SUMMER 2024 ▪ sloanreview.mit.edu

MIT Sloan Management Review

AI in the Wild

Investigating its impact in the real world

skillsoft®

SKILLSOFT'S LEADERSHIP DEVELOPMENT PROGRAM (SLDP) AND SKILLSOFT CAISY™ AI CONVERSATION SIMULATOR

Skillsoft offers transformative learning experiences that equip leaders to navigate unprecedented disruptions and rapid change. Our Leadership Development Program (SLDP) provides cutting-edge, multimodal learning, complemented by live coaching and AI-driven skill assessments. Dashboards and metrics enable assessing, indexing, and development tracking for individuals, teams, or the organization.

With *Skillsoft CAISY™ Conversation AI Simulator*, leaders can practice and refine their conversational skills in a safe environment, receiving instant personalized feedback and recommendations for improvement.

Beyond SLDP, Skillsoft offers customized programs tailored to your organization's needs, covering areas such as Technology Leadership, Leading through Change, New Manager Development, and more.

Learn more about the Skillsoft Leadership Development Program at skillsoft.com

SKILLSOFT LEADERSHIP DEVELOPMENT PROGRAM
POWERED BY
MITSloan Management Review

Know Your Own Strength

We all love a good "what went wrong" story. Analyses of corporate failures are a staple of the management genre, alerting leaders to hidden dangers and unexamined vulnerabilities. But studying our weaknesses and mitigating them only goes so far; it doesn't actually tell us how to succeed. Establishing a lasting advantage requires a deep understanding of our strengths and how to wield them.

That understanding is essential to achieving a fair return on the value that a company creates with its products and services, as Marco Bertini, Oded Koenigsberg, and Todd Snelgrove explain. Their new framework for value-based sales requires companies to go beyond simply quantifying their strengths (in terms of product benefits) to justify pricing — an approach that typically fizzles out over time. Instead, companies must gain insight into their strengths in context: Which of their strengths are meaningful to a specific customer? What are those strengths worth to *that* customer, in terms of productivity, uptime, additional revenue, brand reputation, and other benefits? In this context, knowing your strengths is the basis of building long-term customer relationships where value is cocreated between seller and buyer.

Gaining clarity about your organization's core strengths is also the starting point to gaining the strategic agility needed to quickly capitalize on new opportunities, according to Mark J. Greeven, Howard Yu, and Jialu Shan. They've noted a commonality among high-performing companies that they've tracked since 2021: These businesses have identified core capabilities, modularized them, and offered them to customers and partners, who in turn have built on those capabilities and penetrated new markets on behalf of the company. Achieving this does require a commitment to establishing a digital architecture that enables modularity, with a particular focus on developing and maintaining a suite of APIs through which partners can connect. But once accomplished, this approach offers a hedge against uncertainty, as it positions companies to respond rapidly to new customer demands and changing business conditions.

On a graver note, a sense of collective, shared strengths in an organization can also help teams persevere and grow in difficult times, including in the aftermath of traumatic events. In December, a gunman killed three faculty members at the Lee Business School at the University of Nevada, Las Vegas. The tragic incident moved one of their colleagues, Payal Sharma, to explore the question of whether, and how, leaders can help members of an organization move through and beyond such a horrific experience. Psychology research has established that individuals are capable of post-traumatic growth. That means they are not simply returning to the status quo ante, which cannot be recovered, but are attaining a new level of wisdom and healthy functioning.

Sharma explains what factors support this growth in individuals as they work through trauma and suggests ways that leaders can harness similar dynamics to help facilitate this growth in their organizations. In the process, she illuminates the specific ways in which traumatic events are typically so profoundly destabilizing — but also how, as we come to a new reckoning with what we cannot control in the world, we can emerge with a new appreciation for our strengths.

Elizabeth Heichler
Editorial Director, Magazine
MIT Sloan Management Review

Rubric

Contents

SUMMER 2024 | VOLUME 65, NO. 4

Cover illustration by Paul Garland

RADAR

FRONTIERS

FEATURES

COLUMNS

IN EVERY ISSUE

EDITOR'S NOTE:
Some articles in this issue were originally published online. They have been adapted for print.

MIT Sloan Management Review

Customer Service

General Help
sloanreview.mit.edu/help

Privacy Policy
sloanreview.mit.edu/privacy

Permission to Copy and Reprints
Find the articles(s) you're interested in at shop.sloanreview.mit.edu.

Subscriptions

Individuals
sloanreview.mit.edu/subscribe

Institutions
sloanreview.mit.edu/organization-subscription/

Contact

If unable to find solutions at the URLs above, email or call:
smr-help@mit.edu
+1 617-253-7170
mit-sloan@pelcro.zendesk.com

Mailing Address
MIT Sloan Management Review
MIT Sloan School of Management
One Main Street
9th floor, E90-9200
Cambridge, MA 02142
USA
sloanreview.mit.edu/contact

Sponsorship & Advertising

Direct all inquiries to smr-sales@mit.edu.

Radar

RESEARCH SNAPSHOT

The Long Tail of Social Media Influence

The top social media influencers command top dollar for their endorsements: Italian super-influencer Chiara Ferragni was said to have been paid 1 million euros ($1.07 million) to promote a holiday cake to her followers.

Marketers may assume that paying heavily to place their products into the exquisitely curated digital lives of Ferragni and her ilk is a winning social media strategy. Surely the influencers with the largest follower counts will deliver the best results, right? But new research indicates that social media power isn't all at the top — and marketers might see a better return on investment in paid influence campaigns if they recognize that.

In their November 2023 *Journal of Marketing* paper, "Revenue Generation Through Influencer Marketing," a team of scholars — Maximilian Beichert, Andreas Bayerl, Jacob Goldenberg, and Andreas Lanz — explored whether marketers should target influencers with a large number of followers or those with relatively few. They did so by weighing the results of influencers' social media posts on networks such as Instagram — including how many followers the posts reached, engagement levels, and revenue earned — against the cost of paid endorsements.

The researchers drew on secondary sales data of 1,881,533 purchases and the results of three field studies with hundreds of paid influencer endorsements. They found that influencers with lower follower counts (referred to as *nano-influencers*) delivered dramatically better per-follower returns on marketing investments than macro-influencers — that is, those with a high number of followers. Macro-influencers in the researchers' data set had 32 times more followers but generated only about four times as much revenue compared with nano-influencers.

The finding is not so counterintuitive once engagement is factored in. The researchers posited that macro-influencers would have less engagement with their followers, as measured in followers' likes, shares, and comments on posts and influencers' responses to followers' comments. To evaluate the level of engagement between influencers and followers, the researchers looked at influencers' five most recent social posts before the paid post. They found that follower engagement was an order of magnitude higher for nano-influencers than for macro-influencers.

Marketers who haven't already recognized this dynamic would do well to heed the researchers' findings and begin spreading their budgets more broadly to small-time influencers. As for Ferragni? She has bigger worries than a bunch of small-time influencers eating her lunch: She was fined 1 million euros by Italian authorities when it emerged that there was no basis to her claim that proceeds from sales of that Christmas cake would go to a children's hospital.

— *The* MIT SMR *Editors*

M. Beichert, A. Bayerl, J. Goldenberg, et al., "Revenue Generation Through Influencer Marketing," *Journal of Marketing*, published online Nov. 16, 2023.

Getty Images

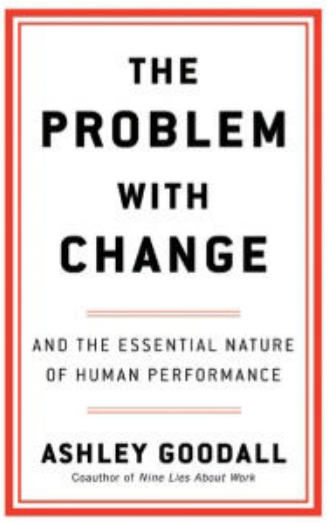
MIT SMR READS

The Problem With Change: And the Essential Nature of Human Performance *Ashley Goodall; Little, Brown Spark; May 2024; 288 pp.*

MANAGING CHANGE IS HARD — SO hard that change management is a cottage industry with books, courses, and hundreds of experts and consultants. After all, moving a complex organization to a new way of doing things is the kind of challenge that warrants plenty of study and discussion.

A new book asks whether any of it is truly necessary. The author, a former Cisco executive and Deloitte consultant, targets a default assumption held by many business leaders and consultants: that their chief responsibility is to effect change. At a time when the rapid march of technology and geopolitical strife threatens all hopes of stability, it seems like sacrilege to suggest that organizations should resist the impulse to continuously adapt.

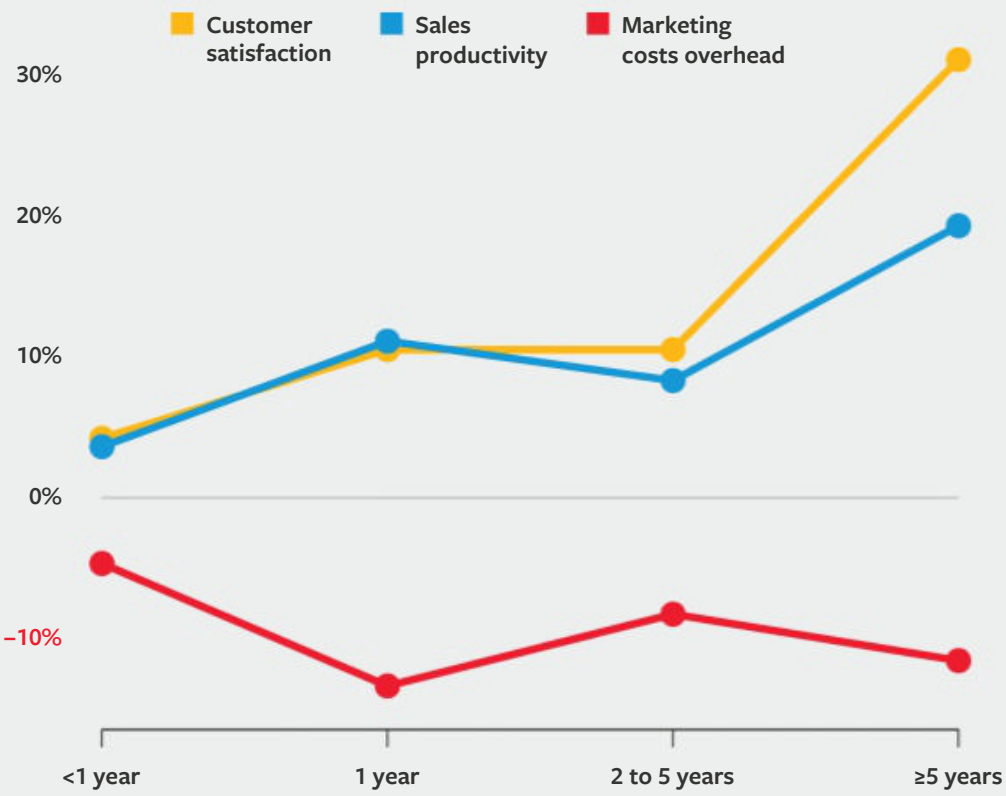
But Goodall argues that too often, change efforts are undertaken not in response to specific challenges but simply for the sake of change itself. The notion that such change might not always be in the best interest of the organization is rarely given the consideration it deserves. Goodall shows how needless change can unmoor and alienate people from their jobs and take a human toll that makes them less effective at achieving the goals leaders set for the organization.



The Problem With Change urges leaders to get back in touch with the individual, human experience of work that collectively forms any enterprise, and offers them a way to do so that employee surveys and KPIs simply cannot. Despite the book's exhortations against change, when it is necessary, readers may find this book may be the most grounded guide to change management yet.

— Kaushik Viswanath

HOW AI ADOPTION TIME INFLUENCES PAYOFFS



SOURCE: THE CMO SURVEY FALL 2023 EDITION

BY THE NUMBERS

How Long for AI to Pay Off?

SPECULATION ABOUNDS CONCERNING AI'S ULTIMATE IMPACT ON ORGANIzations and marketing, but it's tough to discern where companies are achieving results. For insights, the recent CMO Survey —the longest-running noncommercial survey for and about the field of marketing, conducted since 2008 — sampled more than 300 marketing leaders at for-profit U.S. companies. It found that the return on AI investment varies depending on adoption time, the company's digital transformation stage, and its level of experimentation with artificial intelligence. In an online article for *MIT Sloan Management Review*, Christine Moorman, a professor at Duke University's Fuqua School of Business and the founder and director of the CMO Survey, and Colleen Hickey, an MBA student at Fuqua and research fellow at the survey, wrote about their findings.

When respondents rated how the use of AI in marketing has affected outcomes, the results showed gains in three areas: a 6.2% increase in sales productivity, a 7.0% increase in customer satisfaction, and a 7.2% decrease in marketing overhead costs, on average. But the researchers found that returns don't follow a linear trajectory over time: Payoffs accelerate once companies gain a full year's experience deploying AI models, weaken from years 2 to 5, and accelerate again after five years.

Why? During the first full year, many companies deploy minimum viable products and demonstrate ROI to secure the budgets necessary to scale AI projects. Having established executive sponsorship and investment for expansion, they set up AI competencies, such as more data models for different use cases and machine learning operations capabilities, while monitoring and retraining AI models. These efforts flatten positive returns in years 2 through 5 — but payoffs eventually trend positive, doubling by year 5. Although gains might not be realized upon AI's initial adoption, the growth potential of these payoffs soars over time.

Read Moorman and Hickey's article "When AI Investments Pay Off in Marketing" at sloanreview.mit.edu/x/65408. *— The* MIT SMR *Editors*

THREE THINGS TO KNOW ABOUT

PROMPTING LLMs

THERE SEEMS TO BE NO ESCAPING large language models (LLMs) nowadays, and you certainly won't find reprieve in this publication. But if you're having trouble getting LLMs to give you the responses you're looking for, you might want to consider changing how you prompt them. The way that a prompt is structured has a significant impact on the quality of the response provided. Here are three research-based tips to help you improve your prompting strategies and get more out of LLMs.

1. Be polite. LLMs are just software, so it's surprising that the tone of your prompt should have any bearing on their output. But researchers at Waseda University and the Riken research institution in Japan found that across various languages, LLMs' performance on various tasks generally improves as the politeness of the prompt increases (though only up to a point). Being rude to an LLM tends to yield poor results. The researchers suggest that this is

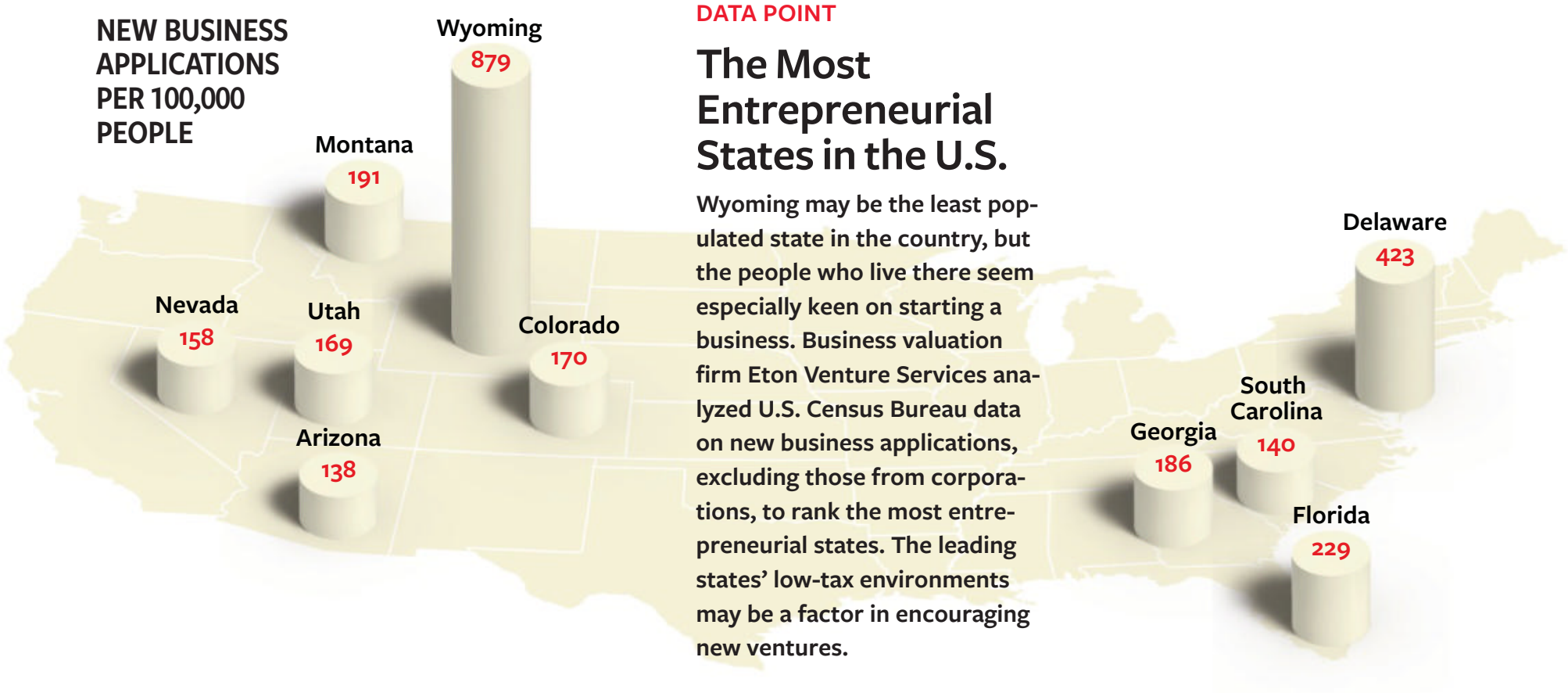
likely because the corpus of data LLMs are trained on shows humans responding better when we are polite with one another.

2. Give it context. Researchers at the University of Maryland found that you can reduce the extent to which LLMs hallucinate, or fabricate information, in response to queries by providing context in your prompt. An LLM that was asked to provide a list of academic publications by an author generally generated more accurate results when provided with that author's CV in the prompt than otherwise. It's worth noting, however, that even when provided with this context, LLMs were still likely to hallucinate some of the time.

3. Assign a role. In his book *Co-Intelligence*, Wharton professor Ethan Mollick writes that prompting the LLM to first assume the role of a subject-matter expert yields better and more specific results on a task than if you were to simply prompt it with the task. For example, before asking an LLM to generate taglines for a new product, try prefacing the request with "You are an expert in marketing."

As LLMs continue to evolve and researchers spend more time working with them, we will undoubtedly discover more and better ways to use them. The tips above should remind us, however, that LLMs are fickle tools and we should be careful about how we use them.

— Kaushik Viswanath



DATA POINT

The Most Entrepreneurial States in the U.S.

Wyoming may be the least populated state in the country, but the people who live there seem especially keen on starting a business. Business valuation firm Eton Venture Services analyzed U.S. Census Bureau data on new business applications, excluding those from corporations, to rank the most entrepreneurial states. The leading states' low-tax environments may be a factor in encouraging new ventures.

SOURCE: ETON VENTURE SERVICES AND THE U.S. CENSUS BUREAU

Getty Images (top)

MANAGEMENT IN PRACTICE

Keeping Innovation Alive at a Legacy Organization

Kevin Nolan, an engineer by training and a 28-year veteran of GE, is the president and CEO of GE Appliances. He became the company's CTO in 2006, 10 years before its acquisition by China-based multinational Haier Group. *MIT Sloan Management Review* spoke with Nolan about the evolution of the company and the sector, and what it takes to be an innovative company today. This interview has been edited for length and clarity.

What was your journey to becoming CEO of GE Appliances?

After studying engineering, I joined GE's Edison training program, where we made industrial circuit breakers. That made me very interested in manufacturing. So I worked at one of our manufacturing sites in North Carolina and did a lot of digitization and restructuring of the product. This was in the late '80s and early '90s, as manufacturing was being moved offshore, which made me wonder what the role of engineering was without manufacturing.

I saw outsourcing as producing a lot of short-term gains, but I saw that disconnect between manufacturing and engineering becoming a long-term problem. When I became the head of engineering at GE Appliances, I started working on reshoring manufacturing. The 2008 [financial] crisis drove home the need to do that as shipping became more expensive and we saw big market fluctuations. It also showed what we had lost in terms of homegrown manufacturing expertise.

Courtesy of GE Appliances

Appliances was sold to Haier. I spent the year following the acquisition learning about what Haier did and how they did it. I realized they were an innovative company and very different from anything I had ever seen growing up in GE. A year after the acquisition, they made me CEO of GE Appliances. I've also had a second role as global technology leader for Haier. It's been a great journey of transformation, taking some of the practices of management from Haier while keeping some of the good things that we did have at GE. And we've done a lot of reshoring and refocusing on consumers.

Have you seen different approaches to management at GE and Haier?

At GE, I was seen as a good technologist but not as someone who could run a company. Haier wanted someone in leadership who had a depth of understanding in tech. That's important because innovation is what drives profitability with this type of business. I think GE was too hung up on professional management; they would take a professional manager from one area and move them into the other, and their understanding of the core technology didn't really matter. To me, that is the fundamental mistake GE was making.

I'm very involved in product development and innovation. Over the past six years, we've been reinventing how we innovate. We started a place called FirstBuild, which is an open makerspace where we focus on cocreation with consumers. At FirstBuild, management has no control over what products are made or invested in. Only the community has a say.

That kind of focus comes out of some of the Haier thinking. Markets don't want companies to invest in R&D because most of that money gets wasted. But we now have an R&D innovation process that is profitable, and I believe somewhat recession-proof.

How do you inject fresh thinking into a legacy company like GE Appliances?

Haier enables what we call micro-enterprises. The company is not top-down at all. It's made of small business units that are market-focused, have autonomy, and are measured on their value creation. Each unit is responsible to the market. They need to listen to customers to form their own strategies and drive their own innovation.

This approach changes the dynamic of the leadership team. We figure out long-term goals and leave it up to the micro-enterprises to devise strategies that get us toward our goal. Most people in my seat want to control the organization, demonstrate their ability to drive a strategy, and make investment decisions. A lot of CEOs will say capital allocation is their job. People in those seats have been striving their whole careers to achieve that kind of control, so they're not going to give it up easily. ■

Frontiers

Aad Goudappel / theispot.com

GENERATIVE AI

The Hazards of Putting Ethics on Autopilot

Research shows that employees who are steered by digital nudges may lose some ethical competency. That has implications for how we use the new generation of AI assistants.

By Julian Friedland, David B. Balkin, and Kristian Ove R. Myrseth

The generative AI boom is unleashing its minions. Enterprise software vendors have rolled out legions of automated assistants that use large language model (LLM) technology, such as ChatGPT, to offer users helpful suggestions or to execute simple tasks. These so-called copilots and chatbots can increase productivity and automate tedious manual work. But if they are not thoughtfully implemented, they risk diminishing employees' decision-making competency, especially when ethics are at stake.

Our examination of the consequences of "nudging" techniques, used by companies to influence employees or customers to take certain actions, has implications for organizations adopting the new

generation of chatbots and automated assistants. Companies implementing generative AI agents are encouraged to tailor them to increase managerial control. Microsoft, which has made copilots available across its suite of productivity software, offers a tool that enterprises can customize, thus allowing them to more precisely steer employee behavior. Such tools will make it much easier for companies to essentially put nudging on steroids — and based on our research into the effects of nudging, that may over time diminish individuals' own willingness and capacity to reflect on the ethical dimension of their decisions.

AI-based nudges may be particularly persuasive, considering the emerging inclination among individuals to discount their own judgments in favor of what the technology suggests. At its most pronounced, this abdication of critical thinking can become a kind of techno-chauvinistic hubris, which discounts human cognition in favor of AI's more powerful computational capacities. That's why it will be particularly important to encourage employees to maintain a constructively critical perspective on AI output and for managers to pay attention to opportunities for what we call *ethical boosting* — behavioral interventions that utilize mindful reflection, as opposed to mindless reaction. This can help individuals grow in ethical competence, rather than allowing those cognitive skills to calcify.

Digital nudges, especially in the form of salient incentives and targets, can lead to subtle motivational displacement by obfuscating the ultimate aims of the team or organization and shifting proximal goals. When a performance measure becomes the main objective, it ceases to function as an effective measure, a phenomenon known as Goodhart's law. For example, copilots might be designed to nudge customer-facing workers to maintain five-star ratings by offering bonus points or financial rewards. But if workers focus entirely on increasing their ratings, rather than on delivering great customer service in the hopes of receiving a high rating, they may be tempted to game the system by misleading customers. In other words, the ratings may become goals in their own right, potentially at the cost of important qualities that are difficult to measure, such as honesty and trustworthy behavior.

The implications of nudging are particularly pernicious in ethically nuanced contexts that require self-awareness of the values we care most deeply about. By uncritically accepting AI copilot guidance, managers may neglect to consider the "why" underlying their decisions. In this article, we'll explain how that leads to the risk that their ethical competence may degrade over time — and what to do about it.

From Reactive Nudging to Reflective Boosting

Nudges tend to exploit what psychologist Daniel Kahneman dubbed *thinking fast*, a reactive mode that contrasts with *thinking slow*, that is, reflective thinking, as described in our 2023 paper "Beyond the Brave New Nudge: Activating Ethical Reflection Over Behavioral Reaction" in *Academy of Management Perspectives*. Such interventions can leverage mild financial incentives or emotional triggers, including joy, fear, empathy, social pressure, and reputational rewards, to induce individuals to act as they arguably should upon ethical reflection. Heavy reliance on these incentives can reactively shift attention toward the extrinsic reward, thereby supplanting and weakening the ethical motives they are intended to encourage. This shift occurs because moral maturity and autonomy are ultimately achieved through instilling good habits aimed at intrinsic — as opposed to extrinsic — rewards.

While nudging interventions can be effective when used carefully and sparingly, leading agents to increased self-awareness and autonomy, the power and pervasiveness of generative AI technology is ripe for overuse. Such overuse could instigate a nudge riot of motivational displacement and dependency, crowding out good habits of ethical reflection. It could also backfire by causing some employees to recoil from what they perceive as excessive paternalism or surveillance. Managers should take care to avoid setting up a virtual version of Aldous Huxley's *Brave New World* in which behavior is perpetually conditioned, via automatic cognitive responses, to follow what is lauded by the AI and its designers.

Though reliance on behavioral nudges cannot be entirely avoided, especially in processes involving risk management or regulatory compliance, the good news is that checking mechanisms can be introduced to keep humans mindfully engaged and to trigger ethical reflection before action. This can guard against the tendency of cognitive skills to atrophy from disuse. Given the many current limitations of LLMs, including tendencies to give biased and inaccurate information, as well as a lack of comprehension and logical coherence, managers should prioritize engagement triggers to keep people thinking critically about AI copilot output, even in the absence of ethical choices or nudges.

How can individuals develop their abilities to think reflectively about ethical choices and resist the easy default options that nudges present, not only in the workplace but also in their many interactions as consumers and citizens? We see promise in ethical boosting, which is rooted in a positive view of the human potential to learn and grow. Whereas nudging promotes reactivity and seeks to steer subjects to choose specific behaviors without much thought of their own, boosting is a long-term developmental exercise to encourage habits of mindfulness and reflection. Boosts could take the form of mental rules of thumb, or heuristics — such as the Golden Rule, the idea of the best for all concerned, and one's own virtuous self-image — that help individuals identify and think through ethical dilemmas.

Boosting principles could also target negative contingencies by correcting unhealthy workplace patterns via reminders at key inflection points. Here, even AI copilots can play a role, if they nudge people to think instead of to just click "accept"

on a recommendation. We found that the Microsoft copilot for its email tools was already fairly adept at warning of subtle, potentially offensive language in emails. But individuals can choose to exercise their brains by rewriting emails in their own words, for instance, rather than accepting the bland system recommendations. To boost such a mindset, messaging apps might invite users to take time before responding to a potentially rude or hostile message chain, thereby allowing tempers to cool and the more reflective mind to engage. An image of a person rage-typing might serve as an effective speed bump, helping users to build virtuous self-awareness. Likewise, training such as the Sirius Program, part of the Office of the Director of National Intelligence's Intelligence Advanced Research Projects Activity, aims to enhance cognitive skills such as gaining competency at recognizing one's biases and assumptions.

