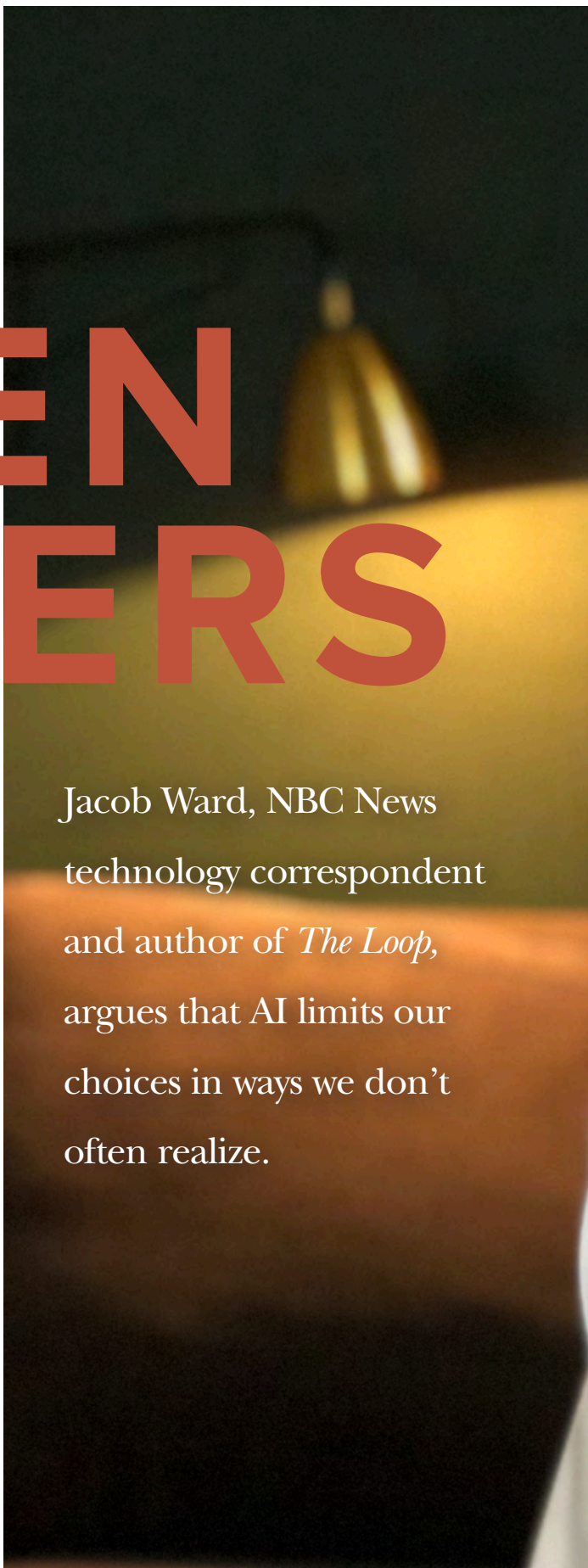




THE HIDDEN DANGERS OF DATA

In the HR field, there has been a more-is-better impulse around artificial intelligence and data analytics. After all, the technology promises quantifiable insights for a function that has often had to rely heavily on more qualitative judgments. But there are risks and limitations that must be taken into account.

David Reimer, the executive editor of *People + Strategy*, and Adam Bryant, articles editor, sat down with Jacob Ward, NBC News technology correspondent and author of *The Loop: How Technology Is Creating a World Without Choices and How to Fight Back* (Hachette Books, 2022), a provocative book on how artificial intelligence can limit our choices, to share his insights.



Jacob Ward, NBC News technology correspondent and author of *The Loop*, argues that AI limits our choices in ways we don't often realize.



People + Strategy: Please explain the concept of “the loop.”

Jacob Ward: I was seeing more behavioral psychologists being hired by companies to think about the deployment of AI to predict human behavior, human taste, human choices, and in the long run shape them. The loop I’m describing is the sampling of our unconscious tendencies by automated systems like AI.

That loop’s effect is to reduce choice in our lives because those choices are being handed to us by those automated systems. I worry that we are creating a world without much in the way of human agency or abilities to make choices on a social or an individual level, in the same way that Google Maps has made us less reliant on our sense of direction.

P+S: Part of understanding your framework requires understanding “system one” and “system two” thinking. Can you define those for us?

