This is a possibility that simply does not, nor cannot, exist with humans.

About World 50

Founded in 2004, World 50 consists of private peer communities that enable CEOs and C-level executives at globally respected organizations to discover better ideas, share valuable experiences and build relationships that make a lasting impact. The busiest officer-level executives and their most promising future leaders trust World 50 to facilitate collaboration, conversation and counsel on the topics most crucial to leading, transforming and growing modern enterprises. Membership is by invitation only.

World 50 communities serve every significant enterprise leadership role. World 50 members reside in more than 27 countries on six continents and are leaders at companies that average more than $30 billion in revenue.

World 50 is a private company that serves no other purpose than to accelerate the success of its members and their organizations. It is composed of highly curious associates who consider it a privilege to help leaders stay ahead.

Inquiries? +1 404 816 5559