Ultimately, managers should be mindful of the rhetorical siren song underlying the generative AI branding as personal *copilots*, in contrast with *decision support* or *assistants*. While the latter terms acknowledge the technology is subservient to the user, *copilot* connotes a more capable, autonomous, and even responsible role. A copilot is fully qualified to fly the plane in a pilot's absence, after all; the implied cachet of competence subtly invites employees to trust in and abide by AI-driven nudges. If AI copilots enable greater managerial control and efficiency at the cost of declining ethical competence in the workforce, managers may want to consider installing some reflective speed bumps. ■

Julian Friedland *is an associate professor of business ethics at the College of Business, Metropolitan State University of Denver.* **David B. Balkin** *is professor emeritus of management at the Leeds School of Business in the University of Colorado Boulder.* **Kristian Ove R. Myrseth** *is a professor of management at the School for Business and Society at the University of York in England.*

Reprint 65410. For ordering information, see page 4.

CLIMATE POLICY ADVOCACY

Why Companies Need to Lobby for Climate Policy

Organizations that want to make real progress on sustainability need to build a business case for climate lobbying.

By Richard Roberts

In recent years, thousands of companies worldwide have made net-zero commitments to meet the urgent challenge of climate change. Many are investing in product and business model innovation or in supplier engagement programs to drive down emissions. But all too often, one of the most powerful tools a company has at its disposal is ignored by corporate leaders who are serious about sustainability: lobbying.

The huge social and environmental challenges we face will not be solved without effective public policy — and business has an indispensable role to play in making that happen. As Alberto Alemanno, founder of The Good Lobby, has argued, "The misalignment between what companies say and lobby for" is possibly the major factor preventing advances on major societal issues. Harnessing the political power of the thousands of companies that have committed to reducing their own greenhouse gas emissions to lobby for stronger climate policies has the potential to be transformative.

A CEO or CFO might agree in principle that their company should be lobbying for climate action, but allocating serious resources to the effort is where it gets tough. In 2022 alone, the oil and gas industry spent an estimated $124 million on lobbying the U.S. federal government, resulting in, for example, concessions to the industry in the 2022 Inflation Reduction Act. While exact figures aren't available, companies in other sectors with much to gain from a successful transition to net zero aren't investing anywhere near that kind of money in pro-climate action lobbying.

Why not? The World Resources Institute (WRI) has identified seven barriers to business leadership on climate policy, some of which are internal (related to organizational structure or technical capacity, for example) while others are external (the role of trade associations or the threat of a political backlash, for example). During a recent workshop with leaders from companies that already invest in some advocacy on sustainability issues, the Volans team that I'm part of asked participants to rank the barriers based on their experiences. The one that most companies ranked first was "competing priorities." Here's WRI's explanation of the problem:

Climate change is not a top advocacy priority for most companies. Many companies see climate policy as "not in our lane" if they are not major emitters, energy-intensive users, or clean energy producers. Instead, they prioritize their advocacy efforts around other issues (for example, tax or trade policy) considered to be more material. A sense of having limited political capital exacerbates this dynamic.

Creating the Business Case

The current scenario points to the need for a stronger business case for investing in advocacy and lobbying on climate and other sustainability issues — one that focuses on value and risk, not just virtue. The business case will look different for each company. However, based on dozens of conversations with companies at the forefront of this emergent field, the Volans team has identified three main value propositions that recur across different organizations in different sectors: growing the market for their products and services, converting the company's sustainability performance into a stronger source of competitive advantage, and mitigating systemic risks.

Which of these value propositions is most applicable at any given company depends on the nature of the products and services it sells, the quality of its environmental and social performance relative to peers', and its level of exposure to risks that cannot be diversified away. Companies that have begun to invest seriously in sustainability advocacy tend to fit into at least one of three categories: solution providers, industry leaders, and residual risk bearers.

Solution providers. These companies have a direct and obvious material interest in speeding up the transition to a sustainable economy, such as renewable energy developers and cleantech companies. Solution providers can use advocacy to grow the market for their products and services. For example, Vestas, the world's largest manufacturer of wind turbines, and Ørsted, the biggest offshore wind farm developer, lobby for policies that will accelerate the energy transition in key markets. Both companies have been vocal in their support for the European Union's Fit for 55 agenda and have been engaged in ironing out the details of key policies, such as the EU's Renewable Energy Directive. They also lobby governments to speed up the approval process for new wind farms — a prime example of solution providers backing reforms that serve both their own self-interest and the wider public interest.

Industry leaders. These companies are ahead of both regulatory requirements and most of their competitors in reducing their negative impacts and/or increasing their positive impacts on the environment. These companies can use advocacy to make their sustainability performance a stronger source of competitive advantage. For example, consumer goods companies like Unilever, Nestlé, and Ferrero supported the EU's regulation on deforestation-free supply chains, which came into force in June 2023. These companies have invested over many years in reducing exposure to deforestation-linked products in their supply chains. As a result, they recognize that rules designed to level the playing field will benefit themselves as well as the planet. They will face lower compliance costs than competitors that have failed to take meaningful steps to tackle deforestation and improve supply chain traceability.

Similar dynamics are at play in the automotive industry. Volvo Cars, for example, pivoted earlier and more aggressively toward producing electric vehicles (EVs) than most other carmakers. The company has backed up this strategic decision with advocacy in support of government-imposed end dates for sales of new internal combustion engine vehicles and opposition to exemptions for such vehicles that can run on e-fuels. The more successful its political advocacy on these topics, the more the bet on becoming a leading manufacturer of EVs will pay off.

Residual risk bearers. Finally, these companies or financial institutions are highly exposed to the costs of social and environmental systems degrading. They can use advocacy as a tool to mitigate systemic risks that cannot be adequately handled via diversification or divestment, such as climate threats to the global food supply and the increasing vulnerability of coastal cities worldwide. Consider the case of Aviva Investors, a U.K.-based investment management firm with more than $250 billion in assets under management. Aviva Investors has pioneered the concept of *macro stewardship*, which it defines as "engaging with regulators, governments, and other entities to 'change the rules of the game,' in favor of … businesses … providing solutions to sustainability problems or supporting the transition to a sustainable future." Aviva Investors recognizes that the likely impact of runaway climate change on investment portfolios cannot be diversified away and that it therefore has a duty to mitigate the systemic risk to its assets via other means. Global food and drink companies like Nestlé and Danone are also residual risk bearers in that their supply chains are highly exposed to physical climate risks. Supplier diversification can mitigate this risk only up to a point. That is why these companies advocate for policies to support a transition to more sustainable and resilient farming practices.

Individual companies might fall into more than one of these categories. Volvo Cars' pro-EV advocacy is about both strengthening its competitive advantage and growing the market for its products. The point is not to shoehorn companies into one or another of these categories but to emphasize that any investment in advocacy should align with at least one of the three value propositions described to go beyond just virtue signaling.

Notwithstanding the examples above, companies that invest in sustainability advocacy are the exception rather than the norm, and even leaders in this field have not yet developed particularly sophisticated ways of evaluating the return on their investments in advocacy. For companies

It takes a coalition — typically spanning both business and civil society — to achieve real policy change.

to step up their sustainability advocacy efforts, it is vital that they develop a robust business case for doing so. The We Mean Business Coalition's Responsible Policy Engagement Framework can be a useful place to start.

From Business Case to Strategy

Having a clear business rationale for investing in sustainability advocacy is only the first step toward developing a winning advocacy strategy. Once the rationale is clear, companies can set concrete objectives and time horizons. These should then inform the selection of issues, tactics, and partners — as well as how success will be measured and evaluated.

One framework that can be useful in this process is the policy funnel, originally developed by climate policy think tank E3G and subsequently adopted by a range of companies and nongovernmental organizations (NGOs):

1. First, an issue starts to surface in public consciousness.
2. Next comes public debate about what to do.
3. Then come concrete policy proposals.
4. Finally, a specific text is negotiated and agreed upon.

If an issue is near the start of the funnel, relevant tactics might include engaging with an audience through media and social media, sponsoring research, and deploying advertising campaigns. At the policy process and policy decisions stages, other interventions become relevant too, such as engaging directly with policy makers; activating key trade associations to do the same, with aligned messaging; and participating in consultations and policy advisory committees.

From Strategy to Execution

When it comes to implementing a winning advocacy strategy, two factors are important: building the right internal capabilities and cultivating relationships with the right external partners.

Internally, sustainability advocacy rarely sits neatly within a single department. It requires close collaboration among teams responsible for sustainability, government affairs, public relations, strategy, and finance. While responsibility for setting strategy should sit with a relatively tight-knit group, companies might also seek to create the conditions for employees at all levels across the business to get involved in amplifying advocacy messages. For example, cosmetics group Natura &Co educates employees on the importance of protecting the Amazon and how to discuss it with friends and colleagues, and it points them toward opportunities to make their voices heard by signing petitions and joining movements. For multinationals, it is essential to have champions in different regions who can translate your global advocacy goals into local campaigns and engagement.

As for the external dimension, advocacy wins are almost never the result of going it alone. It takes a coalition — typically spanning both business and civil society — to achieve real policy change. The archetypal example in the sustainability space is the We Mean Business Coalition, which played a critical role in helping to get the Paris Agreement over the line in 2015. The Business Coalition for a Global Plastics Treaty, which brings together NGOs and businesses from across the plastics value chain, aims to have a similar impact on ongoing negotiations over the details of a United Nations treaty to end plastic pollution.

People responsible for advocacy at large companies told us about the importance of forming partnerships with unlikely bedfellows. If a coalition is perceived as representing too narrow a spectrum of interests — a group of "usual suspects" — its message is more likely to be dismissed. Collaborations that are uncomfortable are often the most effective. The relationships that require work to build trust are often the most fruitful when it comes to doing collective advocacy.

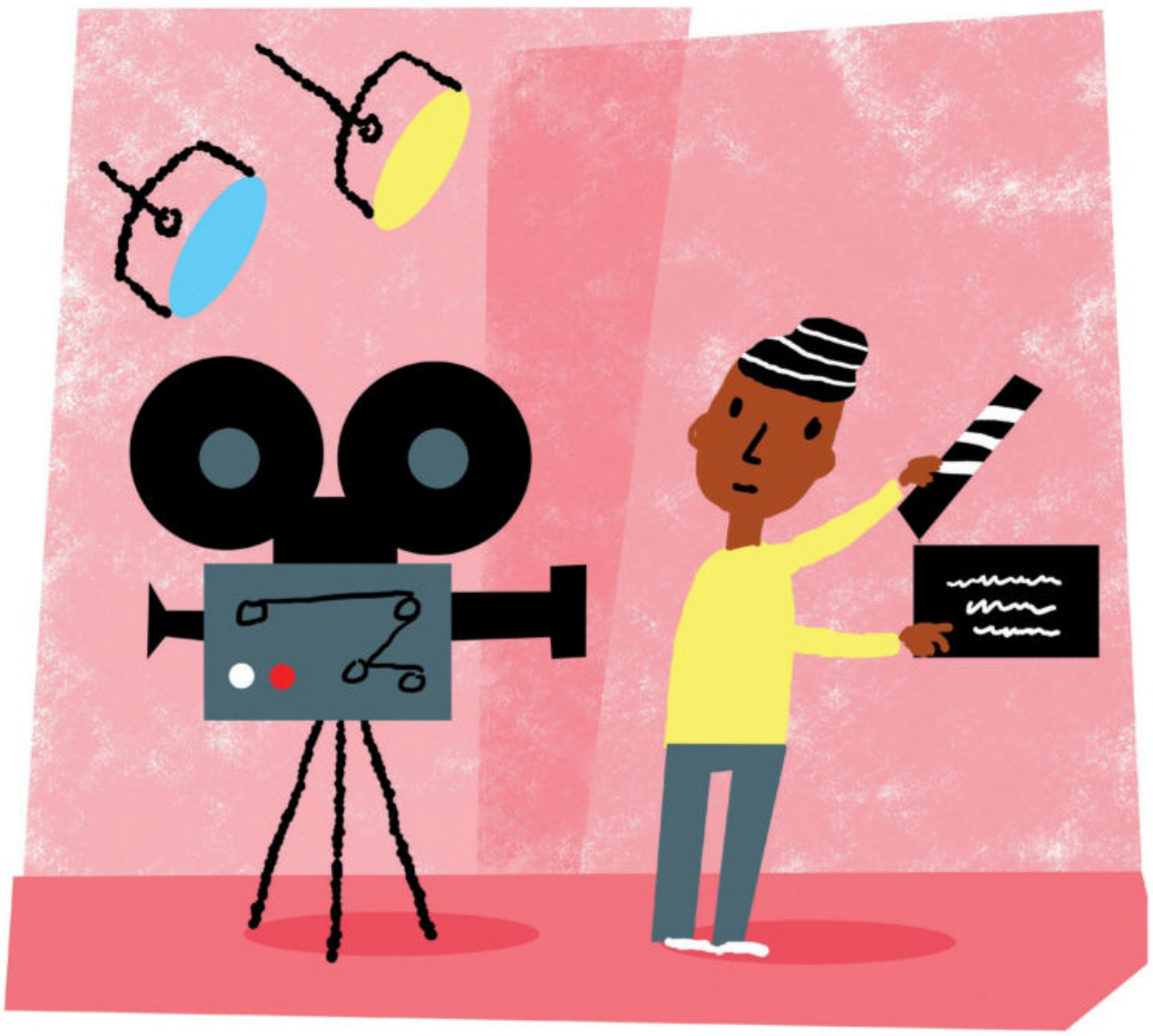
Brazil's Amazon soy moratorium is a case in point. In April 2006, Greenpeace activists targeted McDonald's to raise awareness of how soybean production was driving deforestation in the Amazon rainforest. Instead of going on the defensive, the company agreed to collaborate with Greenpeace to push for an industrywide solution, working in conjunction with the Brazilian government. The resulting moratorium was agreed on within four months and has remained in place ever since.

In summary, lobbying for supportive policies helps establish a virtuous cycle for companies that are serious about making a positive impact on the environment. When the ambition and effectiveness of government climate policies increase, they can enhance the value of existing and planned investments in reducing negative impacts and innovating to create products and services fit for a low-carbon future. The more favorable the policy context, the further companies can go toward achieving their climate objectives and setting more ambitious goals.

The Environmental Defense Fund describes political influence as "the most powerful tool companies have to fight climate change." We can't afford for them not to use it — and use it well. ■

Richard Roberts *is inquiry lead at Volans, a think tank and advisory firm focused on sustainability and innovation. He works as a strategic adviser and leads Volans's work on corporate political activity, which focuses on helping sustainable companies use their political influence responsibly and effectively.*

Reprint 65401. For ordering information, see page 4.



MARKETING

Tap Employee-Creators to Transform Your Social Media Strategy

Businesses that help employees become social media stars have a cost-effective way to generate enormous brand visibility.

By Aaron Dinin

On Jan. 2, 2023, 19 refugees from Cuba who were adrift on a makeshift boat in the Straits of Florida were spotted and rescued by the crew of the *Celebrity Beyond*, a passing cruise ship. Rescuing stranded refugees has become surprisingly routine for cruise liners during an intensifying global migrant crisis. What made this rescue unusual, however, was the way it rippled across social media when the captain of the *Beyond*, Kate McCue, documented it on TikTok.

McCue is well known in the industry for being the first American woman to captain a mega cruise ship (one that can accommodate more than 5,000 passengers). She's also known outside the industry as @captainkatemccue, a TikToker with 3.6 million followers (and another 900,000 or so followers on Instagram). McCue's 67-second video about the rescue generated nearly 30 million views on social media, creating the kind of positive exposure for Celebrity Cruises that any CMO would pay dearly for. Only in this case, it cost almost nothing.

For the sake of contextualizing that value, if Celebrity Cruises had wanted to advertise to the largest TV audience in the United States, the company would have had to spend approximately $882,000 for a 30-second spot during a Sunday night NFL game, watched by an average of 18.3 million people. That's 60% of the reach and half the engagement time of McCue's TikTok video. Even if we were to decide that a viral TikTok doesn't carry the same branding value as an NFL ad, McCue's millions of followers clearly bring amazing visibility to her employer at an almost negligible cost. It's enough to make one wonder whether every company should want its own Capt. McCue.

As a researcher who studies the emerging phenomenon of social media creators who produce and share content about their jobs and employers, I've spent the past few years talking to organizations about how they should be adapting to the rise of employee-creators. This article shares what I've learned about how savvy executives can tap into the vast potential of employees who amass large social media followings by encouraging, supporting, and even training willing team members on how to become successful employee-creators.

Creators With Two Careers

Employee-creators are employees who produce content on their personal social media accounts in ways that deeply entwine them with their employers. They're not just publishing content while at work, which plenty of creators do (often while trying to hide

James Yang / theispot.com

where they are employed). Instead, they have explicit or implicit support from their employer to post about their job and, as a result, publish content for themselves that simultaneously markets the company they work for.

The following are notable examples of employee-creators besides McCue:

- **Cody Rigsby, a cycling instructor for Peloton who has a following of 1.7 million people across platforms.** He has tens of thousands of dedicated fans who bought out the nearly dozen stops on his book tour, passionately cheered for him when he was a contestant on *Dancing With the Stars*, and, of course, keep paying their $59 monthly Peloton membership fee.
- **Dr. Austin Chiang, the chief medical officer of Medtronic Endoscopy and director of Thomas Jefferson University's endoscopy bariatric program.** He's a triple-board-certified gastroenterologist with more than 600,000 followers. He regularly posts informative and humorous content about colon and gut health. As one of the leading producers of medical devices, Medtronic likely gains a significant benefit from Chiang's content on gastrointestinal health and advocacy for procedures that protect against diseases like colon cancer.
- **Aaron Dinin at Duke University (that's me).** I am a faculty member in the Innovation & Entrepreneurship program at Duke University, who, four years ago, regularly began posting articles about startups on Medium.com and have amassed an audience of nearly 100,000 followers on that platform. I have expanded that work to another 150,000-plus followers across TikTok and Instagram, where I regularly post videos about my entrepreneurship classes and collaborations with students. In 2023, my content reached well over 20 million unique users.

Despite the increasing prevalence of such employee-creators and the value they generate, they are still uncommon. Many employers want to have total control over their brands and worry more about the risks that come with employee-creators than about missing out on their value. But even in companies that are open to employee-creators, they are rare for two reasons: Employees are often unsure of how or whether they should post about their jobs, and employers assume that they have to passively wait for employees to build large followings on their own. Both of these issues are resolvable with the right approach.

Create helpful guardrails so that more employees feel comfortable sharing online in ways that reflect positively on your brand.

Here are the seven steps I encourage businesses to take when they engage me to help them cultivate employee-creators within their organizations.

STEP 1: Facilitate creator training. Despite the often casual tone of the finished product, effective social media content creation is a time-consuming endeavor that requires practice, patience, persistence, and a willingness to experiment.

A company that is serious about leveraging employee-creators for its marketing efforts should speed up the process by hiring creator training professionals who are able to identify promising talent within organizations and teach fundamental principles for creating engaging content.

Hiring professionals to train your internal creator talent also helps ensure that the creators you nurture are well integrated into your brand from the beginning so that their content and growth more smoothly align with your brand's desired public-facing image.

STEP 2: Set clear content expectations. Because brands are both valuable and delicate, encouraging employees to create personal social media content that reflects on your organization's brand can be frightening. However, when organizations I work with express concerns about possible brand risks, I like to remind them that the risk of employees creating problematic social media content already exists. Proactively supporting and discussing your organization's content creation policies reduces that risk by allowing you to explicitly address topics such as:

- Expectations for the kind of content shared.
- How to divide time between content creation and other job duties.
- Mechanisms of creative oversight and approval (if any).
- Integration with internal marketing efforts.
- Ownership and usage rights of content.
- Protection of the company's intellectual property.

Remember that if you don't set firm guidelines and educate employees on your policies, the ones most likely to post content are the kinds of risk-tolerant people who are also more likely to post problematic content. In contrast, establishing clear policies and educating employees creates helpful guardrails so that more of them can feel comfortable sharing online in ways that will reflect positively on your brand.

STEP 3: Provide production resources. Effective content strategies rely on an array of production skills (such as videography, photography, editing, lighting, and audio engineering) and expensive equipment (cameras, microphones, lights, and production software).

Creators can learn production skills on their own, but companies can help by providing training on relevant technical skills as well as access to equipment, software, and facilities that can enhance and speed their outputs. In some respects, this kind

of training is similar to the kinds of learning benefits that companies already offer, and using these types of benefits to help employee-creators has a similar payoff of creating value for the company.

In addition, once employee-creators begin seeing meaningful audience growth, I encourage companies to consider hiring additional staff members (such as editors and producers) to help them scale production. It's a relatively low-cost way to quickly get more value from a uniquely effective marketing asset.

STEP 4: Nurture a supportive community. Despite the public visibility achieved by successful content creators, content creation is a challenging and isolating process. Creators who aren't well supported can burn out, which, in the case of employee-creators, can result in the company losing both a valuable marketing asset and an effective employee.

These outcomes can be avoided through proactive policies and structures that establish a supportive creator community. This includes giving sufficient time for employee-creators to produce content alongside their other job duties, ensuring that others in the organization know and understand why creator work is happening, and even having multiple people in the organization create content as part of a community.

In my early days of helping companies train employee-creators, I was often hired to coach a single executive. Progress for these individuals was so much slower than when I trained multiple people in an organization simultaneously that I have found this type of training to be worthwhile only when directed at a community of participants.

STEP 5: Define clear value metrics. Because cultivating audiences takes time, the value of emerging employee-creators won't be immediately apparent. Organizations should track the number and frequency of posts, audience growth, and content engagement to measure the impact of the initiative over time.

Even negative engagement can be a powerful strategy for reaching larger audiences and championing values aligned with the brand.

Companies should also find ways to measure the value of an employee-creator marketing strategy through metrics other than revenue growth. For example, Celebrity Cruises might track booking patterns related to the ship being captained by McCue.

STEP 6: Establish compelling incentives. The most common question I get from corporate leaders considering this strategy is, "What if we commit lots of resources to this and our employees become so successful on social media that they leave the company?"

This question is flawed. If an employee creates social media content promoting your organization for, say, five years and becomes successful enough to leave, it means that the employee created enormous value for five years. Every marketing strategy eventually ends, and an employee-creator strategy is no different.

Instead of focusing on losing employee-creators, I encourage the companies I work with to create incentives that reward success. Remember that successful employee-creators generate enormous brand value. They should be properly compensated and incentivized in the same ways organizations reward all great talent.

STEP 7: Prepare for success. When considering the strategy described in this article, remember that if the public profiles of specific employees are massively elevated at the kind of scale that's possible on social media, it can lead to unfamiliar and possibly uncomfortable outcomes. Effective employee-creators reach millions of people, and nothing that reaches that many people is ever without controversy. A look at the comments section of McCue's refugee rescue video demonstrates this challenge, as commenters with strong views on immigration, the cruise industry, and a host of other issues made their voices heard. Alongside the positive publicity, McCue and Celebrity Cruises also had to endure some criticism.

Still, on social media, even negative engagement can be a powerful strategy for reaching larger audiences, and responding to direct feedback can be an opportunity to champion values aligned with the brand. McCue provides an instructive example. One of her other most viral videos on TikTok is one in which she creatively deflects and devalues a sexist comment about her being a female ship captain. It demonstrates how McCue's enormous audience gives her — and, by extension, her employer — the power to speak out in ways traditional marketing and brand advertising don't allow.

McCue's powerful response to misogyny is a great reminder that, in the age of social media, employees represent more than just a source of labor. They are human embodiments of your company's brand and purpose. Expect to see more and more of your employees creating content about their work, regardless of whether you encourage them to do so. Choosing to partner with them by cultivating a community of employee-creators within your organization empowers your most valuable assets — the very employees doing the work — to become your brand's most influential and effective ambassadors. ■

Aaron Dinin, *Ph.D., is a faculty member in Duke University's Innovation & Entrepreneurship program, where he specializes in teaching marketing and personal brand building on social media.*

Reprint 65421. For ordering information, see page 4.

MIT MANAGEMENT
EXECUTIVE EDUCATION

COVER MORE GROUND

Today's leaders face an increasingly complicated set of challenges and contextual shifts that require a broad set of skills, depth of knowledge, and global perspectives.

At MIT Sloan Executive Education, our portfolio of **90+ courses** is constantly evolving to keep pace with the needs of today's leaders. Choose from a wide range of topics and course formats designed to help you put new methods and frameworks immediately to work, including:

Accelerating Digital Transformation with Algorithmic Business Thinking

Businesses for Inclusive Local Thriving Lab

Communication and Persuasion in the Digital Age

Creating High Velocity Organizations

Data-Driven Teams: The Art and Science of Winning

Digital Strategies for Transforming Your Business

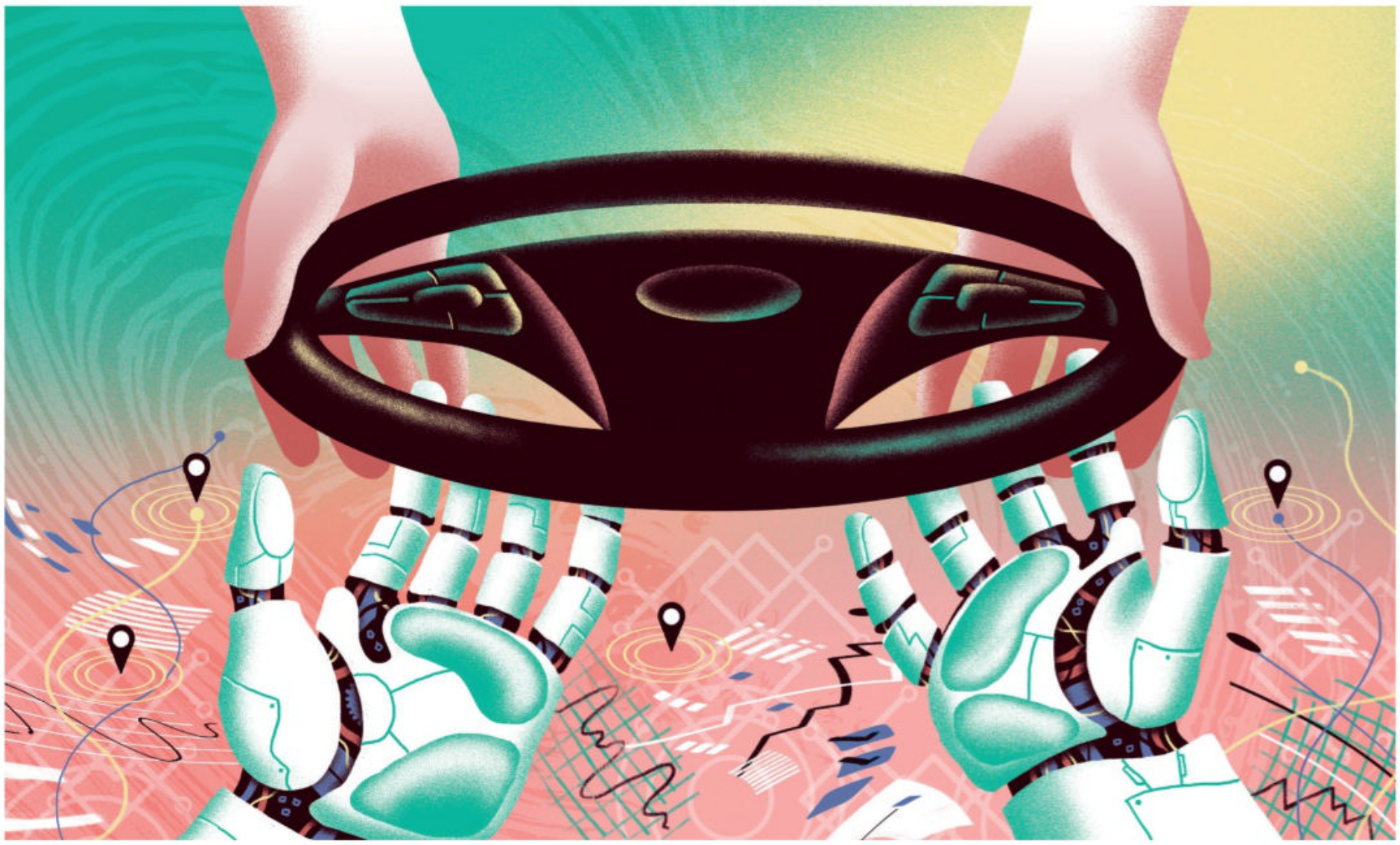
Leading the AI-Driven Organization

Maximizing Your Personal Productivity: How to Become an Efficient and Effective Executive

Navigating Transitions During Disruptive Change

Transforming Your Leadership Strategy

Short courses, Executive Certificates, Custom Programs, and Enterprise Accounts for organizations. **executive.mit.edu/smr**



RESPONSIBLE HUMANS

How AI Skews Our Sense of Responsibility

Research shows how using an AI-augmented system may affect humans' perception of their own agency and responsibility.