Ward: The last 50 years of behavioral science have centered on something called the dual-process theory of the mind, and that is that we essentially have two brains—cognitively, not physically. System one refers to the fast-thinking brain, and this is our oldest, most ancient circuitry. It’s probably at least 30 million years old, and this is our automatic detection system for things that are good and bad for us. We detect fire and snakes and strange tribesmen and calories using this automatic system. It is incredibly advantageous because it allows us to make choices without having to use any of the cognitive burden of making a decision. And it turns out that all kinds of behavioral and social stuff is built on top of that ancient circuitry.

System two thinking is much more recent, only about 70,000 years old, and in the words of one evolutionary biologist, it is a much more glitchy system. It is our system that allowed us as a species to stand up on what is now the continent of Africa, look around and say, “What else is out there and what happens when we die and why are my hands this way?” These are the kind of higher-order questions and musings that we now associate with being human. It is the system that allows us to be creative and cautious and rational and act against our own impulses, and to design systems of law and statistics and other inventions of modern humanity. This dual-process theory has been explored by people like Daniel Kahneman in his famous book *Thinking Fast and Slow*, and it turns out we make the vast majority of our decisions using system one, even though we like to think that we’re making decisions with system two.

And so we have built this huge, modern world of voting for people we’ve never met and trusting the police to make good decisions and all of these sort of artificial and modern conceits on the basis of our confidence that we use system two to make our decisions. But it turns out that system one is really still running the show even today, when we are in many ways allergic as a society to being told that we are not in control of our own decisions. That for me is one of the big problems

with the loop and the way that AI is working its way into our lives. Part of the circuitry in that dual-process theory of the brain keeps us from recognizing when we’re in the grip of system one, and as a result we’re very vulnerable to suggestions that play on it. And nothing plays on it better than AI.

P+S: What was the moment that triggered this book?

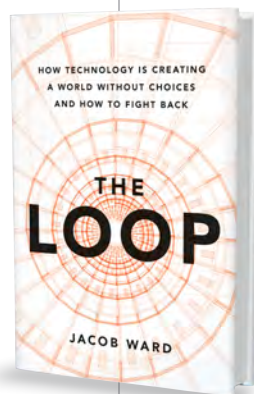
Ward: For three years, I worked on a PBS series called “Hacking Your Mind,” which was basically a kind of crash course in behavioral science. I was learning a lot about addiction, and then I had this experience of going to a dinner with a group of young tech entrepreneurs who would meet monthly to talk about the latest behavioral psychology papers and talk about the best ways to incorporate the findings of new behavioral science into the apps that they were building.

These apps were, for the most part, supposed to be socially redeeming—trying to help you save money or get into shape. That evening, they had invited two experts in addiction to talk to them. These speakers had set up a consultancy and they told the room about a famous experiment. The study focused on people who have been previously addicted to cocaine and tended to do it in nightclubs, but they’ve gotten sober. Then you bring them back to a nightclub, you thump music in their ears and you flash the lights at them and they had all the smells and external stimuli of a nightclub. Then you show them a mirror with lines of baking soda on it laid out like cocaine and you say, “This is baking soda,” and then you ask them how badly on a scale of one to 10 they want to do this baking soda, those people would report an overwhelming desire to snort a line of baking soda—even though their conscious mind, their system two, understands that it’s baking soda. Their system one is in charge.

The speakers told this story not to horrify us about the vulnerabilities of the human mind. They told this story to reinforce their pitch that human behavior is incredibly shapeable using what we know about addiction, and they were offering their expertise in addiction to make apps as addictive as possible. There was a whole discussion that night about whether this was ethical. Should we be doing this?

I was trying to be a journalist. I was trying to sort of stand back from it, but I just couldn’t help it. Because I’d just come from doing all kinds of reporting on the opioid epidemic, and they were being so cavalier about addiction and deploying their insights on people. I realized that it doesn’t take a very powerful command of system one, system two, dual-process theory to deploy it to capture people in the modern world using technology.

And that was the moment I started to realize that we are beginning to build stuff that is playing on really powerful forces in our brains without any ethical framework, much less a regulatory framework. The people who were using the insights, who understand the science in theory, had this very libertarian mindset about how just kind of disposable human



choice really is. They felt it was an infinite resource and therefore we didn't have to worry about mucking with it. And that for me got me started on the book.