By Ryad Titah

As artificial intelligence plays an ever-larger role in automated systems and decision-making processes, the question of how it affects humans' sense of their own agency is becoming less theoretical — and more urgent. It's no surprise that humans often defer to automated decision recommendations, with exhortations to "trust the AI!" spurring user adoption in corporate settings. However, there's growing evidence that AI diminishes users' sense of responsibility for the consequences of those decisions.

This question is largely overlooked in current discussions about responsible AI. In reality, such practices are intended to manage legal and reputational risk — a limited view of responsibility, if we draw on German philosopher Hans Jonas's useful conceptualization. He defined three types of responsibility, but AI practice appears concerned with only two. The first is *legal responsibility*, wherein an individual or corporate entity is held responsible for repairing damage or compensating for losses, typically via civil law, and the second is *moral responsibility*, wherein individuals are held accountable via punishment, as in criminal law.

What we're most concerned about here is the third type, what Jonas called the *sense of responsibility*. It's what we mean when we speak admiringly of someone "acting responsibly." It entails critical thinking and predictive reflection on the purpose and possible consequences of one's actions, not only for oneself but for others. It's this sense of responsibility that AI and automated systems can alter.

To gain insight into how AI affects users' perceptions of their own responsibility and agency, we conducted several studies. Two

Matt Chinworth / theispot.com

studies examined what influences a driver's decision to regain control of a self-driving vehicle when the autonomous driving system is activated. In the first study, we found that the more individuals trust the autonomous system, the less likely they are to maintain situational awareness that would enable them to regain control of the vehicle in the event of a problem or incident. Even though respondents overall said they accepted responsibility when operating an autonomous vehicle, their sense of agency had no significant influence on their intention to regain control of the vehicle in the event of a problem or incident. On the basis of these findings, we might expect to find that a sizable proportion of users feel encouraged, in the presence of an automated system, to shun responsibility to intervene.

In the second study, conducted with the Société de l'Assurance Automobile du Québec, a government agency that administers the province's public auto insurance program, we were able to conduct more-refined analyses. We surveyed 1,897 drivers (mostly of Tesla and Mercedes cars with some level of autonomous driving capabilities) to look at the separate effect of each type of responsibility on the driver's intention to regain control of the vehicle and found that only the sense of responsibility had a significant effect. As in the first study, the more trust respondents reported having in the automated system, the lower their intention to regain control behind the wheel. It's particularly notable that only the proactive, individual sense of responsibility motivated respondents to act, which indicates that the threat of liability will be insufficient to prevent AI harm.

In another study aimed at understanding the use of risk-prediction algorithms in the context of criminal justice in the U.S., a significant proportion of the 32 respondents relied excessively on these tools to make their decisions. We made a determination of overuse based on respondents reporting use of the tools to determine the length or severity of a sentence, strictly abiding by the tools' results, and taking for granted the results provided by the tool. Besides raising fundamental legal and ethical questions about the fairness, equity, and transparency of such automated judicial decisions, this result also points to an abdication of individual responsibility in favor of algorithms.

Overall, these initial results confirm what has been observed in similar contexts, namely that individuals tend to lose their sense of agency in the presence of an intelligent system. When individuals feel less control — or that something else is in control — their sense of responsibility is likewise diminished.

In light of the above, we must ask whether *human in the loop*, which is increasingly understood as a best practice for responsible AI use, is an adequate safeguard. Instead, the question becomes: How do we encourage humans to accept they have proactive responsibility and exercise it?

As noted at the outset of this article, managers tend to exacerbate the problem by encouraging trust in AI in order to increase adoption. This message is often in terms that denigrate human cognition and decision-making as limited and biased compared with AI recommendations — despite the fact that all AI systems necessarily reflect human biases in data selection, specification, and so forth. This position assumes that every AI decision is correct and superior to the human decision and invites humans to disengage in favor of the AI.

To offset this tendency, we recommend shifting emphasis in employee communications from *trust* the AI to *understand* the AI to engender informed and conditional trust in the outputs of AI systems and processes. Managers need to educate users to understand the automated processes, decision points, and potential for errors or harm of the AI. It's also critical that users are aware of and understand nuances of the edge cases where AI systems can flounder and the risk of a bad automated decision is greatest.

Managers also need to position and prepare employees to own and exercise their sense of responsibility. Toyota famously models this by empowering anyone on the factory floor to shut down the production line if they see a problem. This will encourage employees to systematically question AI systems and processes and therefore maintain their sense of agency in order to avoid harmful consequences for their organizations.

Ultimately, a culture of responsibility — versus a culture of avoiding culpability — is always going to mean a healthier and more ethically robust organization. It will be even more important to foster such a culture in the age of AI by leaving clear spaces of possibility for human intelligence.

Otherwise, Roderick Seidenberg's prediction about technologies far less powerful than current AI could materialize:

> *The functioning of the system, aided increasingly by automation, acts — certainly by no malicious intent — as a channel of intelligence, by which the relatively minute individual contributions of engineers, inventors, and scientists eventually create a reservoir of established knowledge and procedure no individual can possibly encompass. Thus, man appears to have surpassed himself, and indeed in a certain sense he has. Paradoxically, however, in respect to the system as a whole, each individual becomes perforce an unthinking beneficiary — the mindless recipient of a socially established milieu. Hence we speak, not without justification, of a push-button civilization — a system of things devised by intelligence for the progressive elimination of intelligence!* ■

Ryad Titah *is associate professor and chair of the Academic Department of Information Technologies at HEC Montréal. The research in progress described in this article is being conducted with Zoubeir Tkiouat, Pierre-Majorique Léger, Nicolas Saunier, Philippe Doyon-Poulin, Sylvain Sénécal, and Chaïma Merbouh.*

Reprint 65427. For ordering information, see page 4.

STAKEHOLDER MANAGEMENT

To Navigate Conflict, Prioritize Dignity

Four interrelated practices can bolster dignity, leading to more constructive problem-solving and collaboration.

By Merrick Hoben

Conflicts between businesses pursuing commercial objectives and communities defending their interests arise regularly and often inevitably, especially when companies don't prioritize engagement with their neighbors. Consider the rapid expansion of the mining sector in Latin America, renewable energy projects that underestimate "not in my backyard" opposition, or the displacement of marginalized groups with unwanted facility siting. In many cases, the work has slogged on despite local protests, and drawn-out conflict has resulted.

Leaders inclined to think strategically and competitively may believe that stakeholder management in these cases is a matter of outmaneuvering the other party through gamesmanship, but that is shortsighted. When corporate interests conflict with the needs and values of communities, we need to build better interactions among people, especially those intensely at odds with one another who also need to collaborate.

This is where augmenting dignity can help. In my years of work in conflict resolution with the Consensus Building Institute (CBI), I've seen the benefit of putting the basic human need for dignity at the center of my efforts. By *dignity*, I mean our inherent sense of value, self-worth, and need to contribute and shape what matters most to us. Emphasizing dignity in conflict resolution doesn't displace the tactics and strategies of classical interest-based negotiation, either; rather, it precedes them, prioritizing what drives human behavior.

Donna Hicks, Roger Fisher, and Daniel Shapiro have written amply of the role of dignity and emotion in conflict resolution. Working in the field, I've found four interrelated practices that distill what's needed most to sustain such dignity, thereby helping people navigate effective problem-solving and healthier collaboration. These four practices are deepening acknowledgment, strengthening agency, building reciprocity, and ensuring clarity of path — acronymically forming an "AARC" that can help leaders across both commercial and public fields.

Deepening acknowledgment can set organizations on a better path to problem-solving. As Donna Hicks writes in *Dignity: Its Essential Role in Resolving Conflict* (Yale University Press, 2021), acknowledgment happens when you grant people "your full attention by listening, hearing, validating, and responding to their concerns, feelings, and experiences," creating a gateway to deeper engagement. Acknowledgment in this sense is not platitudinous empathy. Rather, it's about naming tensions and core issues with an appropriate emotional tenor.

Rafael Lopez / theispot.com

In the coal-mining region of Cesar, Colombia, a place torn by two generations of social and political violence, CBI once coproduced a documentary film to give voice to mining communities and a mining company, all of which were trying to move forward after years of conflict. The company was hoping to make a responsible exit from coal mining, and it wanted to establish a better relationship with the local community and beyond. But that same company had also been accused of human rights violations during years of civil war, and the community's pain, identity, and loss had not been meaningfully understood, heard, or considered. Tensions were high. The film explored commonalities and distinct differences, providing residents, the company, and other stakeholders with a shared focus to address past tensions and envision a way forward. Conversation and collaboration had seemed like remote possibilities, but now, the community and the company have begun to imagine and discuss ways to work together.

Strengthening agency helps parties understand and explore their own opportunities for shaping outcomes. As a facilitator in the Niger Delta, I've seen deep tensions between communities and energy companies take a constructive turn once community members understand their real influence over both process and outcomes. In Chevron's effort to improve community relations in the region, strengthened community agency proved crucial. Chevron's work had polluted the area and proceeded without sufficient participation from local communities.

Violent conflict followed, including kidnappings. Yet improved understanding became possible through the creation of regional development councils, which had representatives of all the major social groups in the area: youth and traditional chiefs, men and women, business leaders, and church representatives. Through direct participation, people were able to clarify and prioritize what mattered most in terms of their livelihoods, development needs,

Acknowledgment is about naming tensions and core issues with the right emotional tenor.

and well-being. Chevron took note. The resulting empowerment produced stable benefit-sharing agreements that have delivered results in terms of jobs, education, health care, and other community services.

Building reciprocity keeps momentum alive through practices that encourage constructive, give-and-take behavior. These practices include the demonstration of genuine curiosity, skillful active listening, and conscientiousness with respect to what's needed to understand the other side's perspective better. A surprising example emerged in the aftermath of Jan. 6, 2021, when CBI worked with Democratic and Republican members of the U.S. Congress to help them hear one another's perspectives on that day's events at the Capitol and find a way to work together despite sharp differences. By facilitating reciprocal listening, acknowledgment, and discussion of issues where there could be bipartisan agreement, the process enabled members of both parties to feel that they had been heard and could again work with the other side — without overlooking the tensions and political rancor — on issues of mutual interest. The resulting success of the Select Committee on the Modernization of Congress has been covered by *The Washington Post*.

Ensuring clarity of path means stating next steps clearly and addressing understandable anticipation from participants. In a recent hydropower conflict in Panama, I helped a major international development institution and indigenous groups rebuild confidence in one another after damage caused by a hydropower project. The project had led to flooding, including areas of local sacred sites, and indigenous communities had been inadequately involved in the planning stages. Specifically, indigenous representatives said they had been given insufficient data before the flooding.

To improve the relationship between the local communities and the development institution, I helped them cocreate a road map for remediating environmental impacts, community development, and social support. In the wake of flooded sacred territories, an AARC mindset helped participants recognize harms, frame joint issues to resolve, identify options for remediation, and, most significant, clarify how next steps would unfold. This approach — in a situation of deep power disparity, cultural divides, and mistrust — led to a precedent-setting acknowledgment of what had happened from the development institution and an overall remediation agreement.

The four AARC elements are essential ingredients in a recipe for dignity that should be modified for each context. And these elements, as in a good recipe, interact with one another. Deepening acknowledgment opens the possibility for joint problem-framing by finding spaces of shared agency. Once we grasp a shared problem, we can identify areas for constructive, reciprocal exchange, and those exchanges can lead to enhanced clarity about decision-making, roles, and responsibilities. When threaded together appropriately, these four elements support the potential for improved interactions, more durable organizations, and more effective problem-solving. In an era when fractures among people seem ever-widening, a bolstered sense of dignity can help bring colleagues and counterparts together constructively, in situations where you'd least expect it. ■

Merrick Hoben *is a senior mediator and director of the Consensus Building Institute's Washington, D.C., regional office. He coleads CBI's corporate-community engagement practice area.*

Reprint 65428. For ordering information, see page 4.

USING GENERATIVE AI

Nudge Users to Catch Generative AI Errors

Using large language models to generate text can save time but often results in unpredictable errors. Prompting users to review outputs can improve their quality.

By Renée Richardson Gosline, Yunhao Zhang, Haiwen Li, Paul Daugherty, Arnab D. Chakraborty, Philippe Roussiere, and Patrick Connolly

OpenAI's ChatGPT has generated excitement since its release in November 2022, but it has also created new challenges for managers. On the one hand, business leaders understand that they cannot afford to overlook the potential of generative AI large language models (LLMs). On the other hand, apprehensions surrounding issues such as bias, inaccuracy, and security breaches loom large, limiting trust in these models.

In such an environment, responsible approaches to using LLMs are critical to the safe adoption of generative AI. Consensus is building that humans must remain in the loop (a scenario in which human oversight and intervention places the algorithm in the role of a learning apprentice) and responsible AI principles must be codified. Without a proper understanding of AI models and their limitations, users could place too much trust in AI-generated content. Accessible and user-friendly interfaces like ChatGPT, in particular, can present errors with confidence while lacking transparency, warnings, or any communication of their own limitations to users. A more effective approach must assist users with identifying the parts of AI-generated content that require affirmative human choice, fact-checking, and scrutiny.

In a recent field experiment, we explored a way to assist users in this endeavor. We provided global business research professionals at Accenture with a tool developed at Accenture's Dock innovation center, designed to highlight potential errors and omissions in LLM content. We then measured the extent to which adding this layer of friction had the intended effect of reducing the likelihood of uncritical adoption of LLM content and bolstering the benefits of having humans in the loop.

The findings revealed that consciously adding some friction to the process of reviewing LLM-generated content can lead to increased accuracy — without significantly increasing the time required to complete the task. This has implications for how companies can deploy generative AI applications more responsibly.

Experiment With Friction

Friction has a bad name in the realm of digital customer experience, where companies strive to eliminate any roadblocks to satisfying user needs. But recent research suggests that organizations should embrace beneficial friction in AI systems to improve human decision-making. Our experiment set out to practically explore this hypothesis in the field by measuring the efficiency and accuracy trade-offs of adding targeted friction, or cognitive and procedural speed bumps, to LLM outputs in the form of error highlighting. We tested whether intentional structurally embedded resistance to the

Neil Webb / Ikon Images

uninterrupted and automatic application of AI would slow the user process and make potential errors more likely to be noticed. We believed that this would encourage participants to engage in what is referred to in behavioral economics as System 2 thinking, a more conscious and deliberative type of cognitive processing than the more intuitive System 1 thinking, akin to accuracy nudges in misinformation research.

The study, a collaborative effort between MIT and Accenture, aimed to explore the integration of an LLM into a task familiar to business research professionals. The objective was to complete and submit two executive summaries of company profiles (Task 1 and Task 2) within a 70-hour time frame and seek and reference any available sources, simulating real work conditions. The research participants were given text output from ChatGPT, along with the corresponding prompts, and were told that they could use as much or as little of the content as they saw fit.

Passages from the provided ChatGPT output and prompts were highlighted in different colors. Participants were informed that the highlighting features were part of a hypothetical tool Accenture could potentially develop and that the highlights conveyed different meanings depending on the color. Text highlighted in purple matched terms used in the prompt and terms in internal databases and publicly available information sources; text highlighted in orange indicated potentially untrue statements that should be considered for removal or replacement; text that was in the prompt but omitted in the output was indicated below the generated output and highlighted in blue; and text that was not identified as belonging to any of these categories was left unhighlighted.

Ideally, this hypothetical tool would combine natural language processing (NLP) techniques and an AI model to query all outputs against a predefined source of truth to highlight potential errors or omissions, but for the purposes of this experiment, the highlighting was done using a combination of algorithmic and human inputs. In addition, we purposely baked in some attention-check errors (nonhighlighted) to measure the circumstances under which adding friction in LLM use led to greater error detection (and improved accuracy) by participants.

Participants were randomly assigned to one of three experimental conditions, with varying levels of cognitive speed bumps in the form of highlighting:

- In the *full friction* condition, the LLM-generated content contained three kinds of highlighting based on the prompt that indicated that information was likely correct, incorrect, or missing from the output.
- In the *medium friction* condition, the LLM-generated content contained two kinds of highlighting based on the prompt that indicated likely errors and omissions of information that should have been in the output.
- In the *no friction* control condition, the LLM-generated content contained no highlighting at all, as per the current generative AI user experience.

Our findings revealed that introducing friction that nudges users to more carefully scrutinize LLM-generated text can help them catch inaccuracies and omissions. Participants in the no-highlight control condition missed more errors than those in either of the conditions with error labeling (31% more in Task 1 and 10% more in Task 2). Moreover, the proportion of omissions detected was 17% in the no-highlight condition but 48% in the full-highlight condition and 54% in the error-highlight condition.

As anticipated, these improvements did come with a trade-off: Participants in the full-highlight group saw a statistically significant increase (an average of 43% and 61% in Tasks 1 and 2, respectively) in the time required to complete the tasks versus the control group. However, in the error-only highlight condition, the average difference in the time taken versus the control was not statistically significant. Considering that each task typically took one to two hours on average without the assistance of generative AI, this trade-off was considered acceptable. Thus, the second condition, which involved medium friction, demonstrated a way to optimize the balance between accuracy and efficiency.

Three Behavioral Insights

The results of our field experiment point to actions organizations can take to help employees more effectively incorporate generative AI tools into their work and be more likely to recognize potential errors and biases.

Ensure thoughtfulness in crafting the prompt — a touch point for beneficial friction — given users' tendency toward cognitive anchoring on generative AI output. Participants' final submissions were lexically very similar to the LLM-generated content (60% to 80% identical content, as measured by NLP similarity scores). This suggests that the participants anchored on that output, even when they were asked to consider it as merely an input to their own writing. This underscores the importance of being thoughtful about the prompt provided to the LLM, since its output can set the trajectory for the final version of the content. Recent research suggests that anchoring may prove beneficial under some circumstances when generative AI content is perceived as high in quality and can play a compensatory role for an error-prone writer. But, given our findings of high similarity between the LLM-generated text and the final submissions from human participants, it could also lead a user down the wrong path.

Recognize that confidence is a virtue but overconfidence is a vice. Highlighting errors did indeed draw participants' attention and improved accuracy via error correction. Yet participants across the three conditions self-reported virtually no difference in response to the follow-up survey item "I am more aware of the types of errors to look for when using GenAI." This presents a reason to be cautious: Users may overestimate their ability to identify AI-generated errors. A tool that adds

friction by making potential errors more conspicuous could help users calibrate their trust in generative AI content by mitigating overconfidence.

Additionally, our findings suggest that highlighting errors had no significant impact on participants' self-reported trust in LLM tools or their willingness to use them.

Experiment, experiment, experiment. Before AI tools and models are deployed, it is imperative to test how humans interact with them and how they impact accuracy, speed, and trust. As indicated above, we observed a difference in self-reported attitudes and actual error detection. We urge organizations to adopt experiments as a means of understanding how best to elevate the role of employees in human-in-the-loop systems and to measure the impact on their understanding, behaviors, and biases.

THE EASE OF USE AND BROAD AVAILAbility of LLMs has enabled their rapid spread through many organizations, even as issues with their accuracy remain unresolved. We must seek ways to enhance humans' ability to improve accuracy and efficiency when working with AI-generated outputs. Our study suggests that humans in the loop can play an important interventional role in AI-enabled systems and that beneficial friction can nudge users to exercise their responsibility for the quality of their organization's content. ■

Renée Richardson Gosline *is head of the Human-First AI Group at MIT's Initiative on the Digital Economy and a senior lecturer and research scientist at the MIT Sloan School of Management.* **Yunhao Zhang** *is a postdoctoral fellow at the Psychology of Technology Institute.* **Haiwen Li** *is a doctoral candidate at the MIT Institute for Data, Systems, and Society.* **Paul Daugherty** *is chief technology and innovation officer at Accenture.* **Arnab D. Chakraborty** *is the global responsible AI lead and a senior managing director at Accenture.* **Philippe Roussiere** *is global lead, Paris, for research innovation and AI at Accenture.* **Patrick Connolly** *is global responsible AI/generative AI research manager at Accenture Research, Dublin.*

Reprint 65431. For ordering information, see page 4.

EMPLOYEE INNOVATION

Why Territorial Managers Stifle Innovation — and What to Do About It

Managers who feel insecure about their status tend not to encourage novel ideas from their employees. Fostering their identification with the organization can change this behavior.

By Vijaya Venkataramani, Rellie Derfler-Rozin, Xin Liu, and Jih-Yu Mao

Gary Waters / Ikon Images

Dilip has been working hard on a game-changing idea for a new product. But when he excitedly approaches his boss to share it and get approval for further development and testing, his proposal is quickly rejected. Instead of offering constructive feedback on how to make the idea workable, his boss vaguely refers to a lack of budget and discourages Dilip from pursuing any further ideas.

Sound familiar? Sadly, such scenarios are all too common in many organizations.

Employee creativity and innovation are critical to the success of organizations today. However, when employees do generate novel ideas, they often fail to receive encouragement or see their ideas materialize. Managers are a significant contributor to this phenomenon; even when they profess to value creativity, they routinely reject innovative ideas proposed by employees, preventing their implementation. Why do managers say no to ideas that could benefit their companies and even themselves?

Researchers have so far focused on personality factors, managers' economic mindsets, or managers' general aversion to uncertainty as explanations for this stifling of employee ideas. For example, one study has suggested that because managers are always focused on the financial consequences of their decisions, they reject novel ideas whose financial outcomes cannot be reliably forecast. However, as we discuss in our paper in *Organization Science*, there are deeper reasons for such rejection, rooted in managers' self-interests, underlying fears, and insecurities.

Managers have typically ascended to their positions by honing their expertise and finding success within specific task domains. They may have pioneered successful new ideas that have contributed to their professional reputations. As a result, it is natural for them to feel a sense of ownership and attachment to these domains, viewing them as their own territories. Being identified with a domain — whether a field of expertise or an area of responsibility — is crucial within organizations, as it grants the individual credible access to resources, decision-making authority, and the final say in domain-related conflicts.

However, this strong attachment to one's domain can manifest in territorial behavior as well. When subordinates propose ideas for new products or processes that can disrupt the status quo, it can trigger feelings of insecurity among managers who fear losing control or being overshadowed by their subordinates. This can prompt defensive territorial actions aimed at safeguarding what they perceive as rightfully theirs.

However, not all managers respond defensively with a "This is my turf — hands off!" attitude. Our research indicates that it is middle managers with low informal social status at work who are particularly susceptible to such insecurities. These are managers who are largely not well respected or held in high regard by their peers (typically measured by asking peers to rate the extent of respect and admiration they feel toward other managers). This underscores the importance of recognizing how these middle managers perceive their status in the workplace, given that our research suggests that there are significant implications and benefits associated with understanding their experiences.

Mind Managers' Power-Status Gaps

A manager requires both formal authority and informal influence within an organization to effect change. Formal authority is bestowed by their title and their place in the organizational chart, which grants them the power to make decisions in a particular domain, such as how to allocate resources,

define roles, and delegate tasks to team members. However, unless they also have informal influence accorded to them by their social status, they will find it difficult to effect change at a wider organizational level. Social status within the organization is defined by how people are perceived and how much they are respected and admired by their peers. Factors such as their performance, prosocial behavior, and the visibility of their successes or failures can affect their standing, as can factors that are outside their control, such as organizational culture, politics, discrimination, and bias.

Regardless of the reasons, because status serves as a marker of one's social worth in the workplace, it profoundly affects individuals' self-esteem, shapes their perception of workplace dynamics, and influences their interactions with others. Since managers with lower levels of social status or respect from their peers become highly sensitive to perceived encroachments into what they consider their territory, they become apprehensive about their employees' innovative ideas potentially infringing upon their domain. This apprehension prompts territorial behavior, as managers withhold support and resources to prevent the advancement of these ideas.

We tested our hypothesis in a series of three experimental studies and one field-based study. In our field study, we recruited work groups of engineers, programmers, and developers from four technology companies in China. First, we asked employees in these organizations to come up with a detailed work-related idea for a new service or product. Each of these ideas was anonymized and given to subject-matter experts, who assessed their novelty. We then provided each idea to 131 managers of the respective teams and asked them to report their own levels of territoriality as well as their willingness to sponsor and endorse ideas. Before this, managers had all rated their peers on their status within their organization.

The results of this study showed that lower-status managers were especially less likely to endorse an employee's idea when it was really novel (as rated by the subject-matter experts). We replicated these findings using different experimental designs that helped in establishing causality — that it is a manager's low status that leads to their territoriality and lower endorsement of novel ideas.

These results indicate that it is important for companies to recognize that how managers feel treated by their peers can have important repercussions on their everyday decisions as key influencers and gatekeepers within the organization's innovation ecosystem. When they feel appreciated and empowered by their peers, they are more inclined to adopt an open and inclusive leadership style that encourages experimentation, risk-taking, and the exploration of new ideas. This, in turn, creates a safe space for employees to voice their innovative ideas with the hope that they will be seriously considered and potentially implemented. On the other hand, if managers experience feelings of exclusion, disrespect, or hostility from their peers, it can lead to insecurity and heightened feelings of threat. This negative dynamic can stifle collaboration, impede knowledge sharing, and discourage employees from contributing their creative insights, thereby inhibiting innovation across the organization.

By acknowledging the importance of how managers feel treated by their peers, organizations can take proactive steps to foster a supportive and collaborative peer environment. This may include implementing training programs to enhance interpersonal skills, promoting a culture of mutual respect and appreciation, and providing resources and support to help managers navigate challenging interpersonal dynamics.

Support Managers as Idea Facilitators

Successful innovation often stems not from solitary geniuses experiencing breakthrough moments but rather from a social process involving collaboration among employees, colleagues, and managers. Middle managers, in particular, play a crucial role as enablers of innovation by identifying promising ideas from employees and offering support for their development. However, many middle managers struggle to transition from viewing themselves primarily as idea generators and innovators to embracing their role as a facilitator of their employees' creative efforts. This might be because they have advanced in their careers by demonstrating their own innovative thinking and problem-solving abilities.

Traditional leadership paradigms have long emphasized individual achievement and innovation, leading middle managers to believe that their value and identity are linked to their ability to generate groundbreaking ideas. This belief is often reinforced by performance metrics and evaluation criteria that prioritize individual contributions over facilitative leadership. Additionally, they may worry that adopting a facilitative role could be perceived as a sign of weakness or incompetence, especially in an environment where individual achievements are highly prized.

A senior manager in a company that one of us consulted with said, "Being a manager used to be about conquering goals myself, closing deals, that kind of thing. Now, it's all about supporting my team. Don't get me wrong; seeing them succeed is incredibly rewarding. But there's a part of me that misses the thrill of that personal win and sense of individual accomplishment. I almost feel I am not viewed as capable of that anymore." Addressing these factors requires a concerted effort from organizational leaders to redefine expectations, provide appropriate training and support, and create a culture that values and rewards collaborative innovation. But taking such steps can encourage managers who feel vulnerable due to their lower status among peers to see the value of promoting their employees' innovations.

Organizations should prioritize rewarding managers for supporting employees' creative ideas, because fostering a culture

of innovation requires more than just individual brilliance — it demands effective leadership that nurtures and empowers the creativity of others. When managers actively support and champion the creative ideas of their team members, it cultivates a collaborative and inclusive environment in which diverse perspectives are valued and innovative solutions flourish. By incentivizing managers to prioritize the development and implementation of their employees' ideas, companies not only encourage a more sustainable approach to innovation but also foster stronger bonds between managers and their teams. This can lead to increased engagement, retention, and, ultimately, organizational success.

Importantly, by shifting the focus away from individual achievement to supporting employee creativity, organizations can alleviate managers' insecurities about being overtaken by their subordinates. When managers are encouraged to act as mentors, facilitators, and champions of their team members' ideas, they are more likely to see their role as empowering others rather than competing with them. This not only reduces the fear of being overshadowed but also fosters a sense of pride and fulfillment in seeing their team members succeed and contribute meaningfully to the organization's goals. Furthermore, recognizing and rewarding managers for their support of employee creativity sends a powerful message that the organization values teamwork and recognizes the importance of cultivating talent from within.

Foster Managers' Organizational Identification

Finally, organizations should help managers understand how their actions contribute to the business's broader goals and facilitate their identification with the company. Managers who have this understanding and identification are empowered to overcome territorial tendencies and prioritize the collective good over individual concerns, becoming more supportive of employee innovation in the process.

When subordinates propose ideas that can disrupt the status quo, it can trigger feelings of insecurity among some managers, who fear losing control.

In our studies, we found exactly this: Insecure managers who were encouraged to think about their attachment to the broader organization and its goals significantly reduced their territorial behavior. In a lab experiment, we assigned 330 participants to play the role of a team leader in a simulated scenario where we manipulated their social status. We primed their organizational identification by asking them to reflect on how it feels to be closely identified (or not) with their company and showed them alleged responses from other participants that reinforced their reflections. Finally, participants were asked to rate their willingness to sponsor and implement a novel product idea proposed by an employee on their team. The results showed that for those participants who reflected on being highly identified with the company, the impact of having low status on the endorsement of novel ideas was diminished.

Taking steps to improve organizational identification will not only facilitate greater cooperation and teamwork but also encourage a more inclusive culture in which diverse perspectives and innovative ideas are valued and supported. Moreover, when managers transcend their self-interests and prioritize the success of the organization, they set a positive example for their teams and contribute to a culture of mutual trust, respect, and collaboration.

Organizations can further enhance managers' attachment by providing opportunities for professional development and growth that are aligned with the business's strategic objectives. Recognizing and rewarding managers who demonstrate a commitment to the organization's goals also reinforces their attachment and encourages a culture that values collective achievement.

THE PHENOMENON OF TERRITORIAL managers stifling innovation within their companies is a pervasive challenge that requires careful examination and targeted interventions. Our research sheds light on the underlying dynamics driving this behavior, revealing how managers' low social status can lead to defensive territorial actions that hinder the implementation of novel ideas. By recognizing the importance of managers' social status and fostering a supportive peer environment, companies can empower managers to embrace a facilitative leadership role, thereby nurturing a culture of collaborative innovation. Furthermore, by emphasizing the broader organizational goals and facilitating managers' identification with the organization, employers can promote a mindset shift from individual achievement to collective success. Ultimately, addressing the challenges posed by territorial managers requires a multifaceted approach that prioritizes collaboration, inclusivity, and organizational attachment. ■

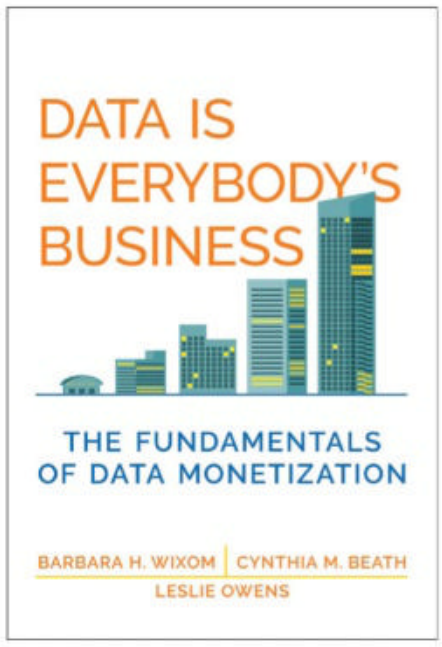
Vijaya Venkataramani *is the Dean's Professor of Leadership and Innovation in the Management & Organization group at the R.H. Smith School of Business, University of Maryland.* **Rellie Derfler-Rozin** *is an associate professor of management and organization at the Robert H. Smith School of Business at the University of Maryland.* **Xin Liu** *is an associate professor at the School of Finance, Renmin University of China, and an affiliated visiting research fellow at the University of Bath School of Management.* **Jih-Yu Mao** *is an associate professor in organizational behavior at the University of Nottingham Ningbo China.*

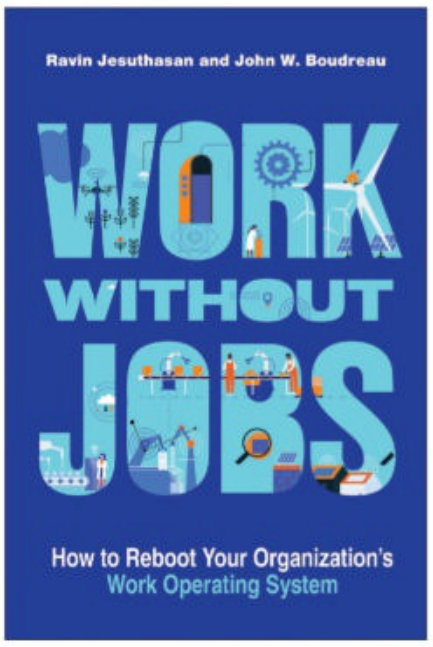
Reprint 65417. For ordering information, see page 4.

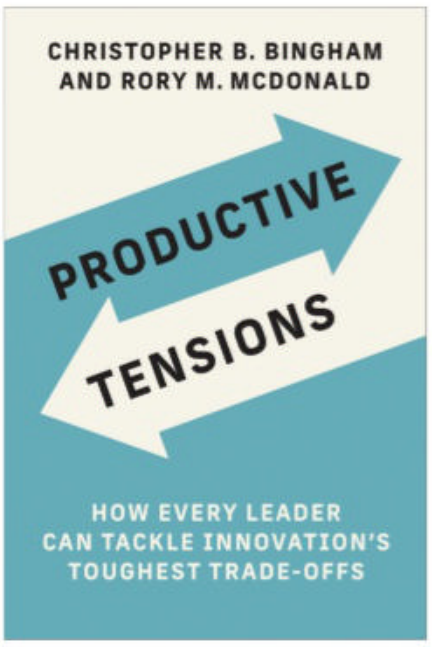
Smart leaders read smart books

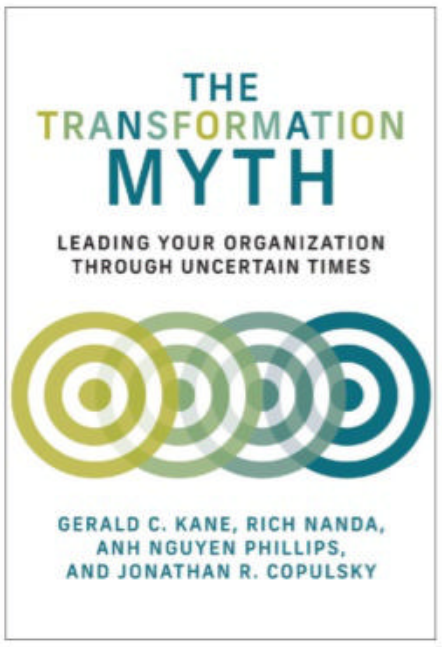
Exclusive offer for *MIT Sloan Management Review* readers — 20% off with code MITSMR20*

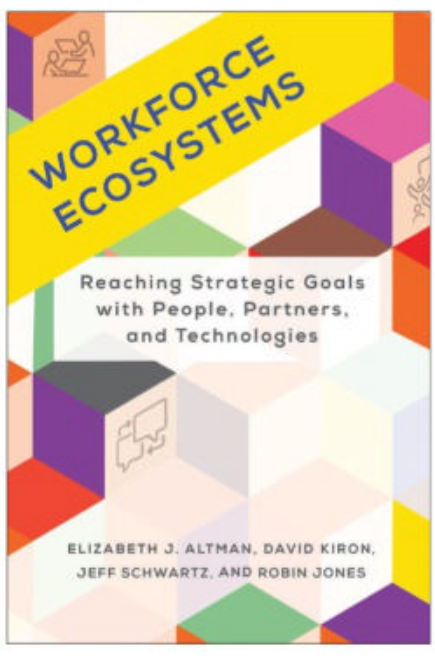
The books in the Management on the Cutting Edge Series—copublished with *MIT Sloan Management Review*—offer practical, proven guidance for the business leaders on the cutting edge of digital transformation, competitive strategy, and the future of work.

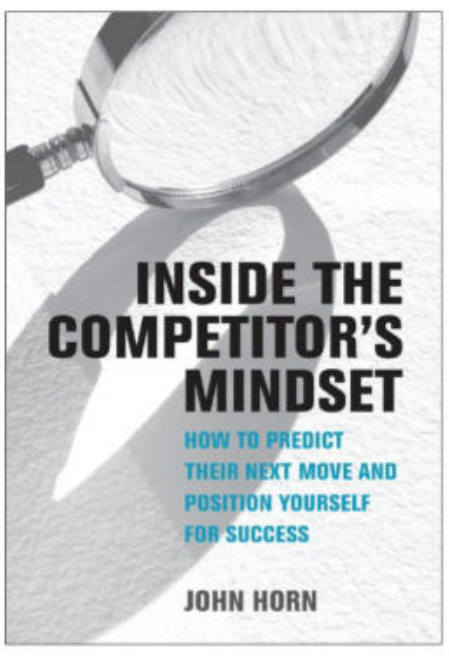


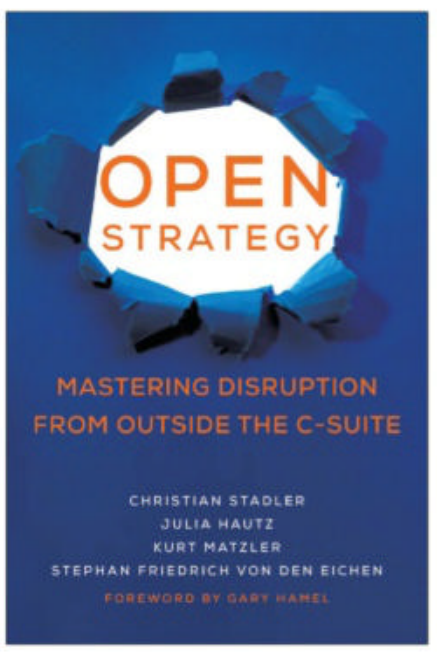


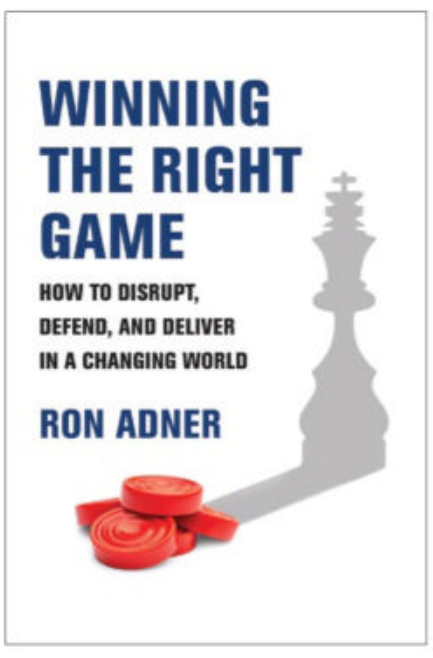












*Valid for purchases on penguinrandomhouse.com shipped to a US mailing address. Visit **mitpress.mit.edu/discount-codes** for full terms and conditions.

[Special Report]

Overcoming the Hard Problems to Advance AI Practice

As excitement around large language models (LLMs) spurs spending on AI, the salient question for business leaders remains, What is the return on our data science investments? In the near term, advanced analytics and machine learning are the workhorse technologies for creating significant value from data assets. Not that doing so is easy; companies face numerous challenges along the way.

Much AI risk becomes apparent when systems are in production, so truly responsible AI isn't just a concern at the front end of the development process. Cathy O'Neil, who posed hard questions about the unintended consequences of algorithmic decision-making in her 2016 book, *Weapons of Math Destruction*, has pioneered the practice of algorithmic auditing. O'Neil and coauthors Jake Appel and Sam Tyner-Monroe walk readers through their approach and discuss how it can be applied to generative AI tools as well.

The trade-off between using data for insights and protecting customers' personal data grows only more difficult as bad actors improve their techniques for re-identifying anonymized data sets. Gregory Vial, Julien Crowe, and Patrick Mesana explain why dealing with this challenge will require data scientists to gain a more sophisticated understanding of data protection and compel cybersecurity staffs to learn a wider range of protection techniques. They draw lessons from emerging practices at National Bank of Canada, where data scientists, data owners, and cybersecurity teams are collaborating to apply data protection practices that don't render data unusable for analytics.

When machine learning projects do get the go-ahead, however, too many initiatives fail upon adoption because data scientists didn't thoroughly understand the original business problem. To find out where such efforts are going wrong, Dusan Popovic, Shreyas Lakhtakia, Will Landecker, and Melissa Valentine studied data science projects that were shelved. They found that convincing data scientists to drop their assumptions and start asking more fundamental questions of their business counterparts is key to avoiding machine learning project failures.

Finally, just as corporations are experimenting with LLMs to figure out where they can add value at relatively low risk, advanced analytics teams can be looking at how they might incorporate generative AI into practice. Pedro Amorim and João Alves see promise for LLMs to take on some data science drudgery, and for their natural language interfaces to make it easier for business managers to collaborate in the development process and understand results.

— *The* MIT SMR *Editors*

Paul Garland

Auditing Algorithmic Risk

How do we know whether algorithmic systems are working as intended? A set of simple frameworks can help even nontechnical organizations check the functioning of their AI tools.

By Cathy O'Neil, Jake Appel, and Sam Tyner-Monroe

ARTIFICIAL INTELLIGENCE, LARGE LANGUAGE MODELS (LLMs), and other algorithms are increasingly taking over bureaucratic processes traditionally performed by humans, whether it's deciding who is worthy of credit, a job, or admission to college, or compiling a year-end review or hospital admission notes.

But how do we know that these systems are working as intended? And who might they be unintentionally harming?

Given the highly sophisticated and stochastic nature of these new technologies, we might throw up our hands at such questions. After all, not even the engineers who build these systems claim to understand them entirely or to know how to predict or control them. But given their ubiquity and the high stakes in many use cases, it is important that

Paul Garland

we find ways to answer questions about the unintended harms they may cause. In this article, we offer a set of tools for auditing and improving the safety of any algorithm or AI tool, regardless of whether those deploying it understand its inner workings.

Algorithmic auditing is based on a simple idea: Identify failure scenarios for people who might get hurt by an algorithmic system, and figure out how to monitor for them. This approach relies on knowing the complete use case: how the technology is being used, by and for whom, and for what purpose. In other words, each algorithm in each use case requires separate consideration of the ways it can be used for — or against — someone in that scenario.

This applies to LLMs as well, which require an application-specific approach to harm measurement and mitigation. LLMs are complex, but it's not their technical complexity that makes auditing them a challenge; rather, it's the myriad use cases to which they are applied. The way forward is to audit how they are applied, one use case at a time, starting with those in which the stakes are highest.

The auditing frameworks we present below require input from diverse stakeholders, including affected communities and domain experts, through inclusive, nontechnical discussions to address the critical questions of who could be harmed and how. Our approach works for any rule-based system that affects stakeholders, including generative AI, big data risk scores, or bureaucratic processes described in a flowchart. This kind of flexibility is important, given how quickly new technologies are being developed and applied.

Finally, while our notion of audits is broad in that respect, it is narrow in scope: An algorithmic audit raises alerts only to problems. It then falls to experts to attempt to solve those problems once they've been identified, although it may not be possible to fully resolve them all. Addressing the problems highlighted by algorithmic auditing will spur innovation as well as safeguard society from unintended harms.

Ethical Matrix: Identifying the Worst-Case Scenarios

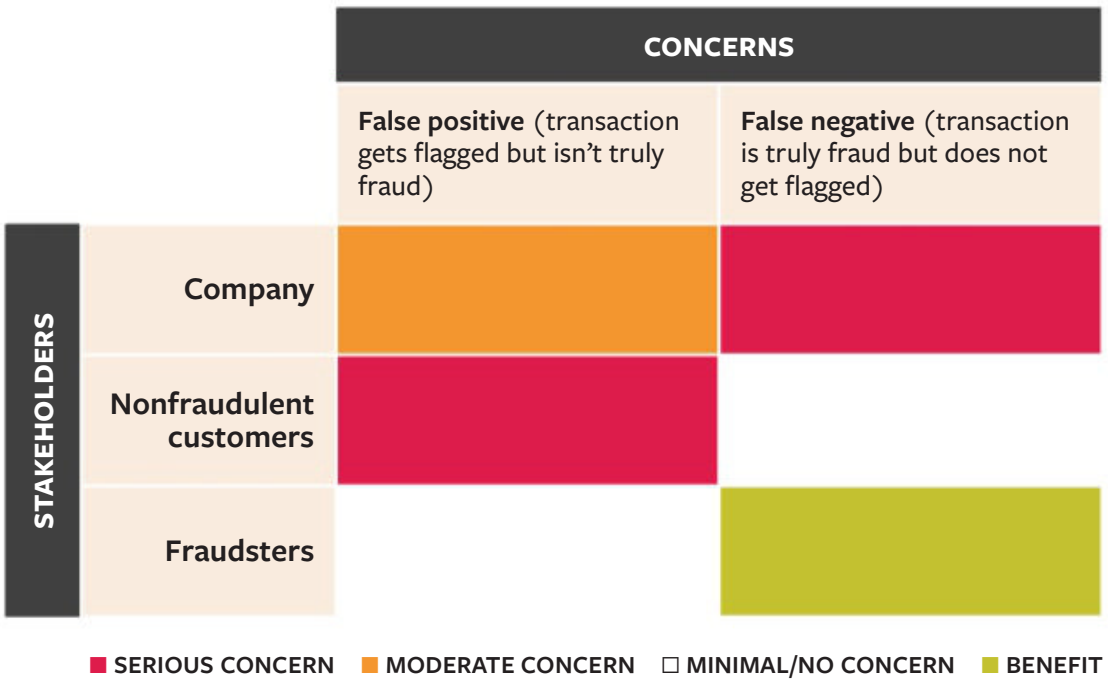
In a given use case, how could an algorithm fail, and for whom? At O'Neil Risk Consulting & Algorithmic Auditing (ORCAA), we developed the Ethical Matrix framework to answer this question.[1]

The Ethical Matrix identifies the stakeholders of the algorithm in the context of its intended use and how they are likely to be affected by it. Here, we take a broad approach: Anybody affected by the algorithm, including its builders and deployers, users, and other communities potentially impacted by its adoption, are stakeholders. When subgroups have distinct concerns, they can be considered separately; for example, if lighter- and darker-skinned people have different concerns about a facial recognition algorithm, they will have separate rows in the Ethical Matrix.

Next, we ask representatives of each stakeholder group what their concerns are, both positive and negative, about the intended use of the algorithm. It's a nontechnical conversation: We describe the system as simply as possible and ask, "How could this system fail for you, and how would you be harmed if this happened? On the other hand, how could it succeed for you, and how would you benefit?" Their answers become the columns of the Ethical Matrix. To illustrate, imagine that a payments company has a fraud detection algorithm reviewing all transactions and flagging those most likely to be fraudulent. If a transaction is flagged, it gets blocked, and that customer's account gets frozen. False flags are therefore a major headache for customers, and the lost business from blocks and freezes (and complaints from annoyed customers) is a moderate worry for

A Simplified Ethical Matrix

Each cell of the matrix represents how a certain concern applies to a particular stakeholder group. Cells that indicate where a stakeholder could be gravely harmed or the algorithm violates a hard constraint are shaded red. Cells that raise some ethical worries for the stakeholder are highlighted yellow, and cells that satisfy the stakeholder's objectives and raise no worries are highlighted green.



the company. Conversely, if a fraudulent transaction goes undetected, the company is harmed but nonfraudulent customers are indifferent. Below is a simplified Ethical Matrix for this scenario.

Each cell of the Ethical Matrix represents how a particular concern applies to a particular stakeholder group.

To judge the severity of a given risk, we consider the likelihood that it will be realized, how many people would be harmed, and how badly. Where possible, we use existing data to develop these estimates. We also consider legal or procedural constraints — for instance, whether there is a law prohibiting discrimination on the basis of certain characteristics. We then color-code the cells to highlight the biggest, most pressing risks. Cells that constitute "existential risks," where a stakeholder could be gravely harmed or the algorithm violates a hard constraint, are shaded red. Cells that raise some ethical worries for the stakeholder are highlighted yellow, and cells that satisfy the stakeholder's objectives and raise no worries are highlighted green.

Finally, zooming out on the whole Ethical Matrix, we consider how to balance the competing concerns of the algorithm's stakeholders, usually in the form of balancing the different kinds and consequences of errors that fall on different stakeholder groups.

The Ethical Matrix should be a living document that tracks an ongoing conversation among stakeholders. Ideally, it is first drafted during the design and development phase of an algorithmic application or, at minimum, as the algorithm is deployed, and it should continue to be revised thereafter. It is not always obvious at the outset who all of the stakeholder groups are, nor is it feasible to find representatives for every perspective; additionally, new concerns emerge over time. We might hear from people experiencing indirect effects from the algorithm, or a subgroup with a new worry, and need to revise the Ethical Matrix.

Explainable Fairness: Metrics and Thresholds

Many of the stakeholder concerns identified in the Ethical Matrix refer to some contextual notion of fairness.

At ORCAA, we developed a framework called Explainable Fairness to measure how groups are treated by algorithmic systems.[2] It is an approach to understanding exactly what is meant by "fairness" in a given narrow context.

For example, female candidates might worry that

Benchmarking and red teaming are two approaches to auditing LLMs in diverse use cases.

an AI-based resume-screening tool gave lower scores for women than men. It's not as simple as comparing scores between men and women. After all, if the male candidates for a given job have more experience and qualifications than the female candidates, their higher scores might be justified. This would be considered *legitimate discrimination*.

The real worry is that, among equally qualified candidates, men are receiving higher scores than women. The definition of "equally qualified" depends on the context of the job. In academia, relevant qualifications might include degrees and publications; in a logging operation, they might involve physical strength and agility. They are factors one would legitimately take into account when assessing a candidate for a specific role. Two candidates for a job are considered equally qualified if they look the same according to these *legitimate factors*.

Explainable Fairness controls for legitimate factors when we examine the outcome in question. For an AI resume-screening tool, this could mean comparing average scores by gender while controlling for years of experience and level of education. A critical part of Explainable Fairness is the discussion of legitimacy.

This approach is already used implicitly in other domains, including credit. In a Federal Reserve Board analysis of mortgage denial rates across race and ethnicity, the researchers ran regressions that included controls for the loan amount, the applicant's FICO score, their debt-to-income ratio, and the loan-to-value ratio.[3] In other words, to the extent that differences in mortgage denial rates can be explained by these factors, it's not race discrimination. In the language of Explainable Fairness, these are accepted as legitimate factors for mortgage underwriting. What is missing is the explicit conversation about why the legitimate factors are, in fact, legitimate.

What would such a conversation look like? In the U.S., mortgage lenders consider applicants' FICO credit scores in their decision-making. FICO scores are lower, on average, for Black and Hispanic people

than for White and Asian people, so it's no surprise that mortgage applications from Black and Hispanic applicants are denied more often.[4] Lenders would likely argue that FICO score is a legitimate factor because it measures an applicant's creditworthiness, which is exactly what a lender should care about. Yet FICO scores encode unfairness in important ways. For instance, mortgage payments have long counted toward FICO scores, while rent payments started being counted only in 2014, and only in some versions of the scores.[5] This practice favors homeowners over renters, and it is known that decades of racist redlining practices contributed to today's race disparities in homeownership rates. Should FICO scores that reflect the vestiges of these practices be used to explain away differences in mortgage denial rates today?

We will not settle this debate here; the point is that it's a question of ethics and policy, not a math problem. Explainable Fairness surfaces difficult questions like these and assigns them to the right parties for consideration.

When looking at disparate outcomes that are not explained by legitimate factors, we must define threshold values or limits that trigger a response or intervention.

These limits could be fixed values, such as the four-fifths rule used to measure adverse impact in hiring.[6] Or they could be relative: Imagine a regulation requiring companies with a gender pay gap above the industry average to take action to reduce the gap. Explainable Fairness does not insist on a certain type of limit but prompts the algorithmic risk manager to define each one for each potential stakeholder harm.

Judging Fairness in Insurers' Algorithms

Let's consider a real example where the Ethical Matrix and Explainable Fairness were used to audit the use of an algorithm. In 2021, Colorado passed Senate Bill (SB) 21-169, which protects Colorado consumers from unfair discrimination in insurance, particularly from insurers' use of algorithms, predictive models, and big data.[7] As part of the law's implementation, which ORCAA assisted with, the Colorado Division of Insurance (DOI) released an initial draft regulation for informal comment that described quantitative testing requirements and laid out how insurers could demonstrate that their algorithms and models were not unfairly discriminating. Although the law applies to all lines of insurance, the division chose to start with life insurance.

The Ethical Matrix is straightforward here because the stakeholder groups and concerns are defined explicitly by the law. Its prohibition of discrimination on the basis of "race, color, national or ethnic origin, religion, sex, sexual orientation, disability, gender identity, or gender expression" means each group within each of those classes got a row in the matrix. As for concerns, algorithms could cause consumers to be treated unfairly at various stages of the insurance life cycle, including marketing, underwriting, pricing, utilization management, reimbursement methodologies, and claims management. The DOI chose to start with underwriting — that is, which applicants are offered coverage, and at what price — and focus initially on race and ethnicity.

In subsequent conversations with stakeholders, however, the DOI grappled with issues related to the Explainable Fairness framework: Are similar applicants of different races denied at different rates, or charged different prices for similar coverage? What makes two life insurance applicants "similar," and what factors could legitimately explain differences in denials or prices? This is the domain of life insurance experts, not data scientists.

The DOI ultimately suggested considering factors broadly considered relevant to estimating the price of a given life insurance policy: the policy type (such as term versus permanent); the dollar amount of the death benefit; and the applicant's age, gender, and tobacco use.

The division's draft quantitative testing regulation for SB21-169 instructs insurers to do regression analyses of approval/denial and price across races, and it explicitly permits them to include those factors (such as policy type and death benefit amount) as control variables.[8] Moreover, the regulation defines limits that trigger a response: If the regressions find statistically significant and substantial differences in denial rates or prices, the insurer must do further testing to investigate the disparity and, pending the results, may have to remediate the differences.[9]

Having looked at how we would audit simpler algorithms, let us now turn to how we would evaluate LLMs.

An LLM red-teaming exercise is designed to elicit unwanted responses.

Evaluating Large Language Models

LLMs have taken the world by storm, largely due to their wide appeal and applicability. But it is exactly the diversity of uses of these models that makes them hard to audit. Two approaches to evaluating LLMs, namely benchmarking and red teaming, present a way forward.

The Benchmarking Approach to LLM Evaluation. Benchmarking measures the performance of an LLM across one or more predefined, quantifiable tasks in order to compare its performance with that of other models. In the simplest terms, a benchmark is a data set consisting of inputs and corresponding desired outputs. To evaluate an LLM for a particular benchmark, simply provide the input set to the LLM and record its outputs. Then choose a metric set to quantitatively compare the outputs from the LLM to the desired set of outputs from the benchmark data set. Possible metrics include accuracy, calibration, robustness, counterfactual fairness, and bias.[10]

Consider the input and desired output shown below from a benchmark data set designed to test LLM capabilities:[11]

```
Input:
The following is a multiple choice
question about microeconomics.

One of the reasons that the government
discourages and regulates monopolies is
that
(A) producer surplus is lost and consumer
surplus is gained.
(B) monopoly prices ensure productive
efficiency but cost society allocative
efficiency.
(C) monopoly firms do not engage in
significant research and development.
(D) consumer surplus is lost with higher
prices and lower levels of output.

Answer:

Desired Output:
(d) consumer surplus is lost with higher
prices and lower levels of output.
```

In this example, the *accuracy* of the model is measured by computing the proportion of correctly answered multiple-choice questions in the benchmark data set. In benchmarking LLM evaluations, metrics are defined according to the type of response elicited from the model. For example, accuracy is very simple to calculate when all of the questions are multiple choice and the model simply has to choose the correct response, whereas determining the accuracy of a summarization task involves counting up matching n-grams between the desired and model outputs.[12] There are dozens of benchmark data sets and corresponding metrics available for LLM evaluation, and it is important to choose the most appropriate evaluations, metrics, and thresholds for a given use case.