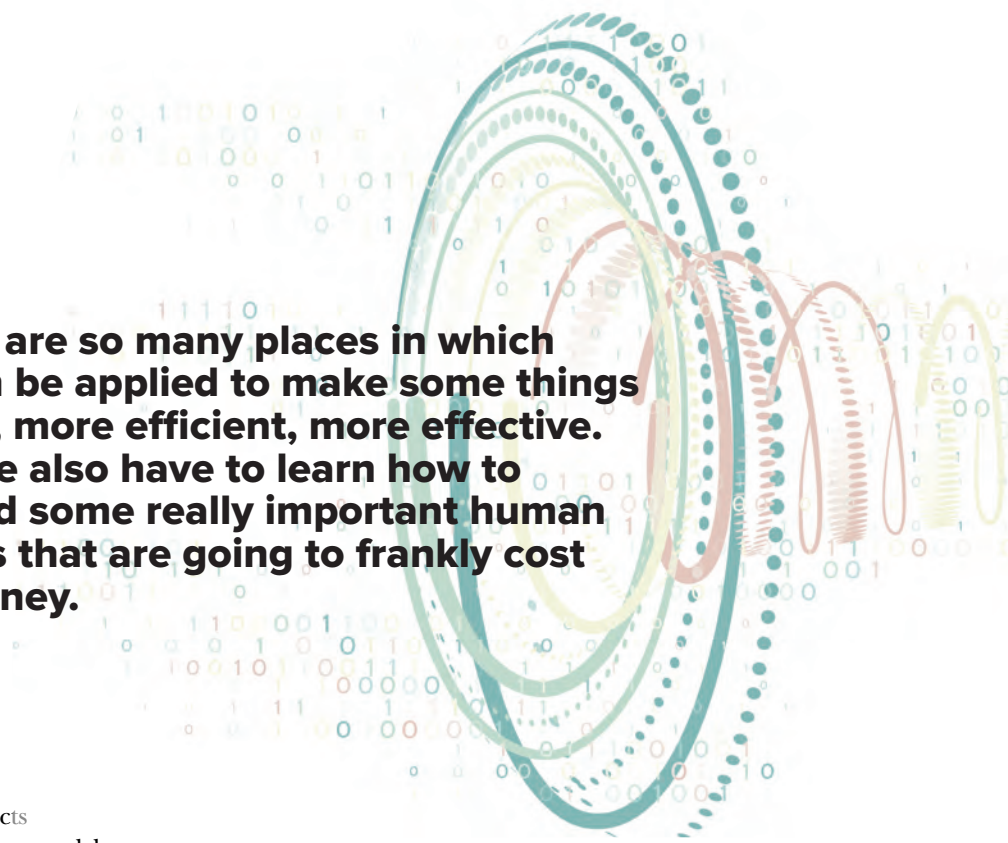
P+S: Your example at the consumer level is pretty stark. Have you studied how this plays out at an organizational level?

Ward: At an individual level, people are beginning to understand it in this more fundamental way. Everyone can wrap their minds around that feeling that your attention is not quite your own and that everybody's brain is kind of changing.

On an organizational level, it may be harder to articulate, but people are beginning to see the effects of automated decision-making based on an efficiency model. I'll give you an organizational example of the inability to withdraw from an automated decision-making process, even when someone influential inside an organization has decided you've got to do so. I've interviewed a leader who badly wanted to hire somebody into a newsroom who doesn't bring traditional news experience. He wanted an outside-the-box thinker—the logic being that we can teach them the fundamental skills of putting news on TV, but we want someone to bring an outside perspective. The leader explained that he'd found it almost impossible, short of physically marching over into the office and looking at resumes manually, to find a process that would help the organization realize this objective. Because the only resumes that would get greenlit and move past the hiring filters were the ones that had the name of a major news network at the top of them.

We're seeing this kind of thing over and over again. Our organizational frameworks have been built on premises of automation and the efficiency that it promises, and so it can feel that there's no way out of it.