Creating a custom benchmark is a labor-intensive process, but an organization may find that it is worth the effort in order to evaluate LLMs in exactly the right way for its use cases.

Benchmarking does have some drawbacks. If the benchmark data happened to be in the model's training data, it would have "memorized" the responses in its parameters. The frequency of this ouroboros-like outcome will only increase as more benchmark data sets are published. LLM benchmarking is also not immune to Goodhart's law, that is, "when a measure becomes a target, it ceases to be a good measure." In other words, if a specific benchmark becomes the primary focus of model optimization, the model will be over-fitted at the expense of its overall performance and usefulness.

In addition, there is evidence that as models advance, they become able to detect when they are being evaluated, which also threatens to make benchmarking obsolete. Consider Anthropic's Claude 3 series of models, released in March 2024, which stated, "I suspect this ... 'fact' may have been inserted as a joke or to test if I was paying attention, since it does not fit with the other topics at all," in response to a needle-in-a-haystack evaluation prompt.[13] As models increase in complexity and ability, the benchmarks used to evaluate them must also evolve. It is unlikely that the benchmarks used today to evaluate LLMs will be the same ones in use just two years from now.

It is therefore not enough to evaluate LLMs with benchmarking alone.

The Red-Teaming Approach to LLM Evaluation. Red teaming is the exercise of testing a system for robustness by using an adversarial approach. An LLM red-teaming exercise is designed to elicit unwanted responses from the model.

LLMs' flexibility in the generation of content presents a wide variety of potential risks. LLM red teams may try to make the model produce violent or dangerous content, reveal its training data, infringe on copyrighted materials, or hack into the model provider's network to steal customer data. Red teaming can take a highly technical path, where, for example,

nonsensical characters are systematically injected into the prompts to induce problematic behavior; or a social engineering path, whereby red teamers try to "trick" the model using natural language to produce unwanted output.[14]

Robust red teaming requires a multidisciplinary approach, diverse perspectives, and the engagement of all stakeholders, from developers to end users. The red team should be designed to assess the risks associated with at least each red cell in the Ethical Matrix. This results in a collaborative, sociotechnical approach that ensures a more comprehensive evaluation of the model, thus enhancing the rigor of the evaluation and the safety of the model. Other LLMs can also be used to generate red-teaming prompts.

Red teaming helps LLM developers better protect models against potential misuse, thereby enhancing the overall safety and efficacy of the model. It can also uncover issues that might not be visible under normal operating conditions or during standard testing procedures. A collaborative approach to red teaming built on the Ethical Matrix ensures a thorough and rigorous evaluation, bolstering the robustness of the model and the validity of its outcomes.

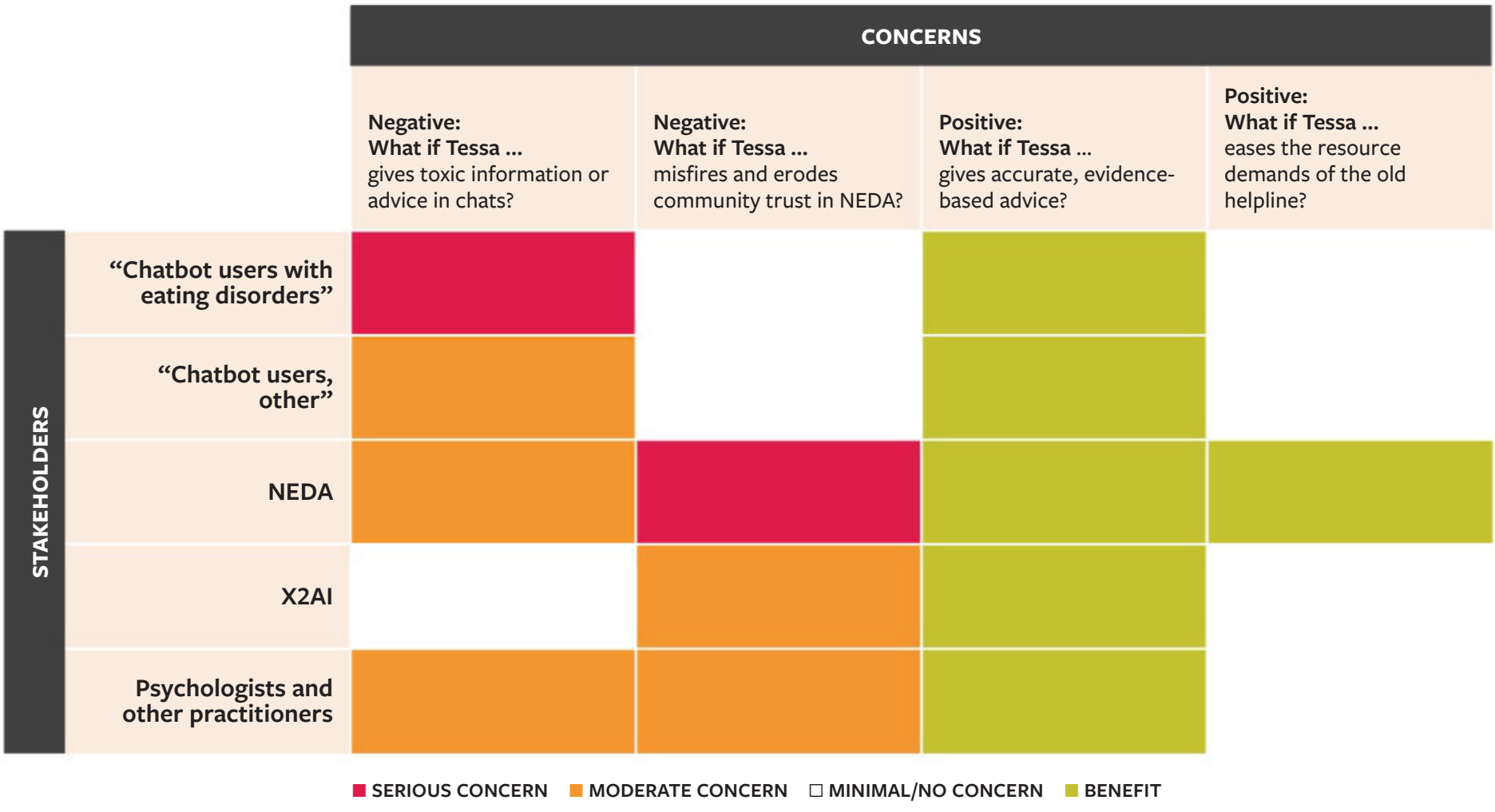
A significant limitation of red teaming is its inherent subjectivity: The value and effectiveness of a red-teaming exercise can vary greatly depending on the creativity and risk appetite of the individual stakeholders involved. And because there are no established standards or thresholds for red-teaming LLMs, it can be difficult to determine when enough red teaming has been done or whether the evaluation has been comprehensive enough. This can leave some vulnerabilities undetected.

Another obvious limitation of red teaming is its inability to evaluate for risks that have not been anticipated or imagined. Risks that are unforeseen will not be included in red teaming, making the model uniquely vulnerable to unanticipated scenarios.

Therefore, while red teaming plays a vital role in the testing and development of LLMs, it should

Sketch of the Ethical Matrix for Tessa in Our Thought Experiment

The National Eating Disorders Association (NEDA) released a chatbot named Tessa that was taken down after it gave out harmful advice. Here we visualize the exercise that may have anticipated such outcomes.



be complemented with other evaluation strategies and continuous monitoring to ensure the safety and robustness of the model.

How Would We Audit Tessa, the Eating Disorder Chatbot?

The nonprofit National Eating Disorders Association (NEDA) is one of the largest organizations in the U.S. dedicated to supporting people who have eating disorders. In May 2023, amid controversy, NEDA took down an LLM-powered wellness chatbot called Tessa from its website. Tessa was designed to "[help] you build resilience and self-awareness by introducing coping skills at your convenience," but screenshots posted to Instagram showed that it sometimes gave harmful diet advice, like adopt a "safe daily calorie deficit."[15] This highly public failure could have been avoided if Tessa had been audited with the frameworks and techniques outlined above.

Before we explain why, two other details are relevant. First, NEDA operated an eating disorder helpline, staffed by employees, for over 20 years; in 2022, nearly 70,000 people used it. Calls to the helpline soared during the COVID-19 pandemic, and, increasingly, callers were in active crisis rather than just seeking information or referrals. NEDA claimed that the human-staffed helpline wasn't set up to handle the growing level of demand, so the organization closed it down in May 2023 and laid off five paid employees who staffed it. Tessa was intended as a replacement for this service.[16] Second, NEDA did not build Tessa in-house. It was built by the company X2AI (now Cass), which offers an AI health care assistant that was customized for NEDA.

Let's sketch an Ethical Matrix for Tessa, shown below.[17] First, we'll define the stakeholders in the context of its intended use in the rows of the matrix. Visitors to the website who chat with Tessa are clearly stakeholders. Visitors who themselves suffer from eating disorders are a distinct subgroup, since the stakes are higher for them. NEDA is also a stakeholder, as is X2AI, the chatbot developer. Finally, psychologists and other practitioners who serve people struggling with eating disorders are a stakeholder group, since they have an interest in the well-being of their patients.

As for concerns, which form the columns of the matrix, anybody who chats with Tessa wants it to give information that is helpful and evidence-based. Visitors who suffer from eating disorders have a heightened concern around bad information or advice that could deepen their disorder or trigger a relapse. NEDA of course agrees that Tessa should give helpful and evidence-based advice. In addition to helping (not harming) individuals, the issue of community trust is at stake. If Tessa misfires and undermines trust in NEDA, then people will look elsewhere for advice on this topic. In this case, NEDA would fail its core mission, practitioners would be losing a valuable resource, and X2AI would likely lose NEDA as a customer. Finally, NEDA also has a concern around efficiency: The old helpline would have needed more resources to handle the increased volume and urgency of calls, while Tessa would allow the organization to cut its staff in favor of a (presumably cheaper) technology expense.

In this Ethical Matrix, we've highlighted two concerns as grave (red). First, chatbot users with eating disorders could be directly harmed if Tessa gives them toxic information or advice. Second, NEDA could lose its standing as a trusted organization if Tessa has a highly public misfire. These scenarios are also concerning to other stakeholders but more moderately (yellow). On the positive side, everybody wants Tessa to give good advice, and NEDA alone cares about the efficiency gain from using Tessa relative to the old helpline. Tracking benefits helps when transitioning from one system to another, to ensure that a system is being replaced with something that works at least as well.

Any organization deploying algorithms in high-impact areas needs to track the risks of stakeholder harms.

The next step in an audit would be creating monitors to track these stakeholder concerns. The LLM evaluation techniques discussed above come into play. Red teaming — trying to trick Tessa into violating its own rules — could address the concern around toxic information.

Benchmarking would address the positive concern around Tessa giving accurate advice. NEDA could create a benchmarking data set of questions on the topic, as well as correct answers. Tessa could be routinely tested on a regularly updated benchmark set to verify its accuracy.

The red-teaming and benchmarking exercises

would have defined target metrics that Tessa would need to meet — or limits it would have to avoid crossing — to be deployed or stay in service.

The NEDA story is hardly an isolated example. LLM-based chatbots are increasingly providing information and advice on important topics, yet they are not being adequately audited in advance, and they are failing in alarming ways. A New York City government chatbot was recently found to be telling users that landlords didn't have to accept tenants on rental assistance and that employers could take a cut of their workers' tips — practices that are against the law.[18] And chatbots deployed by TurboTax and H&R Block were recently found to be giving faulty advice to tax filers.[19]

AUDITING ALGORITHMS, AS PRESENTED here, takes a high-level view: Any organization looking to deploy algorithms in high-impact areas needs to keep track of the risks of stakeholder harms. This should be done in a context-specific way and with generalized methods that encompass everything from old-fashioned flowcharts to classic machine learning to LLMs.

A final note: Sometimes, the risk of AI or LLMs cannot be reliably measured or understood because the results are too stochastic or inconsistent. That might mean that AI simply shouldn't be used in that context. But that's a decision for organization leaders to make, with reference to internal rules or external laws and regulations; it's not the role of the auditor to fix problems, just to locate and measure them. ■

Cathy O'Neil *is the CEO of O'Neil Risk Consulting & Algorithmic Auditing (ORCAA) and the author of* Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy *(Crown, 2016).* **Jake Appel** *is chief strategist at ORCAA.* **Sam Tyner-Monroe***, Ph.D., is the managing director of responsible AI at DLA Piper.*

REFERENCES

1. The Ethical Matrix is based on a bioethical framework originally conceived by philosopher John Mepham for the sake of running ethical experiments. For a detailed presentation, see C. O'Neil and H. Gunn, "Near-Term Artificial Intelligence and the Ethical Matrix," ch. 8 in "Ethics of Artificial Intelligence," ed. S.M. Laio (New York: Oxford University Press, 2020).
2. C. O'Neil, H. Sargeant, and J. Appel, "Explainable Fairness in Regulatory Algorithmic Auditing," West Virginia Law Review, forthcoming.
3. See N. Bhutta, A. Hizmo, and D. Ringo, "How Much Does Racial Bias Affect Mortgage Lending? Evidence From Human and Algorithmic Credit Decisions," Finance and Economics Discussion Series 2022-067, Federal Reserve Board, Washington, D.C., 2022. Table 6A is particularly relevant.
4. M. Leonhardt, "Black and Hispanic Americans Often Have Lower Credit Scores — Here's Why They're Hit Harder," CNBC, Jan. 28, 2021, www.cnbc.com.
5. B. Luthi, "How to Add Rent Payments to Your Credit Reports," myFICO, Dec. 14, 2022, www.myfico.com.
6. The four-fifths rule is not a law but a rule of thumb from the U.S. Equal Employment Opportunity Commission, saying that selection rates between groups of candidates for a job or promotion (such as people of different ethnicities) cannot be too different. In particular, the rate for the group with the lowest selection rate must be at least four-fifths that of the group with the highest selection rate. See more at "Select Issues: Assessing Adverse Impact in Software, Algorithms, and Artificial Intelligence Used in Employment Selection Procedures Under Title VII of the Civil Rights Act of 1964," U.S. Equal Employment Opportunity Commission, May 18, 2023, www.eeoc.gov.
7. "SB21-169 — Protecting Consumers From Unfair Discrimination in Insurance Practices," Colorado Department of Regulatory Agencies, accessed April 24, 2024, https://doi.colorado.gov.
8. "3 CCR 702-10 Unfair Discrimination Draft Proposed New Regulation 10-2-xx," Colorado Department of Regulatory Agencies Division of Insurance, accessed April 24, 2024, https://doi.colorado.gov/.
9. The draft regulations also define these terms. "Statistically significant" means having a p-value of <0.05, and "substantial" means a difference in approval rates, or in price per $1,000 of face amount, of >5 percentage points. The details of the further tests are beyond the scope of this article, but the main idea is to inspect whether "external consumer data and information sources" (that is, nontraditional rating variables, such as cutting-edge risk scores, which insurers often purchase from third-party vendors) used in underwriting and pricing are correlated with race in a way that contributes to the observed differences in denial rates or prices. If inspection shows they are, then the insurer must "immediately take reasonable steps developed as part of [its] risk management framework to remediate the unfairly discriminatory outcome."
10. P. Liang, R. Bommasani, T. Lee, et al., "Holistic Evaluation of Language Models," Transactions on Machine Learning Research, published online Aug. 23, 2023, https://openreview.net.
11. D. Hendrycks, C. Burns, S. Basart, et al., "Measuring Massive Multitask Language Understanding," arXivLabs, published online Sept. 7, 2020, https://arxiv.org.
12. Liang et al., "Holistic Evaluation of Language Models."
13. B. Edwards, "Anthropic's Claude 3 Causes Stir by Seeming to Realize When It Was Being Tested," Ars Technica, March 5, 2024, https://arstechnica.com.
14. A. Zou, Z. Wang, N. Carlini, et al., "Universal and Transferable Adversarial Attacks on Aligned Language Models," arXivLabs, published online July 27, 2023, https://arxiv.org; and D. Ganguli, L. Lovitt, J. Kernion, et al., "Red Teaming Language Models to Reduce Harms: Methods, Scaling Behaviors, and Lessons Learned," arXivLabs, published online Aug. 23, 2022, https://arxiv.org.
15. L. McCarthy, "A Wellness Chatbot Is Offline After Its 'Harmful' Focus on Weight Loss," The New York Times, June 8, 2023, www.nytimes.com.
16. K. Wells, "National Eating Disorders Association Phases Out Human Helpline, Pivots to Chatbot," NPR, May 31, 2023, www.npr.org.
17. By "sketch," we mean we are imagining the stakeholders and their concerns. Truly creating an Ethical Matrix for this use case would entail interviewing real representatives of these stakeholder groups. In this article, we approach it as a thought experiment.
18. C. Lecher, "NYC's AI Chatbot Tells Businesses to Break the Law," The Markup, March 29, 2024, https://themarkup.org.
19. G.A. Fowler, "TurboTax and H&R Block Now Use AI for Tax Advice. It's Awful," The Washington Post, March 4, 2024, www.washingtonpost.com.

Reprint 65435. For ordering information, see page 4.

Avoid ML Failures by Asking the Right Questions

Machine learning solutions can miss the mark when data scientists don't check their assumptions. Adopting a beginner's mindset in any domain can help.

By Dusan Popovic, Shreyas Lakhtakia, Will Landecker, and Melissa Valentine

IN OUR COLLECTIVE DECADES OF EXPERIence building, leading, and studying companies' machine learning (ML) deployments, we have repeatedly seen projects fail because talented and well-resourced data science teams missed or misunderstood a deceptively simple piece of the business context. Those gaps create obstacles to correctly understanding the data, its context, and the intended end users — ultimately jeopardizing the positive impact ML models can make in practice.

We have discovered that small mistakes and misunderstandings are much less likely to cascade into failed projects when development teams engage with colleagues on the business side and ask enough questions to deeply understand the process and the problem at hand. Asking questions might seem like a simple step, but that might not be part of a company's, team's, or an industry's culture. Appearing to be in command of all the information needed may be one of the ways employees signal competence in the organization. And while data scientists might possess technical mastery, they can lack the soft skills to reach a deep, accurate mutual understanding with business partners.

At the same time, business partners often hesitate to ask questions themselves and don't necessarily know what information or context would be helpful to share with a data science team. It's hard work on both sides to have the kinds of interactions that allow everyone to surface and question assumptions, and identify the most important elements of business context.

Setting ML projects up for success with those kinds of useful interactions requires leaders to foster a culture that normalizes and values asking questions with a beginner's mindset — and putting aside ego and past expertise. One data scientist we've worked with became very intentional about this after noticing that he makes the fewest mistakes when he is in a new context and must ask a lot of questions. But what are the right questions to ask? In this article, we present three examples where significant ML projects failed and explore how asking the right questions with a beginner's mindset could have improved the interactions between data scientists and business partners, and helped their ML deployments succeed in creating business value.

Paul Garland

SCENARIO 1:
Ask 'What is the business process?' not 'What is the data set?'

Facing the first economic downturn prompted by the COVID-19 pandemic, a local finance team at a multinational retail company had a hunch that some customers would weather the storm with a little help whereas others

were at risk of bankruptcy. The team wondered whether the company's data science team could help them predict which customers were likely to file for bankruptcy each month. This information would allow the finance team to identify solvent customers and temporarily extend more credit to assist them during the downturn while limiting the company's exposure to customers who were very likely to default. The local finance team requested this analysis from the local IT partner team, which in turn engaged the company's data science center of excellence to produce the model. Using the data provided, this central data science team successfully developed a model that seemed to perform well: During offline tests, using historical data, the data scientists could accurately predict which customers were likely to default on payments.

However, when the model was deployed with the local finance team, it no longer performed well. In fact, it was essentially useless for predicting customer bankruptcy each month, despite performing well during testing and prototyping.

The missing link: Process understanding. This central data science team received and analyzed a compelling and complete data set but, having had little interaction with the team that had commissioned and would be using the model, failed to grasp the underlying business processes. They did not understand the legal process for bankruptcy in the country the finance team was concerned with, or how the timeline of bankruptcy was recorded by the company. The data scientists built the model based on a variable that flagged customers as either having defaulted or not and trained the model to detect the typical pattern of transactions right before the customer was flagged as being in default.

There were three key events on the timeline of a customer declaring bankruptcy: the customer not meeting financial obligations, the customer then filing for bankruptcy in court, and the court finally making the bankruptcy ruling. What the data scientists did not know was that customers were not flagged as having defaulted when they began missing payments; rather, the flags were entered on their accounts only after the bankruptcy court ruling. The data scientists missed that because they were using training data in which the default had already been reported for all customers in the data set; they did not realize that live customer accounts would not be flagged as defaulted until about a year after customers started missing payments. In other words, the model used data to make a prediction that in reality would be unavailable to the prediction system when the model was run against live data — a problem that data scientists call target leakage.

As Kate Jordan, a data scientist at Octagram Analytics, told us, data scientists are often trained to think in terms of the data set in front of them, as well as perhaps some other data that's accessible and might be relevant to their analysis. By focusing their questioning on the data set, they overlook the context of the operational system that the model will be placed into. Jordan has seen other cases similar to our example above, where data scientists analyzed a data set that included all of the variables, not understanding that one of those variables was not actually recorded in the data in the live system at the time they were programming the model to analyze and act on it. She has seen data science teams focus on variables that they could not actually put into the algorithm in the live operational system. She warned against teams handing data scientists "sterile" data sets and encouraged teams instead to ask "What is the process?" and "What is the system and the system flow that this model is going to be placed into?"

One industry-standard practice that helps data science teams find answers to these questions and develop deep business understanding is to shadow the entire business process. We think that regardless of where the analytics team is located within a company — even those with a global center of excellence — a data scientist working on a problem should be temporarily deployed to the function in question. There, they should spend a meaningful amount of time observing how the job is normally done, what tools are used, and where the inefficiencies arise. Shadowing the business process and being embedded among users is not only a great way to develop a detailed understanding of the problem space itself but also a means of gaining a solid foundation for subsequent adoption of the solution. A related process is for the teams to prioritize creating a diagram that walks through the process.

Business leaders can value and normalize the value of these processes and this level of understanding rather than handing centralized data scientists a sterile, decontextualized data set that they must analyze without understanding the business process and operational systems.

SCENARIO 2:
Ask 'Who are the decision makers, and what are their decisions and incentives?' not just 'What should we predict?'

The revenue management team at the headquarters of a large multinational bank was facing a serious problem. The profit margins of its home mortgage businesses had been steadily eroding for several years in a row. As the team investigated this trend, they learned that customer-facing credit officers who worked in

local branches had been offering interest rates toward the lower end of the assigned discretionary ranges. The revenue management team hypothesized that a data-driven approach to setting the terms of mortgage loans would help improve profits. They commissioned a loan price optimization system from a centralized data science team, which developed, tested, and shipped an ML system that had been shown to successfully determine profit-maximizing terms for each individual loan.

During the initial live A/B test, the system displayed performance superior to that of most individual credit officers in terms of realized profits. However, none of those credit officers used the system after the testing was completed.

The missing link: Competing organizational priorities. As in most companies, the bank's executive board defines and communicates organizationwide strategy. In this case, the board's strategic focus was profit maximization, a priority that cascaded down to the top-level functions such as retail banking and revenue management. To directly address this strategic goal, the revenue management team commissioned the development of the profit-maximizing pricing model. However, the intended users of the model were housed in the retail banking function, which had its own operational KPIs for local branches. In this case, retail banking had set several KPIs around growth, and consumer-facing credit officers in the local branches were given financial incentives to sign more loans to fuel the desired growth. The credit officers received higher bonuses if they sold more loans — regardless of the loans' profitability — and in most cases the most straightforward way to achieve that was to apply the largest allowed discount. From a technical perspective, the data scientists had created a model that effectively optimized the given metric. However, from an organizational perspective, their incentives were not aligned with those of their users. The data scientists were optimizing a metric that their sponsors had asked for but their end users did not care about.

To avoid this kind of failure, it's essential for business leaders and data scientists to better understand the decision makers using the ML system, and the factors informing their decisions. Before engaging in a full-blown modeling exercise, a data scientist and their business stakeholder should make a comprehensive map of decision makers and decisions. This can be another output from shadowing and part of creating the business process diagram. They should seek to understand what decisions are under the control of the project sponsors, the intended end users, their partners, and their business customers. This might include asking users how they might act in response to the kinds of results that a data scientist can anticipate the model producing. This question can help identify gaps in understanding a problem. In the bank's case, the data scientists' sponsors were focused on optimizing one metric (profit), but their end users were incentivized to optimize a different metric (revenue growth) and thus did not follow the ML recommendations. This failure could have been prevented if the data scientists had sought to understand the decision makers and their incentives rather than just asking what variable to predict.

Jordan told us that in her previous role at Zurich Insurance, she and her team would sit with users for days and ask questions as they interacted with the data, such as "What would you look at?" and "What do you do with that data insight?" They even rescued a failed project using this method, after a data scientist (who had never been to the intended users' office) delivered a sophisticated neural net to predict fraud in a disembodied data set, and the model was never adopted by the users. As Jordan and her team questioned the intended users, they came to understand that the users were actually responsible for collecting and producing evidence of fraud that would be sufficiently substantive for regulatory or court proceedings.

A neural net prediction did not meet the standard of evidence; the users needed to construct an account of the fraudulent activities using the actual bills that proved fraud. In other words, their decisions needed to be based on documentary proof; they could not forward cases for potential enforcement action on the basis of an analysis that simply predicted the likelihood of fraudulent behavior.

SCENARIO 3:
Ask 'Who are the stakeholders, and what actions do they control?'

At Anheuser-Busch InBev's European operations, the team responsible for the company's B2B e-commerce platform sought to improve conversion and repurchase rates. Online promotions were their primary tool to achieve this goal, and the team was responsible for designing the key aspects, or mechanics, of a promotion: what products it would include, how long it would run, what type of discount would be offered, and so on. Category managers at the company decided which brands would be promoted, and then promotions were typically executed in bulk each month, using a single mechanism determined separately for each brand.

After running a number of promotions, the platform team saw signs that different customers preferred different types of promotions. This indicated that personalizing promotions to the preferences of each B2B customer

might capture additional value by increasing overall conversion and repurchase rates. The e-commerce team engaged a local data science team to produce a personalized promotion model. The deployed system consisted of two layers: A model predicted the probability of conversion for every customer that was given a fixed setup of promotion mechanics, and an optimization wrapper simulated different mechanics for each customer in order to identify the one that resulted in the biggest increase in conversion probability for that particular customer. For example, the system might recommend increasing the range of products included, with the intent of increasing the likelihood that a particular customer would make a purchase from 90% to 95%.

However, after a series of live A/B tests, the data science team was surprised to see that the system was failing to increase customers' conversion or repurchase rates. While the underlying model had been extremely good at estimating conversion probabilities for each customer and promotion combination, in live tests the system as a whole failed to move the needle on the salient KPIs.

The missing link: A key variable outside of the team's control. After investigating the output of the model, the data science team discovered that promotional mechanics — the levers that the platform team could control — were not the strongest factors determining customer purchases. Whether they were offered a direct 33% discount or a "buy two, get one free" promotion, for example, did not materially affect conversion probability for most customers. Instead, the most significant variable turned out to be which brand was promoted: If a certain customer was offered the right brand with a discount, they would convert regardless of the mechanics applied. Unfortunately, the choice of which brands to promote was made at a higher level in the organization, which meant that this insight could not be immediately operationalized. Organizational alignment and tight cross-functional collaboration, not a technological solution, had to be implemented before the overall approach could pivot to personalizing the brand offer and increasing conversion rates. From a technical perspective, data scientists were perfectly capable of modeling conversion, and they successfully identified levers that affected it. But from an organizational perspective, their direct users had limited control to act on the recommendations the model surfaced, at least during the initial iteration of the project.

To avoid this type of failure, a best practice is for business leaders and data scientists to understand the stakeholders that are directly and indirectly involved in their work. One way to do this is to ask, "Who are the stakeholders relevant to the process at hand, and what actions do they control?" The answer should result in a clear map of the decision process, with responsibilities unambiguously attached to each junction where human input is involved. For business leaders, this helps clarify where to build relationships and whom to inform in the context of the project scope. This also creates an opportunity for data scientists to educate nontechnical partners and build support and awareness of their work. Finally, it helps both business leaders and data science leads ensure the actionability of insights delivered, by determining who controls which levers and how those tie to the data science analysis at hand.

Especially when responsibilities cross functional boundaries, all relevant decision makers should be onboarded from the beginning of a data science project, regardless of which particular function the initial request originated from. The key is to ensure that from the very outset, any insight resulting from the initiative can eventually be acted upon. Moreover, there is an inherent trust-building value in bringing potential stakeholders on board at the onset of data science work rather than once it is done. Such engagement builds cross-functional trust — giving teams not just an opportunity to learn whether insights can be acted on but also the reassurance that when they are, they are done with goodwill and complete buy-in that maximizes the business returns for the organization at large.

Foster a culture of questioning through hiring and training. Strengthen the organizational capability to think like a beginner and ask fundamental questions by reinforcing these practices with training. Zurich Insurance, for example, ran an intensive summer school for all new data science interns and hires, where they worked through templates that helped them map business processes and better understand the full set of incentives and decision makers in play.

These practices can help managers diagnose and avoid machine learning project failures. But diagnosing issues is only one part of it. Business leaders also need to solve the problems that data scientists discover around misaligned incentives or competing priorities. This requires not only strong sponsorship but also a high level of cross-functional collaboration and alignment, which is where business leaders can excel. ■

Dusan Popovic *is head of data science at Anheuser-Busch InBev, Commercial Analytics Europe.* **Shreyas Lakhtakia** *is a graduate student at Stanford University.* **Will Landecker** *is the former AI ethics lead and data science tech lead at NextDoor.* **Melissa Valentine** *is an associate professor of management science and engineering at Stanford University.*

How Generative AI Can Support Advanced Analytics Practice

Large language models can enhance data and analytics work by helping humans prepare data, improve models, and understand results.

By Pedro Amorim and João Alves

Paul Garland

THE GLARE OF ATTENTION on generative AI threatens to overshadow advanced analytics. Companies pouring resources into much-hyped large language models (LLMs) such as ChatGPT risk neglecting advanced analytics and their proven value for improving business decisions and processes, such as predicting the next best offer for each customer or optimizing supply chains.

The consequences for resource allocation and value creation are significant. Data and analytics teams that our team works with are reporting that generative AI initiatives, often pushed by senior leaders afraid of missing out on the next big thing, are siphoning funds from their budgets. This reallocation could undermine projects aimed at delivering value across the organization, even as most enterprises are still seeking convincing business cases for the use of LLMs.

However, advanced analytics and LLMs have vastly different capabilities, and leaders should not think in terms of choosing one over the other. These technologies can work in concert, combining, for example, the reliable predictive power of machine learning-based advanced analytics with the natural language capabilities of LLMs.

Considering these complementary capabilities, we see opportunities for generative AI to tackle challenges in the development and deployment phases of advanced analytics — for both predictive and prescriptive applications. LLMs can be particularly useful in helping users incorporate unstructured data sources into analyses, translate business problems into analytical models, and understand and explain models' results.

In this article, we'll describe some experiments we have conducted with LLMs to boost advanced analytics use cases. We'll also provide guidance on monitoring and verifying that output, which remains a best practice when working with LLMs, given that they are known to sometimes produce unreliable or incorrect results.

Applying LLMs in Predictive Analytics

Predictive analytics lies at the heart of processes that are increasingly data-driven for many companies. It's rare to find a marketing department that isn't discussing shifts in customer churn predictions and how to react, or commercial teams that aren't considering how to boost next month's sales in response to a dip that's been forecast by predictive analytics. We see opportunities to expand the impact of such approaches by tapping LLMs in the following ways to increase the variety of data used to train and execute models or better communicate with business stakeholders who use predictive analytics outputs in decision-making.

Incorporating complex data types. In the development phase of predictive projects, challenges arise when decision makers regularly consult and monitor data sources that are difficult to incorporate into predictive algorithms. For instance, customer reviews detailing negative experiences are complex to use in churn models directly, yet they have valuable predictive power. To utilize such data in predictive models, significant time must be invested in distilling and structuring relevant information from each source. This leads to a trade-off between the investment to make that data usable and the anticipated improvement in the predictive model performance. Of course, there is already natural language software to help with data structuring, but its use is usually circumscribed to particular cases, such as sentiment analysis.

LLMs can significantly reduce the time invested in data wrangling and make it easier to analyze complex data types. A precise prompt can instruct an LLM to review a given data set for key themes and return its answer as data formatted with standard labels that is then suitable for use by predictive models. (See "Labeling Unstructured Data.") This ability of LLMs to expedite the processing of complex data types — from weeks to mere days or hours — may seem quite simple, but it represents a notable leap forward in the practice of advanced analytics.

Since the rise of LLMs, we have seen that the increased ease of incorporating unstructured data sources into analyses has led to a substantial increase in the share of this type of data in various predictive applications. In a recent project with a telecommunications company that focused on predicting the next best action (NBA) in a debt collection and recovery process, there was an untapped data source: written complaints made by customers that were often linked to this process. Because there was no certainty that a thorough analysis of this information could yield a substantial benefit for the NBA project, it was left unused. However, once the team understood that LLMs could accurately filter and categorize the complaints related to the debt collection and recovery process, it started considering this data source — a source that eventually steered the project substantially in terms of what actions to consider to improve this business process.

Explaining predictions. During the deployment of predictive projects, such as the previously mentioned telco project, we've often faced communication challenges in decoding and explaining the inner workings and consequent outputs of machine learning models. One tool data science teams often use to understand and communicate the relevance of the different input variables to a predicted output is Shapley additive explanations (SHAP) analysis. This analysis can be translated into a visualization that describes the relevance and the directional impact of the input variables on a given outcome. For example, it may be used to understand that purchase frequency is the most relevant variable when predicting churn: The less-frequent customers tend to churn more than the others. However, explaining the findings of a SHAP analysis to colleagues who aren't data scientists is often tricky because of the technical knowledge required.

LLMs may help tackle this communication gap. We've noticed that the data sets scraped from the web to train generative AI models, particularly LLMs, encompass extensive knowledge about machine learning models and the analyses used to explain them. (See "Explaining Prediction Results," p. 44.)

Labeling Unstructured Data

With this prompt, an LLM can help wrangle complex data types.

```
#Prompt template with zero-shot learning, no need to provide examples
to the model for it to reason.

enlarge_data_inputs_template = f"""
"Analyze the following set of customer data {data_type}.
Identify key themes, sentiments, and any specific products mentioned.
Provide the data in the following format:
{{
    "customer_id": <id>,
    "customer_name": <name>,
    "customer_review": <rawdata>,
    "customer_sentiment": <sentiment>,
    "customer_theme": <theme>,
    "product": <product>
}}"
"""
```

Consequently, LLMs can provide useful output in response to a well-crafted prompt that specifies the prediction topic, the analytical model employed, the results of analyses, and the technique used to understand the results (such as a SHAP visualization). This information allows LLMs to articulate a plausible explanation for changes in predictions for decision makers and highlights the main contributing factors.

Sticking to our churn example, we experimented by giving the information listed above to the LLM and prompting it to explain, in simple terms, the most relevant variables. It returned accurate output: "NumOfProducts and Age are consistently the most impactful features across all iterations. This suggests that the number of products a customer has and their age are strong indicators of potential churn. For example, if the distribution of these features changes (like offering more products to customers or the demographics of the customer base aging), it could significantly impact the model's predictions." This output may still sound too technical for some, so we were excited to see that the LLM kept expanding upon the topic and eventually mentioned the business impact of the results of our predictive (banking) churn model: "The bank should consider examining their product offerings and customer engagement strategies, particularly for older customers, as these seem to be significant factors in predicting churn."

Applying LLMs in Prescriptive Analytics

Prescriptive analytics are typically employed for business problems involving limited recourses and multiple decision options, such as in supply chain management. Mathematical programming and optimization techniques are the go-to approaches to solve complex decision-making problems such as production and distribution plans that have myriad possible decisions constrained by finite resources, such as production and transportation capacities. Analytics teams can use LLMs to support and streamline the development and deployment phases in the following ways.

Crafting model mechanics. In our experience, mathematically representing a business challenge with all of its nuances is a formidable challenge. It requires an understanding of decision makers' precise goals. For example, when planning store assortments, do category managers give precedence to profitability or to market share, or try to balance both? Defining the boundaries for the underlying decisions is also very hard: In the store assortment problem, is the shelf space assigned to a category a constraint, or can this limitation be overruled in some situations? Missing these pieces of information in the development phase usually results in ineffective models. It's not uncommon for data scientists to miss something important; the ability to pose the right questions and translate the corresponding answers requires a rare combination of business acumen and analytics expertise.

Recently, we have started experimenting with LLMs to help design the mechanics of optimization models, augmenting the capabilities of analytics translators responsible for these tasks. With a carefully designed prompt, it is possible to instruct the LLM to engage in a conversation that can effectively identify decision makers' understanding of the business problem and write the first version of the prescriptive model. (See "Developing Model Mechanics.") An exemplary prompt to the LLM is: "Objective Clarification: What is the primary goal

Explaining Prediction Results

This LLM prompt can return information that helps business users understand a model's output.

```
communicate_prediction_results_template = f"""
"Your goal is to translate the results and changes in results of a
machine learning model that seeks to predict {problem_description}.

To achieve this, the model {machine_learning_model} was applied and
the following metrics from the training and test split were achieved:
{model_training_results} # These metrics are model specific

The dataset that supports this model has the following structure, and
a sample of a few lines is below:
{dataset_columns}
{dataset_data_sample}


To understand the model results the technique {technique_name} was
applied.
The current results of the model and the results of the {technique_
name} are provided, and then the same results of previous iterations
are also provided.

Current results:
{model_results}
{technique_results}

Iteration-1 (Previous iteration):
{model_results}
{technique_results}

Iteration-2 (Previous iteration):
{model_results}
{technique_results}

Use this information to provide possible explanations for why the
results are changing.
"""
```

Developing Model Mechanics

This prompt can return a starter version of a prescriptive model.

```
prescriptive_craft_model_template = f"""
You are an expert in optimization models and your goal is to help
decision makers to create an optimization model.

Your goal is to write the model.

To create a model, you need to define the following components:
- Decision variables
- Objective function
- Constraints

You can use the following syntax to define the components:
- Decision variables:
    - x = model.continuous_var_list(keys, lb, ub, name)
    - y = model.integer_var_list(keys, lb, ub, name)
- Objective function:
    - model.minimize(model.sum(x))
    - model.maximize(model.sum(y))
- Constraints:
    - model.add_constraint(model.sum(x) <= 100)
    - model.add_constraint(model.sum(x) >= 100)
    - model.add_constraint(model.sum(y) == 100)

The problem you are helping is {problem_description}. Start with 20
questions to understand the problem. Keep asking questions while you
still have doubts about the model.

Finally, before writing the model, explain the model using plain
English, and when the decision maker's feedback is positive, you can
start writing the model.
"""
```

of the supply chain optimization? Is it to minimize costs, maximize profits, ensure timely delivery, or something else?" The prompt guides the LLM to identify and define the decision variables, the objective function, and any constraints of the business problem such that it can generate the mathematical formulation of the challenge.

Alongside the dialogue, the LLM may explain in plain English its understanding of the problem and ask the decision maker for clarifications to sort out missing information. We asked: "We're creating a linear programming model that decides on the optimal number of units of Product A and Product B to produce. ... Before proceeding to write the model, could you please clarify the time required for Product B in both the cutting and finishing departments?" These interactions with the LLM resulted in rapid and accurate output.

Understanding model results. Even if LLMs can help teams craft rigorous and applicable prescriptive models to help decision-making, there is another barrier that, in our experience, is even more severe: the difficulty in deciphering the solutions such models produce. This complicated task often leads business users to distrust the results. Moreover, we have frequently seen technical teams spend many more hours than were budgeted to go back and forth with business teams explaining the results. There are numerous underlying reasons, but one rather obvious one is that decision makers have a more subtle approach to decision-making than algorithms, which often get stuck on corner solutions. For example, if the prescriptive algorithm finds a small savings in changing a supply chain network, it will suggest this movement independently of all of the change management efforts that such an action may imply.

LLMs can be an interesting aid when deploying prescriptive analytics to help teams understand model results. Analytics teams can feed into generative AI the mathematical notation representing the prescriptive model as well as the internal metrics, such as the unused capacity (or slack) in constraints that did not limit the solution and the opportunity costs associated with each limiting condition. (See "Provide Insight Into Model Workings," p 46.)

Decision makers can then ask questions to understand the results that interest them, and the LLM can explain these results in plain English until the user is satisfied. This conversation enables generative AI to identify and explain areas where the model's trade-offs may seem counterintuitive to decision makers. Moreover, this method facilitates the collection of additional feedback from decision makers, which can be used to refine the model's mechanics.

Such an approach can be beneficial for all company stakeholders responsible for the success of analytics initiatives. In our experiments with this use case, we understood that increasing the level of autonomy for business owners would have a drastic impact on their sense of empowerment and control over the quality of the proven prescriptive methods. For technical teams, the fact that they didn't have multiple conversations with decision makers about topics that an automated system could instead explain was a huge relief.

A classic example of necessary interaction between analytics and business teams around prescriptive analytics projects is the need to understand why a given algorithmic run is not yielding a solution that respects all defined constraints. See how smooth that interaction can be based on the answer we obtained after asking what the source of infeasibility was in a retail distribution problem: "The primary constraint leading to the infeasible state is the insufficient supply to meet the combined demand and minimum stock requirements. The total available

supply from all warehouses is not enough to satisfy the demands of the retail stores."

Monitoring Model Quality and Business Impact

Companies should already be employing processes to monitor the performance of their advanced analytics models to detect errors and drift resulting from, for example, changes in variables or the business environment that deviate from the model's original assumptions. However the quality of LLMs' output (especially in the absence of additional techniques that further constrain or check that output) is, by design, somewhat unpredictable. The integration of LLMs introduces additional opportunities for errors due to the potential for random or objectively false responses.

The best approach for controlling the output quality of the LLM will depend on which opportunity to integrate generative AI with advanced analytics is pursued. When incorporating complex data types into predictive models, the approach may be relatively straightforward: Companies can investigate the end result and make note of the improvement (or not) of the accuracy metric of the prediction task after adding the unstructured data sources. In the case of using LLMs to explain predictions and understand prescriptive results, technical teams have to heavily test the prompt and the answers that are being generated for different possible questions that business stakeholders may pose. Finally, for crafting model mechanics — an opportunity that is mainly focused on augmenting the development phase of prescriptive models — the outputs of LLMs will always have to be supervised by modeling experts, who must review them critically.

Despite these caveats, bear in mind that the success of any organizational analytics effort depends on how relevant it is to the enterprise, as measured by usage frequency, adoption rate, and internal customer satisfaction. By allowing decision makers to interact with analytics models, the integration of generative AI could lower barriers to adoption, ease change management, and promote trust in outputs.

The quality of LLMs' output is, by design, somewhat unpredictable.

Provide Insight Into Model Workings

This prompt can generate output that reveals and explains trade-offs made in optimization models.

```
prescriptive_understand_model_template = f"""
You are an expert in optimization models and your goal is to help
decision-makers to understand the results of an optimization model.
The problem you are helping is {problem_description}.

Your goal is to explain the results of the following model:
---
{model}
---

The model results were:
- Objective value: {result}
- Decision sets of variables values (variable name, value):
 {decision_variables_values}
- Constraints values, slacks, and shadow prices (constraint name,
value, constraint slack, shadow price):
 {constraints_information}

Wait for the decision maker to start asking questions about the
results. Then, explain them using plain English. Finally, ask
the decision maker if the explanation is clear. If it is not, ask
questions to refine your understanding and explain the results until
the decision-maker is satisfied.
"""
```

As a result, we would expect to see improvements in these metrics.

USING LLMs TOGETHER WITH ADVANCED analytics tools can increase efficiency by streamlining the labor-intensive processes of explaining the validity of predictions and developing prescriptive models. LLMs can also make analytics more effective by aiding in the incorporation of complex data sets for predictive modeling and in understanding prescriptive model outputs. By leveraging the potential of unstructured data sources and identifying opportunities for model refinement, companies can enhance the quality of their outcomes.

We know from current AI and analytics practice that creating multidisciplinary teams that involve both business owners and data science specialists is essential to take full advantage of these opportunities. Thanks to the accessibility provided by LLMs' natural language capabilities, integrating generative AI into analytics should empower business users to take a more active role in the development and monitoring of analytics applications. ■

Pedro Amorim *is a professor at the University of Porto, partner at LTPlabs, and coauthor of* The Analytics Sandwich.
João Alves *is a senior digital manager at LTPlabs.*

Managing Data Privacy Risk in Advanced Analytics

Cybersecurity techniques that keep personal data safe can limit its use for analytics — but data scientists, data owners, and IT can partner more closely to find middle ground.

By Gregory Vial, Julien Crowe, and Patrick Mesana

"HOW CAN WE PROTECT THE PRIvacy of our customers' personal data while leveraging that data via AI and analytics?" This question reflects a growing internal dilemma as companies pursue advanced analytics and artificial intelligence.

The troves of data that customers' ever-more-digitalized lives produce can be a rich source of insight for organizations using advanced analytics tools. At the same time, this data is a deep source of concern to IT staffs committed to meeting both regulatory agencies' and consumers' expectations around data privacy. Both are important objectives — but meeting them simultaneously requires confronting an inherent conflict. Increasing data privacy in the context of analytics and AI involves using techniques that can reduce the utility of the data, depending on the task and the privacy preservation technique chosen.

The issue is one that an increasing number of organizations will face as the fields of analytics and AI continue to quickly evolve and lead to the widespread availability of an array of tools and techniques (including turnkey and cloud-based services) that enable organizations to put data to work more easily than ever. Meanwhile, customers have increasing expectations that companies will take all necessary precautions to protect the privacy of their personal data, especially in light of reports of large-scale data breaches covered by mainstream media outlets. Those expectations are backed by regulations on personal data and AI across the globe that make it critical for companies to keep personal data protection practices in compliance.

Paul Garland

The Nuances of Protecting Personal Data

Fundamentally, data privacy is about assessing the probability that one or more attributes, or pieces of information, about an individual whose data has been anonymized and included with others in a data set can be used to re-identify that specific individual. Some of these attributes are obvious:

Direct identifiers that enable almost immediate identification include name and Social Security number. *Quasi-identifiers* do not generally enable the identification of a single individual on their own, but their uniqueness or their combination with other attributes may do so. For example, the combination of a person's age and their address may enable their re-identification. Or consider a data set held by a bank's fraud alert team on customers' card transactions. That data set contains both direct identifiers (such as the customer's name) and quasi-identifiers (such as credit card transaction information).

In the context of analytics and AI, quasi-identifiers are often highly valuable because they can help organizations uncover shared characteristics and patterns that may help them better find or serve customers. But even seemingly innocuous quasi-identifiers, such as marital status, can be combined with other pieces of publicly available information to re-identify a specific person. Consequently, companies are already being challenged to go beyond protecting just personally identifiable information and consider how to protect quasi-identifiers as well.

Finding the optimal solutions to the privacy-utility conundrum will also require a broader understanding of data privacy throughout the organization, beyond IT and cybersecurity functions. Managers seeking to better understand the scope of options available in balancing data privacy with utility should be broadly familiar with the array of approaches available. Each has its own advantages and disadvantages, with varying implications for data privacy and data utility. (See "Five Approaches to Preserving Data Privacy.")

Privacy Versus Utility Trade-off

To understand how organizations are confronting the complex matter of protecting personal data in their care while also leveraging it for analytics and AI, we'll look at initiatives recently undertaken at National Bank of Canada. (Note that Julien oversees artificial intelligence at the bank; Gregory and Patrick have studied the organization's practices.) Founded in 1859, National Bank is one of the largest financial institutions in Canada. Like its competitors, it must comply with stringent federal and provincial regulatory requirements. Customers trust that National Bank manages their money and the wealth of personal data they share with the bank (when they execute transactions or apply for loans, for instance) with the utmost care.

As a financial institution, National Bank considers customer trust to be its greatest asset, and so it has built a culture in which protecting the privacy of its customers' data is a core value. In addition to driving significant efforts and investment in cybersecurity and organizationwide training, it has also increasingly prioritized analytics and AI. Here, new techniques and approaches increase the potential to leverage personal data to improve services for customers. This increasing use of AI techniques also requires heightened protection efforts, given that new approaches can also be used to compromise the privacy of personal data.[1]

Data protection had traditionally been treated as a security matter that was the responsibility of cybersecurity experts at National Bank. Under this logic, personal data protection would be guaranteed using tried and proven techniques. However, some of those techniques may not readily achieve the required balance between data privacy and data utility. For example, cybersecurity teams can encrypt entire files, but doing so prevents data scientists from being able to use the data contained within those files. Using a more granular approach, direct identifiers could be protected using tokenization (to achieve de-identification), leaving the data science team able to leverage quasi-identifiers, but this does not address the risk of re-identification associated with those quasi-identifiers. To simultaneously satisfy requirements for both data privacy and data utility, teams must find a common ground that allows them to move beyond techniques that favor an either/or approach. In the case of National Bank, we have identified three important steps that contribute to its ability to achieve this objective.

STEP 1: Bridge the gap between IT and data science. In most organizations, cybersecurity and AI/data science teams don't work together. Each has its specialty, and trying to put data to work requires collaboration between experts who tend to work in silos. National Bank realized that this division led to inefficiencies, frustration, and an overall lack of mutual understanding of teams' respective priorities and concerns, and it set out to mitigate the issue. Managers fostered close collaboration between cybersecurity experts and AI delivery team members — including those in roles such as AI architect, data scientist, machine learning engineer, and data engineer — to evolve their competencies and skill sets in each other's domain of expertise.

One illustration of the importance of building this mutual understanding is the example of using synthetic data, where there may be a probability of

Five Approaches to Preserving Data Privacy

Each approach to preserving the privacy of personal data will have an effect on the degree to which the data set remains useful for AI and analytics.

APPROACH	DESCRIPTION	TYPICAL APPLICATION	TYPICAL USE CASE	SEEN BY REGULATORS AS	IMPACT ON DATA USABILITY
Masking	Hide attribute values in whole or in part with modified characters	Direct identifiers	Credit card numbers, email addresses	De-identification	**High** (loss of original information)
Tokenization	Replace sensitive attributes with nonsensitive substitutes (tokens)	Direct identifiers	Social Security numbers, bank account numbers	De-identification	**High** (regain usability through detokenization)
Data anonymization	Remove or modify (for example, swap or generalize) personal information to prevent re-identification	Quasi-identifiers	Health care records, location data	Anonymization	**Moderate to low** (original properties of the data can be closely approximated)
Data synthesis	Create new data that mimics the properties of the original data set without personal information	Quasi-identifiers	Data science research, data sharing	Anonymization	**Low** (original properties of the data can be preserved)
Data encryption	Convert data into nonreadable, unstructured text using an algorithm and an encryption key	Entire files	Secure storage and transmission	Anonymization	**High** (only possible through decryption)

re-identification, depending on the type of algorithm used to generate the data, the data used to train the system, fine-tuning of the parameters, and the attributes to which this approach is applied.[2] This marks a significant departure from the use of techniques such as data encryption, which provide great security at the expense of any data utility. National Bank's cybersecurity and AI delivery teams worked together to develop a common understanding of both the issue and the fact that they would have to evaluate the potential for personal data to be de-anonymized against the degree of utility provided by synthetic data. This process gave the cybersecurity team insight into how re-identification techniques are now increasingly rooted in data science, while the AI team improved its understanding of the work that cybersecurity teams do to ensure the proper sharing and use of personal data.

STEP 2: Formalize and document data-privacy decision-making. Decisions regarding data privacy need to be clearly motivated and justifiable to regulators in the eventuality of an audit — a scenario in which organizations must demonstrate that they have done everything they could to protect the privacy of customer data. That means they must be able to justify why they decided to use a given privacy preservation technique over another in a given situation.

Collaboration between cybersecurity and AI delivery teams at National Bank has led to ongoing efforts to quantify the impacts of various approaches on data privacy and data utility to better inform such decisions. (See "Modeling Data Privacy and Data Utility," p. 50.) Teams simulate audits on data sets that have been protected using different data privacy approaches and parameters to calculate the probability of re-identification within those data sets. At the same time, they evaluate the utility of those data sets based on the same approaches and parameters. For example, certain data anonymization techniques work by making quasi-identifiers more general (such as substituting broader income brackets for actual income values). While that increases the privacy of customer data, it's important not to compromise all of the nuances contained in the original data set that make

it valuable for the organization. Using the example of income brackets, those should be neither narrow enough to enable re-identification nor too broad to be useful in analysis. In evaluating these two variables simultaneously, data managers can quantify and document them to make an informed decision in specific data-sharing contexts. The key here is that the combination of data privacy and data utility can be acknowledged as a risk factor that can be mitigated with sufficient confidence.

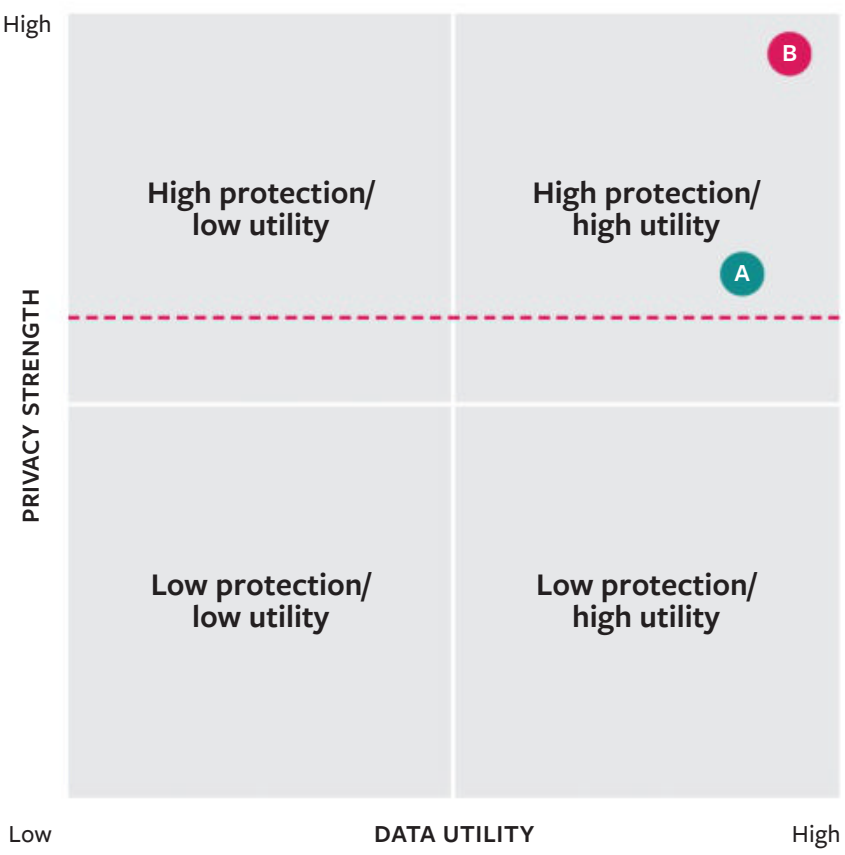
STEP 3: Keep informed on technology, regulations, and evolving threats. As one might expect, data privacy regulations don't prescribe an approach; they mandate an outcome: keeping people's personal data secure. While regulations may vary across jurisdictions, they generally define criteria that can be applied regardless of the scenario under consideration, such as what constitutes data anonymization. Then organizations are responsible for devising data protection strategies that meet those criteria. Given the rapid advances in de-anonymization practices by bad actors, data privacy is a moving target. Organizations need to understand the risks associated with how they protect data, beyond the bare minimum required by regulation. So it's essential that they proactively stay up to date on not only regulations but tech developments.