One of the best examples of how a process like that can take over the minds of everybody involved is a famous case from 2017 of a United flight out of Chicago where four people were asked to volunteer to give up their seats because crew members needed to be on that flight. They had raised cash offers over and over but nobody would say yes. And finally they said that they were going to choose names and seats out of a computer for the people who would have to get off. They announced four names and three passengers left, but not the fourth one, a pulmonologist named David Dao. He refused to give up his seat because he had rounds in the morning to see his patients. And he was right. There is a legal framework to guard against such instances, but United didn't know that. So the story became famous because they brought in Chicago aviation police who dragged him off the plane. They bounced his head off a metal armrest and he had a concussion. It was a PR nightmare for

An abstract graphic featuring a background of binary code (0s and 1s) in various colors. Overlaid on this are several concentric, overlapping circles and loops in shades of teal, green, and red, creating a sense of motion and complexity.

There are so many places in which AI can be applied to make some things faster, more efficient, more effective. But we also have to learn how to defend some really important human values that are going to frankly cost us money.

the airline, and they settled out of court with him. The lesson was that because a computer had chosen the names, everybody involved—the gate agent, the manager on duty, the aviation authority, everybody—suddenly abandoned their faculties. They didn't then go sit and talk to David Dao or anyone else to say, "What would it take to have you give up your seat?" They just resorted to this computerized decision. And then everybody lost their minds to the point where they actually beat up a customer, a doctor trying to get to his patients.

P+S: And yet there seems to be a widespread push to adopt more analytics and data in the world of HR.

Ward: There are some brilliant and influential optimists about AI. Their thesis that AI is going to somehow make up for our inability to make consistently good decisions and is going to be this wonderful prosthetic decision-making technology. I think that's wrong because I don't think it takes into account the profit motive. There are so many places in which AI can be applied to make some things faster, more efficient, more effective. But we also have to learn how to defend some really important human values that are going to frankly cost us money. Here's an example. Up until about five years ago, every year in the United States about 30 to 60 children were accidentally run over and killed in their family's driveway because their parents couldn't see behind them as they were backing out.

That's a small number of people compared to the 30,000 to 45,000 Americans who die every year because of things like opioids and guns and car accidents. But Congress decided that even 30 to 60 children dying that way was too many, and they passed a set of rules requiring back-up cameras on all cars. It's the kind of accident that data-driven decisions would say is not really worth looking at compared to all the other causes

of death. But this bipartisan group of senators looked at that number and said that this is totally unacceptable as a society. And today, when you buy a new car in the United States, it comes with a backup camera. The point is that there are going to be moments when we decide that something should be done, even if the data might suggest otherwise. We have to carve out a language beyond spreadsheets to make it a variable in the decision-making process.

There has been an attitude within organizations that somehow handing off these decisions to these automated systems absolves us of responsibility for the choices made. But there are risks in that, and it's creating a practice area for lawyers to guard against the risks created by companies relying too much on AI systems for important decisions. Hiring is one of the biggest liability areas—AI systems are just as biased as everybody else, except the problem is that it's very difficult to detect in these sleek and opaque systems that we are deploying.

P+S: So if you were advising the head of HR at a large company about the smartest questions they should ask about their AI systems—or should ask of a vendor pitching them an AI application—to uncover those risks, what should they be?

Ward: The questions would be: What is the dataset the AI system was trained on? What other applications are trained on the same dataset? That question is important because these outside companies, for the most part, cannot afford to train a new piece of AI for your task. They're repurposing insights from other applications. So what are they?

Then the next question is, how auditable is it? How explainable is it? Another way you ask that is, how do we understand the inner workings of this system so that we can inoculate

ourselves against a lawsuit if it makes a terrible decision for our company? Is it a black box or can I actually see how it works? And the last question to ask is, what can it not do? A responsible vendor of these systems is going to be able to tell you the line past which it will not go.