One way that National Bank addresses this challenge is by reducing the distance between the legal teams that have visibility into upcoming regulations and the AI teams that work with data. This can happen early on by including legal experts in discussions in which project team members explain their data needs to the internal owners of the data. The data owners are typically connected to the legal team and can bring legal in to discuss how the needs of a particular project fit in with the existing data governance framework at the bank.

The bank also collaborates on multiple projects with universities and academic researchers who specialize in data privacy and security. This gives the bank's relevant teams access to cutting-edge scientific knowledge on recent techniques to support their own research and development while advancing knowledge to incorporate into their practices. Similarly, academic researchers also find collaborations with industry valuable because they often lead to more practical work with real-world impact.

Modeling Data Privacy and Data Utility

The figure below is a simplified example of visualizations used at National Bank in deciding on trade-offs between privacy versus utility. Dots A and B represent two different data protection techniques, with dot B predicted to better protect personal data than dot A. The red line represents the baseline compliance with legal requirements for personal data protection.



Ramping Up Data Privacy for Data Science Practice

For many companies that are investing in AI and analytics in the hope of gaining valuable business insights from their customer data, the implications for potential exposure of personal data are just emerging. To effectively manage the trade-offs between data privacy and data utility, we suggest the following practices and approaches.

Teach data privacy as part of data literacy. In many organizations, data literacy is still uneven or lacking, and substantial efforts are still required to address this issue.[3] In the context of data privacy, this challenge is even more glaring: One cannot assume that managers who possess basic data literacy skills have a clear understanding of data privacy concepts such as direct identifiers and quasi-identifiers. They also need to understand the risks of re-identification associated with these identifiers, and the characteristics of the approaches typically used to address these risks.

In the case of National Bank, data governance and data literacy initiatives have been implemented for several years, and, like many other financial institutions, the bank was an early adopter of analytics and other

Multiple stakeholders must contribute to informed decisions on protecting personal data.

approaches to improve decision-making. However, it has had to further develop data privacy literacy as a competency that transcends specific domains of expertise. Experts working in cybersecurity, legal, and AI delivery all had their own understanding of data privacy, its implications for their department, and the approaches available to mitigate its associated risks. For example, members of one team would use terms referenced in regulations (such as de-identification), while data scientists would consider specific technical approaches to data privacy (such as k-anonymity or differential privacy). Fostering collaboration across functional units has been an important part of developing data privacy literacy at the organizational level.

Treat data privacy as a business issue. Developing data privacy literacy as an organizational capability also supports an organizational culture in which data privacy is treated as a business issue, not a purely technical matter. That is, there should be a widespread understanding that the imperative to manage personal data carefully is founded on the need to maintain customer trust — and is thus directly related to the bottom line. Connecting the dots between personal data protection, company reputation, and performance is possible only if personal data protection is explicitly acknowledged as a strategically relevant matter that requires dedicated time and resources.

Doing this may require a rethink at companies that have relegated data privacy to the cybersecurity team. However, data privacy involves a variety of stakeholders with different expertise and concerns, and they all must be able to communicate in a common language and participate in discussing and designing data privacy strategies.[4] Cross-disciplinary collaboration is essential — and when something is understood to be a business issue, it is understood as being important to everyone in the organization.

Formalize your approach to balancing data privacy and data utility. As we described above, multiple stakeholders must contribute to informed decisions on how to protect personal data in a given situation. Establishing a systematic approach to working through the issues and communicating the implications of different privacy techniques for data utility and data protection is essential.

National Bank has accomplished this by evaluating the impact on data privacy and contextualizing it against data utility. That has enabled the creation of tools, such as the matrix presented earlier, that communicate the privacy and utility implications of different conditions in given situations over and above regulatory requirements. With such visualizations, data managers don't need to know the intricacies of data privacy preservation techniques, but they can see their outputs and rely on their data literacy skills to ask pertinent business questions. In addition, computation of quantitative measures can be integrated into the analytics/AI model creation/validation pipeline so that it becomes part of a standard process. This maintains an awareness of the need to continuously improve data privacy approaches as re-identification techniques continue to improve as well.

DATA PRIVACY SHOULD BE AN IMPORTANT area of concern for organizations managing personal data. But it is also a complex business matter that has important technical implications. The quick evolution of the science of data privacy, coupled with modernized regulatory requirements, makes it challenging for companies to optimize their strategies on this front. Ultimately, as data managers gain a deeper understanding of this topic, they can design and evolve strategies that will help them optimize both data privacy and data utility, forgoing the idea that we necessarily need to sacrifice one for the other. ■

Gregory Vial *is an associate professor in the Department of Information Technologies at HEC Montréal.* **Julien Crowe** *is senior director of artificial intelligence at the National Bank of Canada.* **Patrick Mesana** *is a doctoral candidate in the Department of Decision Sciences at HEC Montréal.*

REFERENCES

1. C. Dwork, A. Smith, T. Steinke, et al., "Exposed! A Survey of Attacks on Private Data," Annual Review of Statistics and Its Application 4 (March 2017): 61-84.
2. T.E. Raghunathan, "Synthetic Data," Annual Review of Statistics and Its Application 8 (March 2021): 129-140; and S.L. Garfinkel and C.M. Bowen, "Preserving Privacy While Sharing Data," MIT Sloan Management Review 63, no. 4 (summer 2022): 7-10.
3. T.H. Davenport and R. Bean, "Action and Inaction on Data, Analytics, and AI," MIT Sloan Management Review, Jan. 19, 2023, https://sloanreview.mit.edu.
4. Raghunathan, "Synthetic Data," 129-140; and Garfinkel and Bowen, "Preserving Privacy While Sharing Data," 7-10.

Reprint 65433. For ordering information, see page 4.

Acing Value-Based Sales

To get the best returns on innovative products, collaborate with customers to define and share the commercial opportunity.

By Marco Bertini, Oded Koenigsberg, and Todd Snelgrove

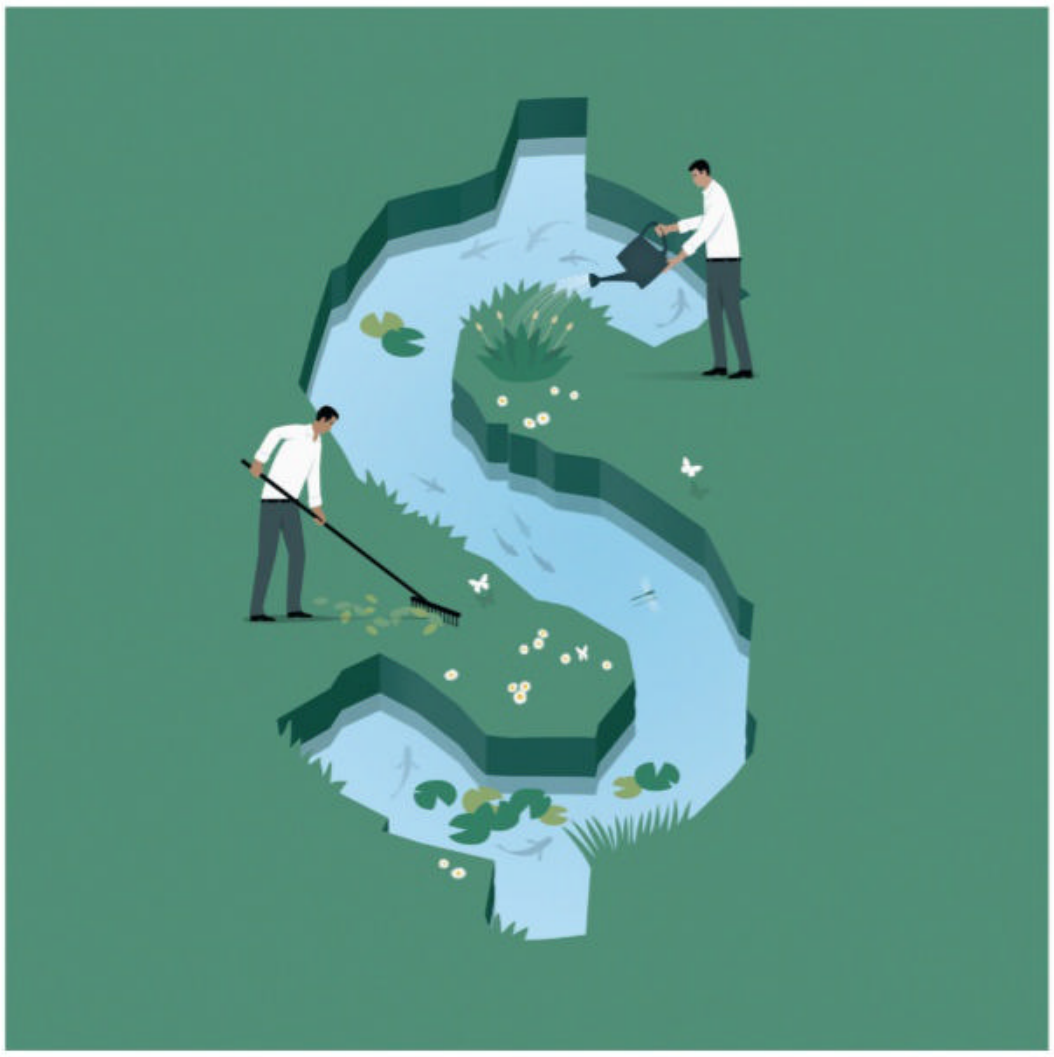
IMAGINE THAT BURST OF ENTHUSIASM when a senior executive unveils a plan that promises a significant and lasting impact on the organization's financial performance. "Our new product creates more value for our customers than anything else on the market, and we should get paid accordingly," they proudly declare. "If we measure and communicate that value precisely, then we can finally get the return we deserve."

The sales team buys into the logic. Now armed with a superior product, a sophisticated value calculator, and a host of new selling arguments, they quickly turn the senior executive into a prophet. Price, margin, and profit all improve significantly in the first quarter of implementation, with a smaller but still important increase the next.

A year later, however, few traces of the highly touted program remain. The selling arguments have gone stale. The state-of-the-art value calculator is dismissed as theoretical and complex. The early financial gains have evaporated as the sales team, under constant pressure from fierce buyers, returns to aggressive discounts to win the volume and customers lost at the outset.

Such sagas are an open secret in B2B markets around the world. Many companies embrace value-based selling programs, but few succeed in a significant, sustainable way. They invest heavily in product development, but frustration spreads when customers don't understand the points of differentiation, never mind wanting to pay for them. The dream outcome of selling on value rather than cutting good deals never materializes, leaving a trail of demotivated salespeople and missed opportunities. Worst of all, lacking a clear return, companies postpone or cut back on investments — which reassures competitors that they can get away with mediocre offerings.

In our experience, value-based selling initiatives often start with the wrong premise. Senior executives assume that their company's differentiated offerings deserve a higher price. In turn, this sense of entitlement lulls them into thinking that the only challenge is to quantify and communicate every advantage. Faced with hard numbers, they reason, any customer in their right mind would happily pay a premium.



Mark Airs / Ikon Images

However, in most commercial relationships, value is not at the sole discretion of the seller but cocreated with customers who play an active role in realizing desired outcomes. It follows that successful value-based sellers not only prove the added value of their products and services but also gain commitment from customers and foster a shared understanding of the opportunity.

Our framework, developed over years of helping companies achieve long-term returns from product development, grounds a five-part process for successful value-based sales in three key principles essential to such

initiatives: commitment, understanding, and proof. (See "Engaging Customers in Value-Based Selling," p. 54.) Establishing *commitment* is required at the start to motivate customers to work with the sales team, and again at the end to reward customers appropriately. The drive to reach mutual *understanding* shapes the second step, where salespeople educate customers about the range of benefits offered by their solutions, as well as the fourth step, where they translate expected customer savings and gains into a compelling business case. At the center of our framework and the sales process is *proof* — because, ultimately, companies attract and retain customers only when they can quantify the value that their solutions add.

Best-in-class value-based selling realizes that "showing customers the money" is not enough. In this article, we will expand on each of these steps and explain how to tackle them in your business.

MOTIVATE: Set the Context for Value Cocreation

The potential of a commercial opportunity may be clear, but it will fail to materialize if your company or the customer lacks the motivation to collaborate in creating and sharing value. The two key dimensions to motivation, the first step in our framework, are the balance of power between sides and the customer's financial wherewithal.

Conventional wisdom calls for the party with the greater bargaining power to shape the deal in its favor. This results in higher prices and more stringent conditions when the seller has the advantage, and lower prices and more generous conditions when the buyer does. In this view, any codependence between the two parties relies more on force or acquiescence than on actual need or mutual choice.

But this perspective is myopic. The overconfidence that comes from having leverage tempts the dominant side to appropriate as much value as it can in the short term, meanwhile neglecting the opportunity to create greater value in the long run by working with the other side. When your company has the advantage, you should exercise self-restraint and pass up short-term gains for longer-term value creation. An excessively short-term focus fosters an arm's-length relationship that keeps high prices top of mind for customers and encourages them to look for alternative suppliers. If the customer switches suppliers, you lose not only a revenue opportunity with a partner but also insights into product performance and customer benefits that can accrue only from long-term, repeated engagements. When the customer has the advantage, look to reduce its leverage by developing other customer relationships. You can also look for ways to influence the customer's own customers to demand the value that you provide, such as via a stronger branding effort. The "Intel Inside" tag on personal computers is an example of how to drive awareness and end-user demand.

Buyers and sellers, then, have the best basis for joint value creation when the balance of power between them is relatively even. Yet, even with a level playing field, value creation will not happen unless customers have the money, time, and talent to purchase and properly implement a demonstrably better solution. Innovative offerings, especially those whose value accrues over time, are irrelevant if customers cannot afford them or cannot get them to perform to their potential.

Affordability was a barrier for global tire manufacturer Michelin when it introduced new products to its industrial customers. Its solution was to change the price metric such that, instead of selling customers the tires they needed for their fleets of trucks, Michelin retained ownership of the tires and invoiced on a per-mile-driven basis. The company not only improved access to better tires by spreading the burden of payment across time but also improved relationships: Low-consumption customers assumed less commercial risk (since they paid only when trucks were on the road) and could rely on Michelin's expertise in the selection, repair, replacement, and recycling of tires to ensure that their fleets always had optimal equipment at hand. Michelin and its customers improved productivity and profitability because the common metric became a way to align their interests and not merely a clever way to pay for tires.

Motivating customers to participate in joint value creation is important, but it isn't the only element of context-setting that your company needs to attend to. It's also essential that customers understand the full extent of benefits that your products or services can bring.

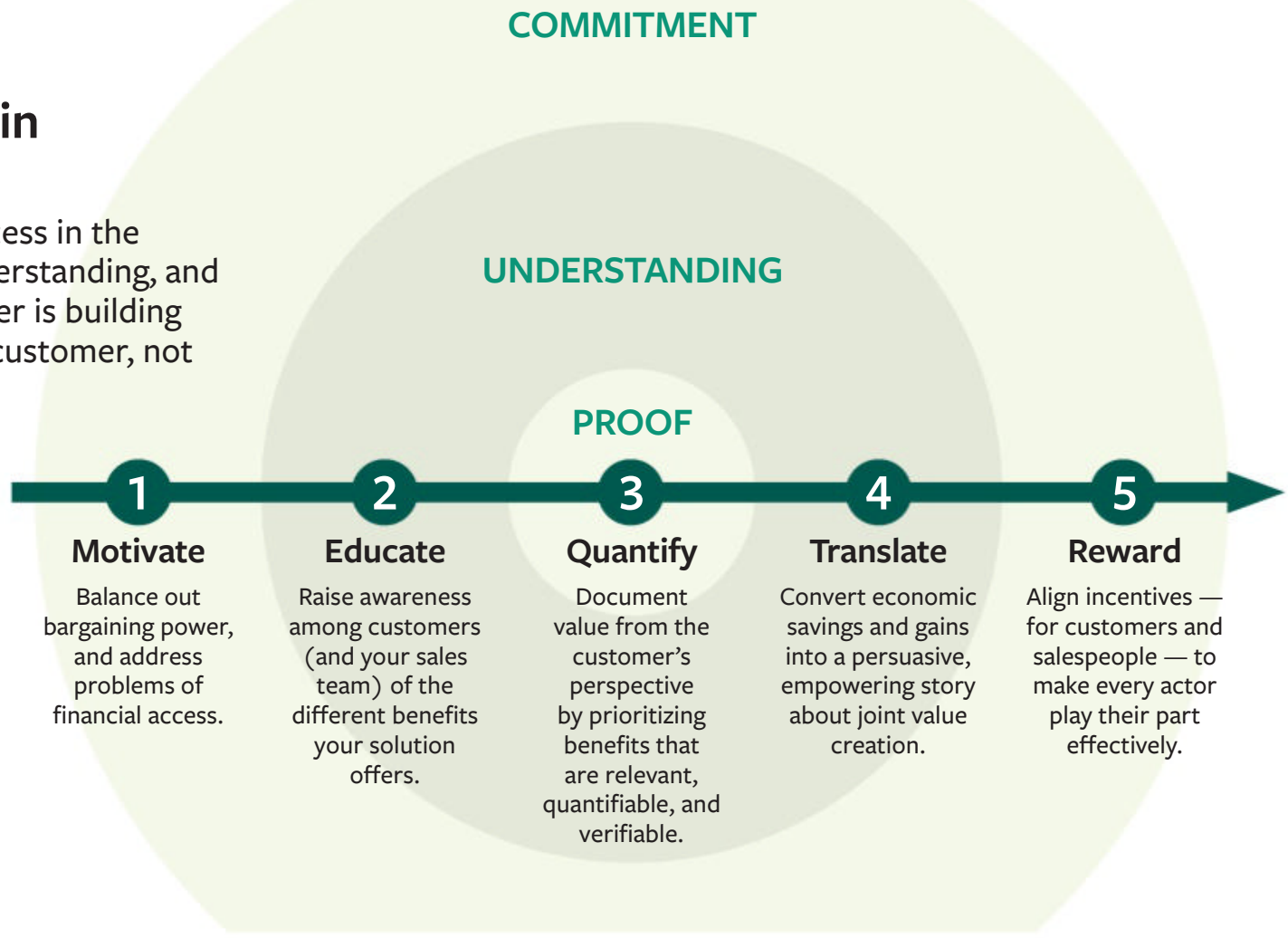
EDUCATE: Raise Awareness of Customer Benefits

A commercial opportunity relies on how well your company and your customers understand the benefits that drive value creation. The assistance of marketing and communications teams is critical at this point, in particular because benefits often accrue not only when a product is used but also at acquisition and disposal. At this step, the goal is simply to make customers question the wisdom of pushing on price.

Scot Forge, a 130-year-old metals forging company, succeeds with its value-based selling efforts because its teams think about value and price separately. They achieve this hard separation by focusing on a prospective customer's biggest needs and challenges. In a joint process of discovery and exploration, Scot Forge and its customers create a standard vocabulary to define benefits

Engaging Customers in Value-Based Selling

Rooting the five-step sales process in the principles of commitment, understanding, and proof underscores that the seller is building a deeper relationship with the customer, not merely seeking to close a deal.



and how the two parties will then quantify them.

They aren't starting this process with a blank slate: Scot Forge maintains a long list of potential benefits for its products and services that is based on how the benefits are initially defined and measured. This standardization facilitates the discussions with buyers by providing a starting point. Many customers can intuitively judge whether the initial estimate of a benefit is too high or too low, but if asked directly, they would struggle to provide an accurate estimate of their own.

This process narrows down the list of benefits to what matters to each customer. Agreement on how to quantify the benefits is about more than mathematical formulas. The parties also agree on how accurate measurements should be. In reality, "close enough" is acceptably accurate for many measurements.

Such a process allows salespeople to set aside their assumptions regarding "this is what our solution does" to reach consensus on "this is how our solution helps you." Salespeople in companies with a strong engineering culture may know, for example, that their new equipment has a smaller footprint and half the moving parts compared with a similar product offered by their main competitor. But outperforming competitors on one or more attributes does not guarantee that every customer will benefit from it.

Salespeople may also know that their solution has helped other companies reduce labor costs. But that benefit may not matter as much to a customer whose shops employ union labor or a customer that relies on offshore manufacturing to keep labor costs competitive. In another case, a company's solution may yield a step change down in energy costs only after a certain volume threshold is reached. Salespeople need to know whether a prospective customer can at least meet that threshold, which would make the benefit worth prioritizing.

The number of benefits, and the mix between tangible and intangible benefits, is specific to an individual customer or, at most, to a segment of similar customers. Benefits such as less downtime, lower per-unit labor costs, or reduced energy consumption are easier to understand and quantify than potential increases in demand or higher brand awareness.

The business and technology consulting firm West Monroe succeeds with an approach similar to Scot Forge's. The firm delivers value by advising organizations and implementing new technologies and ways of working, but those benefits are achieved only by changing people's mindsets and behaviors. It's about helping clients understand how new and different ways of working can deliver value across stakeholders.

Indeed, West Monroe saw an opportunity across the consulting industry. "Approximately half of our competitors' clients didn't feel that their consulting firm's value delivered was greater than the fees charged," said Casey Foss, West Monroe's chief commercial officer. "That's a big opportunity, but competing on financial results required us to be better at creating, optimizing, and communicating our value to clients." The firm now tracks value-creation scores that rate the value

received compared to fees paid. Those scores can be 32% higher among clients with whom it jointly quantifies value, because both parties are focused on the initiatives expected to have the greatest impact.

We've found very few companies engaging with customers the way Scot Forge and West Monroe do. We often encounter salespeople who lack the basic training to engage buyers in discussions about value and acknowledge the right trade-offs. Professional buyers are quick to point out that there are cheaper solutions available in the market; of course, there are always going to be competitors offering better deals. The issue is whether salespeople can help customers see — and quantify — the results they will be sacrificing to enjoy a better price.

QUANTIFY: Solve the ABC Problem

Proof sits at the center of our framework because it is the fundamental part of any business transaction. Every commercial opportunity must be expressed in the medium customers use to gauge value. In most business markets, this medium is money. Your company's task, then, is figuring out how your solution yields savings and gains — a challenge that we call the "ABC problem" because the path leads from product or service *attributes* to customer *benefits* to shared *cash*.

Solving the ABC problem is a numerical exercise that feeds on data and creativity. The data element is clear: Your company cannot substantiate claims without evidence of how an attribute or feature generates a benefit and how that benefit impacts the customer's finances. This evidence, however, often arises when a solution is used, which implies that the customer must be enticed to share information. These enticements can include exclusive access to future innovations, consulting or other insights, temporary discounts, or other rewards.

But sellers should not underestimate the role of creativity and discovery. It is not always clear how the technical specifications of a given attribute can be expressed numerically as a benefit. One common misconception is that tangible benefits such as higher productivity or reduced energy consumption are relatively easy to convert into dollars and cents. Sometimes the tangible benefits are indirect. When a manufacturer introduced a new generation of machines that had a significantly lower operating temperature than its predecessors, its sales and marketing teams struggled at first to express the advantage in monetary terms. It turned out that the lower temperature meant the equipment required less lubrication, which translated into measurable savings for customers in terms of materials, labor, and waste disposal.

Another misconception is that intangible benefits such as brand equity or ease of use are nearly impossible to articulate in monetary terms. While there is some truth to this, blanket assumptions are dangerous if they lead companies to give up on quantifying intangible benefits too soon and use them as add-ons to sweeten the deal rather than as inputs to a better price. In many cases, intangible benefits represent a large part of the added value of an innovative product. If your company undersells or underestimates them, then it effectively renounces a good part of the margin it stands to make from the sale.

In our view, the impact of a benefit depends on the extent to which it fulfills the following criteria.

- **Relevant:** A customer needs to understand how it can capitalize on a given benefit. Superior towing capacity and cold-weather performance are wonderful innovations, but they mean nothing to a company that does little towing and never operates in cold weather. If your sales team doesn't understand which benefits are relevant for a prospective customer, they risk dumping features into the offer indiscriminately or making every offer "all you can eat" for simplicity's sake.
- **Quantifiable:** Is the relevant benefit observable and countable? Ideally, the customer can easily perceive a direct relationship between cause and effect. A heavy vehicle manufacturer may, for example, document that it has twice the towing capacity of its competitors and has superior cold-weather performance. But finding a concept and an equation to quantify a benefit is not sufficient. Quantification is possible only with a reliable flow of credible data. Does your company and the customer have the means in place to monitor performance and collect data? Do you have a central collection point that can serve as a single source of truth? The better the data, the better the company can estimate the value of the benefit.
- **Verifiable:** The customer still needs to realize the estimated benefits. This depends on cooperation between the buyer and seller to share data and agree on the calculations. Following up regularly after the sale gives your team a chance to observe how successfully the customer has implemented the solution and verify that the estimated benefits exist. You can work on corrective actions if the implementation or the benefits are falling short of expectations. Follow-up visits also provide additional data you can use to adjust claimed benefits for future sales. One way to support verification is to make the outcomes visible through a tool such as a value calculator tailored to the customer's business.

These three criteria should be considered sequentially when deciding where to focus the sales effort. First, your company must identify the benefits that are relevant to the customer in question. From these relevant benefits, priority goes to the ones that are quantifiable. Then, among the quantifiable relevant benefits, the emphasis should be on the verifiable ones.

These criteria also determine the nature of the sales effort. The more a benefit satisfies all three criteria, the greater the underlying certainty that it is achievable. Your sales team can then focus on making direct comparisons to alternative and presumably inferior solutions. Indirect comparisons and general statements such as "a 1% improvement in this benefit yields $X million in savings" may not be meaningful to most customers, given that estimated and actual results often deviate significantly from an average. One buyer told us that if his company were to achieve all the savings that potential sellers have promised him in those "1%" claims, its savings would be greater than its actual costs. Such cynicism has prompted some buyers to insist on guarantees or other forms of risk sharing that compel sellers to take a more active role in ensuring that they achieve the estimated benefits.

For benefits that are relevant and quantifiable but not verifiable, your company owns the burden of proof. You can use mechanisms such as benchmarking, pilot tests, or samples to offset perceived risk by the customer. Finally, for benefits that are relevant but neither quantifiable nor verifiable, look to market research for support. When introducing data that comes from customer surveys, though, it is important to make sure respondents have sufficient prior experience with the product to validate their judgments, and an incentive to answer truthfully. You want credible and accurate data at hand as you move toward building a strong case in the next step.

TRANSLATE: Turn Impact Into a Compelling Business Case

Specialty chemicals company Borealis conducts value workshops with its potential customers to help them connect the prioritized list of quantified benefits to their day-to-day operations. "We build a customized model, and we present it step by step, validating our assumptions during the conversation," said Paolo De Angeli, head of customer value management. This not only gets buy-in from customers but also enables Borealis and the customer to reach a common understanding of the differential value that the company creates. It also allows both parties "to identify opportunities for further value creation by working together as partners," De Angeli added. The common metrics keep conversations focused on quantified value delivered.

We talk about translating impact because numbers seldom speak for themselves. Calculating the monetary gains from adopting a particular product is a must, but wrapping dry figures in language, visuals, and interfaces that resonate with customers is also critical. The more familiar the story feels, the more agency the customer will feel it has. Your company can use that as a call to action to cocreate value.

The most powerful business cases are stories that travel well, meaning that they retain meaning beyond the people directly involved with negotiating a deal — such as the senior executive who has final control over budget approval, the local manager who is responsible for implementing the company's innovative product, and so on. This extended audience usually has limited time, a limited understanding of the context, and its own vernacular. Therefore, successful business cases feature the following.