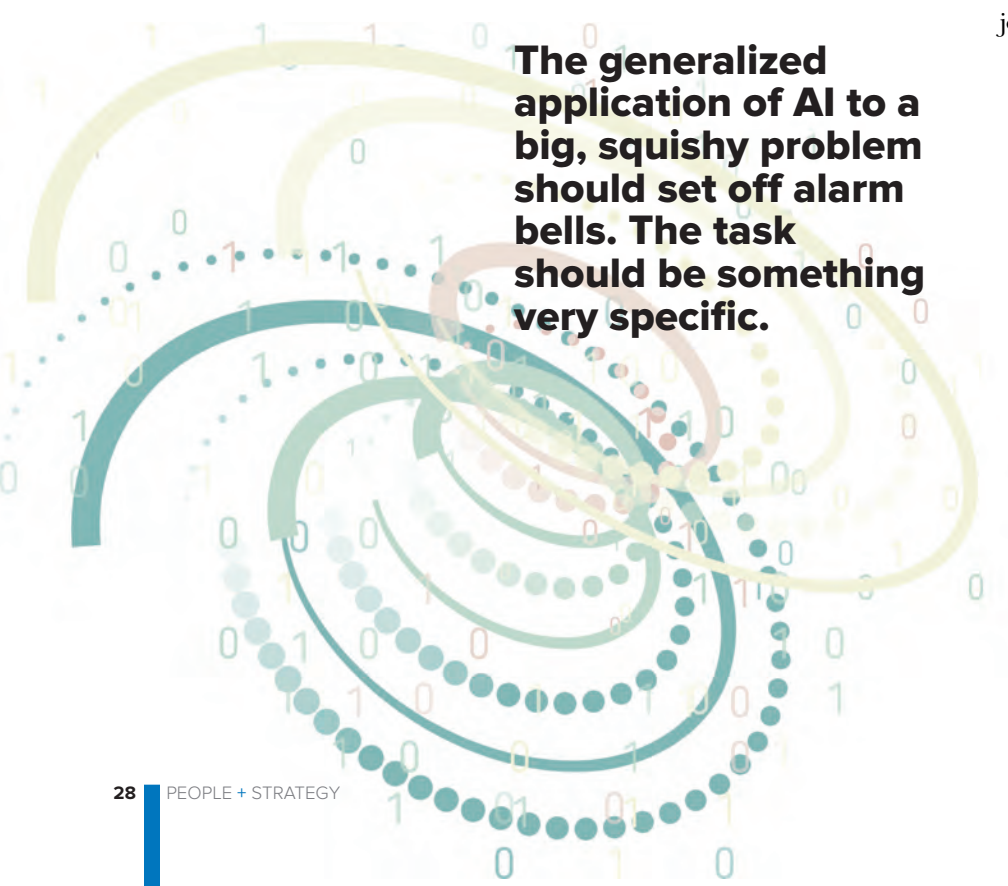
The point is that the generalized application of AI to a big, squishy problem should set off alarm bells. The task should be something very specific. AI essentially is good at telling the difference, for example, between dogs and cats in photos. It doesn't know anything about dogs and cats. It just uses the patterns it finds in 10,000 photographs of the two to do that. The problem with the way a lot of AI is built is that you don't get to see how it does what it does. It doesn't show its work. But it turns out that one of the most common ways that these systems discern between pictures of dogs and cats is that dogs tend to be photographed outside and cats tend to be photographed inside, so a green background was one of the variables that it relies on.

So again, the system is not very smart. It's just good at spotting patterns. And so you're going to have all kinds of people who make money off AI trying to convince you this thing can handle a broad set of things for you. Not only would I say that that is not technologically true, it's also important to remember that even the people who have studied AI in an academic setting all their lives cannot really agree on most of it anyway.

I went to a meeting once between a group of AI theorists and a group of behavioral and political scientists, and the AI theorists were debuting a system that they had created that they thought could somehow absorb universal human values from huge amounts of written prompts. After they talked for a while, a political scientist said, "I have three questions. What is universal? What is human? And what are values?"

As the head of any *Fortune* 50 company's HR department would admit, it is a squishy world. Human resources—which involves finding the right people for certain jobs—is also super squishy. So be tremendously wary of anybody that says that AI can, say, find your perfect candidate. What it might be able to do, however, is find common patterns in why things are going wrong. But something as all-inclusive as we will simply find you great candidates—nobody should be claiming that.

And there is so much disruption now, with unprecedented events, that it is making it harder for AI to be applied to a lot of the challenges we are facing now. AI does nothing except forecast the future from the past, and we have no precedent for so much of what we are living through right now. We are bringing more AI into our lives, because it's available, at a moment when things are less predictable than ever. I think that's going to lead a number of companies to make some weird and bad decisions using it. That's a tension that everybody in the field of HR and beyond is navigating. The past is not a good indicator of the present right now. ■■



The generalized application of AI to a big, squishy problem should set off alarm bells. The task should be something very specific.