- **A small number of key messages:** Benefits are not necessarily additive; less may be more. The choice of what benefits to communicate, in which order, can affect a buyer's value perception. For example, juxtaposing a strong or resonating benefit with a less impactful one can diminish the perception of the more valuable benefit.
- **The right amount of detail:** It is better to share just enough detail to enable a decision and keep the information specific and relevant to the customer. When discussing the business case directly, your company can share more information if the listener indicates a need for it or asks a question. To help the business case travel better, you could include a brief Q&A in an appendix that anticipates and answers additional questions. Refraining from telling the whole story may be difficult for sellers who know their product back to front, because they may struggle to put themselves in the position of someone looking at the product for the first time. Salespeople need to resist the temptation to smother the customer with extraneous details.
- **Clear next steps:** The business case should include an operational framework that specifies the KPIs and the tools and data to measure and validate them. Creating a joint agreement for documentation and validation keeps your company involved after the sale closes and draws the attention of senior management beyond the buying team that negotiated the deal.

REWARD: Align Interests With Smart Incentives

Any value-based selling effort must be coupled with strong incentives for the company and customers to

create and allocate value over an extended period. With respect to the customer, incentives operate at two levels: the price metric and the price points.

Michelin and its customers in the mining sector agreed on a price metric — cost per mile driven — that aligned with the way the parties create value together. The more material the mine can transport more frequently over longer distances, the more productive the mine will be. Contrast this approach with setting a price per product and offering discounts ex post to adjust for observed outcomes. The money paid and received may be the same, but the latter process is far less collegial.

Early-stage motivation and late-stage incentives both belong in the commitment layer of our recommended process because the price points — the split of the value pie — need to reflect and maintain the stable balance of power established at the outset of the process. The split should reflect the proof behind benefits and the time frame of the agreement. A split that favors one party too strongly will eventually give the other party a disincentive to create additional value.

After moving through the first four steps of our framework, salespeople should be well prepared to counter requests for lower prices or larger scopes by making the appropriate price-value trade-offs. But they also need incentives to negotiate those trade-offs in a way that secures the best price for your company, which is usually neither the highest possible price nor a low price that cedes too much value to the customer. Specific incentives can include rewards for negotiating a price within a target range or for growing the value pie, or rewards that are inversely proportional to discount levels. This marks a significant shift for companies that have trained and rewarded their salespeople to meet revenue or volume goals rather than goals related to profit or prices.

How the buyer pays also makes a difference, and these payment methods and terms must be codified in advance. Let's say a seller creates $10 million in mutually validated value for a customer over a two-year period. No matter how appreciative buyers are, few of them will have the willingness or even the authority to write a lump-sum check based on that value. Sellers can overcome these issues with financing options or with price metrics based on consumption, outcomes, or other parameters.

Closing the deal doesn't mark the end of the transaction but rather the initiation of a new phase of the relationship. Tracking actual performance versus projected value allows your company to identify and resolve problems early and build a robust database for benchmarks. To ensure that it has a natural role in ongoing validation, for example, Michelin builds its models together with its customers, based on a common vision and goals.

IN OUR EXPERIENCE, COMPANIES THAT implement our process for value-based selling start with two initiatives in parallel. First, they improve their ability to provide proof. Rigorous quantification is the core task for a reason. If your company does nothing else, this change will improve your sales efforts. Thinking through what is relevant, quantifiable, and verifiable ensures that business cases have a robust basis.

Second, they need to appoint a full-time value leader who is independent of sales and marketing. Marketing professionals tend to overshoot "soft" benefits such as brand and expertise. Apart from being wrong, this loses the trust of salespeople and complicates collaboration: Marketers blame sales for not being able to sell value; salespeople blame marketing for sending them into battle unprepared. If there is collective ownership of value throughout the organization, no function feels accountable for it, and each applies its own definitions and tools.

Appointing an independent value manager, then, marks the first move toward establishing commitment and understanding, the two outer layers of the framework. Marketing and communications support the process in promoting understanding, while finance and strategy have influence over commitment. The value leader also provides a powerful cross-check or counterweight by representing the customer's interests as your company develops its plans for each step of the process.

Perhaps the most important task of the value leader, though, is to align the organization on a shared and respected definition of value: how to define it, measure it, and adapt it for an individual customer. This is the ultimate prerequisite for successful value selling because the term *value* can mean everything and nothing at the same time. Continuous attention to the ABC problem makes value a habit and enables salespeople to adapt the definition to an individual customer's needs.

Companies that make our framework the basis for their value selling can also expand their target markets. We refer to this as *key account management at scale*. A standard definition of value, supported by the right tools and mindset, enables your company to define and quantify value for any customer — even small ones — and decide the best way to share that value. ■

Marco Bertini *is a professor of marketing and the director of open executive programs at Esade, in Barcelona, and a senior adviser to Globalpraxis.* **Oded Koenigsberg** *is executive dean, Dubai; professor of marketing; and chair, marketing faculty, at London Business School.* **Todd Snelgrove** *is the founder of consultancy Experts in Value and a fractional VP of sales at Sales Xceleration..*

Reprint 65425. For ordering information, see page 4.

Find a Circular Strategy to Fit Your Business Model

Products and services that maximize use and reuse of materials and other resources can be both growth opportunities and sustainability measures.

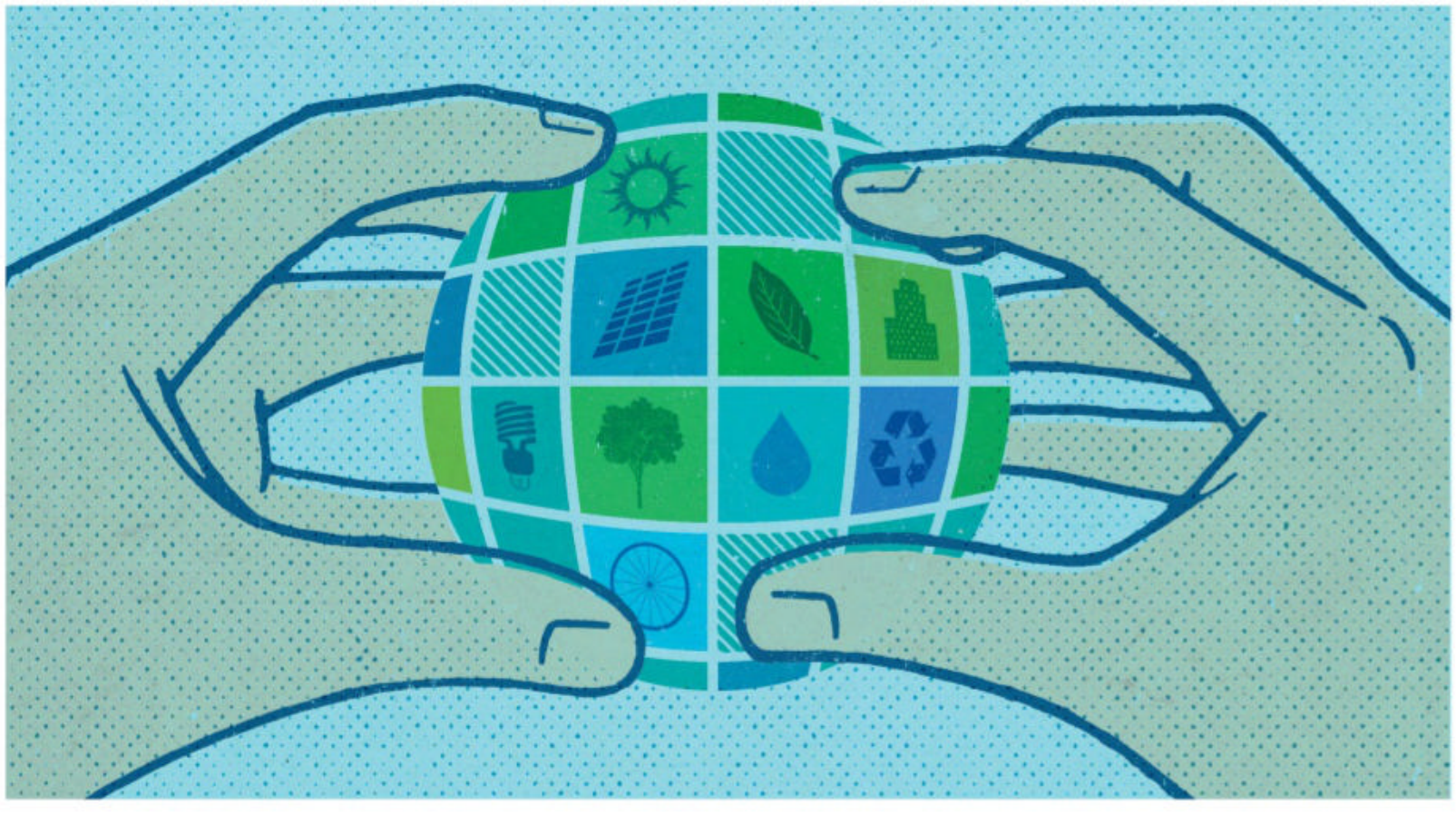
By Samsurin Welch and Khaled Soufani

COMPANIES SEEKING TO MEET AMBITIOUS sustainability goals to reduce fossil fuel consumption and reduce waste must go beyond a shift to renewable energy. Roughly half of global emissions are linked to the production and consumption of goods such as food, packaging, buildings, and textiles.[1] Accelerating decarbonization will require that we rethink the materials and services sourced from suppliers, the distribution and use of products by customers, and what happens to products at end of life.

That's where circular models come into focus. They aim to optimize the use of material resources in organizations and thus help reduce carbon emissions and mitigate strain on natural systems. For businesses, circularity promises to enhance competitiveness by improving bottom lines through resource efficiency and boosting resilience against resource market volatility through closed-loop supply chains. Moreover, compelling circular offerings, such as refurbished furniture or new items made from recycled materials, appeal to customers striving to meet their own sustainability goals.

Despite these advantages, circular businesses remain uncommon, largely due to the challenge of identifying and operationalizing suitable models that align with a firm's overall strategy and capabilities.[2] In this article, we draw on our research into organizations implementing these models to explain four different routes to circularity, and key considerations for effectively implementing them.

The practical principles underlying circular business models can be stated as simple goals: *use less*, *use*



Dan Page / theispot.com

longer, *use again*, *use differently*.[3] (See "Four Approaches to Circularity," p. 61.) Most circular models emphasize one of these paths, but often in combination with others.

Extend Product Lifespan

Reuse, resale, repair, or refurbishment are among the use-longer tactics that help companies maximize the usable life of products, reducing both waste and consumption of new resources. The challenge for companies that pursue this approach is preserving the benefits customers derive from newly manufactured products. Some ways to operationalize this include the following:

Design products for longevity. To enable repair or refurbishment, products must be designed for durability, modularity, and disassembly. That may require using more durable materials, avoiding adhesives in assembly, and ensuring that standard tools can be used in repairs.

The Dutch Ministry of Defense (MoD) has adopted a circular approach to the uniforms it provides, cleaning and repairing them to extend their usable life. It redesigned uniforms to allow easy replacement of components like damaged sleeves without destroying the whole item. This not only reduced waste but also led to an annual savings of 8 million to 10 million euros that would have been spent on new purchases and disposal.[4] Smartphone manufacturer Fairphone similarly designed phones with modular, swappable components. Customers wanting a better camera or battery can simply upgrade those components rather than having to buy a new phone, delaying obsolescence.

Designing for disassembly and repair does involve trade-offs. More durable and modularized components or detachable fasteners can increase manufacturing complexity and cost. Despite the complexities, this approach is a key principle in enabling circular models from electronics to buildings.[5]

Establish repair and refurbishment capabilities. Extending product lifespans requires building repair and refurbishment capacity, either in-house or through partnerships. The Dutch MoD chose the latter route. With item volumes in the hundreds of thousands, the MoD needed an industrial-scale repair facility. Rather than build it themselves, it partnered with Biga Groep, a social enterprise that provides employment to members of disadvantaged communities. The two parties collaborated to develop industrial-scale repair capabilities. This partnership grew over time as Biga Groep gradually expanded its capabilities, also creating new jobs.

It is crucial that product development teams work closely with those responsible for repair and refurbishment to better understand design requirements for longevity and repair. When shoe manufacturer Vivobarefoot launched a repair and refurbishment program, it learned that its shoes could be difficult to repair. It subsequently started having its design teams engage with repair partners to gain feedback for future product.

Manufacturers must also plan for longer-term availability of spare parts and repair services, including how repair and refurbishment will be provided, such as through expanded warranty plans or as chargeable services that could be additional revenue streams. Buyback plans can encourage customers to return products to manufacturers for refurbishment and resale.

Depending on the sector, independent companies may already offer resale, repair, and refurbishment services, for example, in the mobile phone market. Manufacturers can support this market with spare parts, repair kits, and manuals. This may become a regulatory requirement under emerging Right to Repair laws across different jurisdictions. For example, the EU's right to repair rules require manufacturers to make spare parts available for professional repairers for up to 10 years for certain product categories. A looming issue on this front is support for software components in so-called smart products.

Leverage data and digitalization. Data can enable more effective maintenance, repair, and refurbishment. Some manufacturers use internet-connected sensors and predictive analytics to monitor product health and schedule maintenance. Data analysis can also shed light on customer usage, common faults, and wear-and-tear patterns to improve product designs.

Additionally, digital platforms provide a channel for resale of used and refurbished products, especially if supported by comprehensive data on product condition and history, enhancing transparency and trust in the resale market. Brands may find that their used or refurbished products are already available on marketplaces like eBay, Vestiaire Collective, or Back Market.

THE RESEARCH

- This research was based on qualitative case analyses of organizations that have implemented circular economy models.
- Data collection involved a mix of structured interviews with company managers, review of company documents, and review of in-depth case studies.
- The analysis identified key patterns in successful adoption of circular models and the strategic decisions, challenges, and outcomes associated with circular economy practices.

Reclaim and Regenerate Resources

For companies following the use-again approach to capture value from waste materials, the distinction between technical and biological resources is key. Originally defined in the cradle-to-cradle model,[6] biological resources include raw materials such as food, wood, cotton, and other animal or plant-based materials, while technical resources include metals, plastics, synthetic chemicals, and other artificial materials. The two require different approaches.

Because biological resources are in principle renewable, the focus should be on recovering and reusing nutrients to regenerate natural systems and ensure sustainable production. Plant and animal waste can be composted and used to enrich soil health, and biogas can be recovered for energy production through anaerobic digestion.

Remanufacture or recycle technical materials. Metals used in electronic products, such as copper, lithium, aluminum, and gold, could theoretically be reused indefinitely, particularly if recovered into their elemental forms. On the other hand, plastic is a nonrenewable resource, and plastic packaging has limited reuse. A use-again approach through recycling can be a viable option to ensure circularity of technical materials. U.K.-based packaging manufacturer Charpak uses recycled plastic as raw material for its products, which are themselves recyclable.

Turn waste into revenue. Use again can also involve turning waste resources into products through intermediate processing. For example, materials science companies Biohm and Ecovative are using agricultural residues to grow mycelium, the rootlike structures of fungus, as bio-based materials for packaging, fashion, and building materials. Biotech company Entomics Biosystems uses organic waste to cultivate insect larvae, which is turned into a sustainable protein source for animal feed in livestock farming.

Given the vagaries of fashion, unsold inventory can be a significant source of waste for apparel manufacturers. Luxury brand LVMH has spun out a platform, Nona Source, where others can purchase deadstock yarn, fabric, and leather. Aside from avoiding waste, this enables smaller, independent designers to purchase high-quality material that they typically would not have access to. Rype Office remanufactures used office furniture into high-quality, low-carbon products for corporate clients. As furniture makes up to 30% of embodied emissions in commercial buildings, shifting to remanufactured items can help organizations reduce their carbon, resource, and waste footprint.

Design products with recycling in mind. Like use longer, use again starts with product design. Most products today suffer from low recycling rates, and most recycling is really *downcycling* into lower-quality products, such as garments into rags or insulation materials.

One reason for this is that many products are not designed for recyclability. They are often composed of a mix of materials, each with different degrees of recyclability and requiring different treatment. For example, clothing often contains a mix of natural and synthetic fibers; electronic devices comprise a host of different metals, plastics, glass, and other materials; and packaging frequently includes labels or linings. Ideally, products should be made from single materials, such as monofiber textiles or polyethylene terephthalate (PET)-only packaging, although this is often not possible.

Designing for use again involves several approaches. The first step is to maximize the proportion of recycled or regenerative materials that are themselves recyclable or biodegradable in order to reduce the use of virgin materials and production of waste. This includes eliminating toxic elements. For example, flooring company Tarkett's carpet brand Desso and carpet manufacturer Hook & Loom designed out toxic chemicals and materials that would limit the reusability or biodegradability of their materials.

Designers must also consider disassembly. Just as with use-longer models, products should be designed to be easily taken apart so that components can be remanufactured or recycled. The E.U.-funded Buildings as Material Banks project applied this principle, developing standards and tools such as the Reversible Building Design Protocol so that buildings can be deconstructed to recover and reuse materials.

This design approach will often require development of new materials and processes that meet quality and performance criteria. Charpak addressed this by working with innovation partners on modular design and replacing black plastic with 100% recyclable gray plastic. This required a full life-cycle analysis of product design for recyclability, including considering the post-use phase and managing trade-offs. For example, a certain percentage of virgin material was required to meet functional and aesthetic requirements.

Design products to be easily taken apart for remanufacturing or recycling.

Four Approaches to Circularity

Organizations developing circular business models typically begin with one of the following paths, though they may use several in combination.

 **EXTEND** *USE LONGER*	 **RECLAIM** *USE AGAIN*	 **MAXIMIZE** *USE DIFFERENTLY*	 **MINIMIZE** *USE LESS*
Extend the usable life of products and materials	Recover and reuse materials and energy	Make idle products or capacities available for use	Optimize resource efficiency and minimize waste in operations and product life cycles
Examples: more durable materials, product repair, refurbishment	**Examples:** component remanufacturing, recycling, waste-to-energy	**Examples:** products-as-a-service, sharing, renting, virtualizing	**Examples:** energy efficiency, zero-waste, predictive maintenance

Establish reverse logistics systems and partnerships. Designing circular products does not necessarily result in actual circularity in the absence of effective reverse logistics systems. Modern supply chains are typically optimized for producing and distributing products into the market, with less emphasis on getting products and materials back for reuse. For manufacturers transitioning to a circular model, the challenge extends beyond merely ensuring effective recycling of their products to also guaranteeing certainty of supply of raw materials to create their products.

Accomplishing this may require innovation in the waste-processing segment of the value chain, including technologies and processes for effective collection, sorting, and recycling of waste into new raw materials. In some segments, like post-consumer plastic waste, direct recovery of products can be challenging. Plastics enter the waste management stream and need sorting and processing before they can be reused. Charpak addressed this in the U.K. by partnering with city councils, waste management providers, and reprocessors to establish a local circular plastics loop. Under the Cambridgeshire and Peterborough Waste Partnership (RECAP) program, household plastic waste is collected, sorted, cleaned, and sent to be reprocessed into new materials, which Charpak then sources.

Rype Office grappled with uncertainty in its supply of used furniture, which typically depended on office clearance events. To mitigate this, it established take-back channels in partnership with building retrofitters and offered its customers leasing and buy-back options. For situations where remanufactured furniture is insufficient to meet client demands, Rype Office recommends new pieces that are easily remanufactured over multiple uses, further ensuring future supply.

Maximize Product Use

In many cases, consumers may not feel the need to own a product. Companies can sell access rather than ownership; sharing, renting, and service-based approaches can replace product sales. These models boost circularity by increasing asset utilization and aligning incentives to encourage customers to use less, use longer, and use again. The use-differently approach also can broaden access to products that might otherwise be out of reach for some consumers.

Service-based models come in different forms, such as very short-term rentals of cars (Zipcar) or designer clothing (Rent the Runway), or peer-to-peer platforms like Hello Tractor, which enables farmers to rent idle equipment to other small-scale farmers who can use the machinery to improve crop yields. Another model is providing products-as-a-service, where providers retain ownership and customers pay based on usage. Lighting multinational Signify offers a pay-per-lux service: The company will install, maintain, and decommission lighting equipment, including reuse, repair, and recycling.

Design services based on customer jobs-to-be-done. Identifying the customer's goals can inspire innovative circular models.[7] As Signify illustrates, customers want reliable and efficient lighting, and a "lighting as a service" model both circumvents the complexities of managing assets and reduces capital expenditure on equipment. Similarly, many city dwellers prefer to not own a car, but they occasionally need transport with more flexibility, convenience, and range than a bike, bus, or train. This idea underpins shared mobility models like Zipcar — rather than multiple households owning cars that are infrequently used, customers can access shared cars on demand.

Leverage digital platforms. Hello Tractor's sharing model illustrates how data, connectivity, and digital platforms can enable sharing, rental, and service models to monitor and optimize assets and connect supply and demand. Hello Tractor's mobile app connects small

farmers with tractor owners to schedule and manage bookings and payments. The company also helps tractor owners outfit their equipment with low-cost internet of things devices to track location, fuel levels, tractor operations, and maintenance needs. That data can support new services such as route optimization for tractor owners and help them to better manage their equipment assets. Importantly, Hello Tractor understands that access to tech is uneven among small farmers, so it also engages local booking agents to help with bookings and aggregate local demand.

Align incentives. Service design critically hinges on aligning incentives between the service provider and the customer around economic, functional, and sustainability factors. To achieve circularity goals, providers and customers must actually use items responsibly and efficiently.

Signify's lighting-as-a-service contracts include delivery of performance metrics for lighting quality and energy efficiency. Customers are concerned with only the reliability and performance of lighting. Extending the use of equipment and materials through monitoring, durable design, reuse, repair, refurbishment, and recycling saves costs for the provider. A service-based model also facilitates closed-loop reverse logistics systems, as the provider retains control and chain of custody of the assets.

Minimize Resource Use

Figuring out how to use less of any resource should always be a priority when conservation is the objective. With this goal, organizations can think broadly about how they can minimize the use of energy, materials, and water in operations.

Focus on efficiency. Look for practices and technologies that optimize resource use while minimizing waste. Lean manufacturing principles, for instance, can significantly reduce material waste and energy consumption. Designers can experiment with ways to deliver the same value using fewer materials. Additionally, integrating energy-efficient technologies in operations, such as LED lighting or optimized heating and cooling systems, can reduce the environmental footprint significantly.

Efficiency optimization can also be applied to capacity utilization of assets. Shared resources, like coworking spaces or shared manufacturing facilities, can optimize space and equipment use. Designing buildings for reconfigurability can enable flexibility for multiple uses. In logistics, fleet optimization ensures full vehicle loads, reducing the number of trips and thus lowering emissions and costs. Similarly, demand-responsive services like cloud computing can adapt supply efficiently to meet actual demand, avoiding excess capacity.

Design out waste. Naturally, using less begins with product or process design. Consider packaging: Significant efficiencies can be found here before looking at the product itself. U.K. supermarket retailers Tesco and Waitrose, for example, reduced plastic waste by removing plastic wrapping from some multipack food and beverage items. Food technology companies Mori and Apeel Sciences are developing edible coatings that extend the shelf life of produce, replacing plastic packaging. Bulky liquid products that contain a great deal of water, like cleaning or beauty products, can be reformulated as concentrate or tablets. This reduces product size and weight, which lowers the carbon impact of shipping and the cost and footprint of storage and delivery.

Tap data analytics. Insights afforded by digital tools are as important here as they are in other approaches. Advanced data analytics pinpoint ways to minimize resource consumption and waste production. Ikea reduced food waste by 54% in its in-store restaurants by using smart scales and AI to track and analyze food waste. Ikea also partnered with food platform Too Good to Go to further reduce food waste through selling its surplus at discounted prices.

Key Success Factors for Circularity

Numerous opportunities are available for companies to embed circularity into their products and operations. While approaches may vary, our research has found that successful initiatives pay attention to the following:

Strategic fit. The cornerstone of circular transformation is how it fits with your business strategy. Identify a model that aligns well with your company's strategic positioning and brand ethos and leverages core strengths and values. A chosen model should enhance both sustainability and profitability. For example, a luxury brand should provide premium retail experiences for resale, repair, or refurbished offerings.

Committed leadership and circular mindset. Circularity transformation starts with leadership, mindset, and culture. The company's leadership needs to champion and drive the circularity agenda and provide resources and training to the whole organization. A key lesson is that circularity must go beyond specific initiatives; it requires a mindset where everyone is constantly

Most companies will struggle to build circular capabilities by themselves.

looking at waste as a valuable resource. Sugar processor British Sugar has turned almost every by-product stream from its beet sugar manufacturing process into inputs for new products, including biofuels and industrial and agricultural products, as well as a new business segment for animal feed.

Combined models. The four models — use less, use longer, use again, and use differently — are not mutually exclusive and should be combined to create viable circular offerings. For example, the Dutch MoD recycles items once they can no longer be repaired. A similar model could be applicable for electric vehicle batteries. With electric cars expected to account for more than two-thirds of global car sales by 2030, pressures will increase from mineral scarcity and electronic waste. A more circular approach could keep batteries in use in vehicles through repair and refurbishment (use longer). Once degraded, used batteries could be remanufactured into energy storage systems for renewable power grids (use differently), and finally recycled to recover materials to make new batteries (use again).

Compelling value proposition. Successful circular offerings must go beyond purely environmental benefits to also meet customer needs such as functionality, desirability, aesthetics, accessibility, and economics. Consider the end-to-end value chain design and trade-offs required to ensure customer acceptance of a circular offering. Rype Office recognized that corporate clients would not compromise on furniture quality and aesthetics for lower carbon options, so it provides interior design services to help clients integrate remanufactured pieces into existing office aesthetics. Hello Tractor's network of local agents was in part based on user insight on ensuring accessibility and supporting the needs of small farmers.

Value-based collaborations and partnerships. Most companies will struggle to build circular capabilities by themselves. While organizations may embed circular requirements into green procurement criteria, challenges arise when suppliers cannot meet these demands in the short and medium term. A more effective approach may require co-innovation partnerships with suppliers, peers, and other partners. A prime example is Jaguar Land Rover's RealCar project, where it worked with industrial, academic, and public-sector partners to develop aluminum recycled from postindustrial waste for use in its cars. This required innovation in material chemistry to develop aluminum grades with suitable performance, recyclability, and applicability. A critical key success factor was the approach to collaboration around shared goals for the entire value chain for recycled aluminum.[8]

Journey of learning and iteration. The circularity transformation should be approached as a journey of learning and iteration. The winning circular models are often not obvious, emerging through experimentation and engagement with customers, suppliers, and other stakeholders. Explore what works and what doesn't, and adapt strategies based on these findings. Equally important is understanding and managing trade-offs and unexpected consequences that arise. Starting with small-scale pilots allows companies to test ideas in a controlled environment, learn from these experiments, and then scale up successful initiatives, building capabilities over time.

The journey to circular economy models is complex. However, many brands may already have active markets for resale, repair, or refurbishment of their products, as well as potential partners in different sectors. Beyond meeting environmental goals, engaging in circular models offers businesses opportunities to control customer relationships and experiences, gain valuable market insights, and capture additional revenue. ■

Samsurin Welch *is a research associate at the Circular Economy Centre at Cambridge Judge Business School, University of Cambridge, and senior associate at the University of Cambridge Institute for Sustainability Leadership. He is also cofounder and chief operating officer at HyveGeo.* **Khaled Soufani** *is a management practice professor of financial economics and policy, director of the Circular Economy Centre, and director of the Cambridge Executive MBA Program at Cambridge Judge Business School. He is also a fellow of Clare Hall.*

REFERENCES

1. Ellen MacArthur Foundation, "Completing the Picture: How the Circular Economy Tackles Climate Change," 2019, www.ellenmacarthurfoundation.org.
2. K. Soufani and C. Loch, "Circular Supply Chains Are More Sustainable. Why Are They So Rare?" Harvard Business Review, June 15, 2021, https://hbr.org.
3. Adapted from M. Geissdoerfer, M.P.P. Pieroni, D.C.A. Pigosso, et al., "Circular Business Models: A Review," Journal of Cleaner Production 277 (Dec. 20, 2020): 123741.
4. K. Soufani, T. Tse, M. Esposito, et al., "Bridging the Circular Economy and Social Enterprise: The Dutch Ministry of Defence and Biga Groep," The European Business Review, March 18, 2018, www.europeanbusinessreview.com.
5. Circular Electronics Partnership, "Circular Electronics System Map: An Industry Blueprint for Action," Sept. 22, 2022; and A. van Stijn, L.C.M. Eberhardt, B. Wouterszoon Jansen, et al., "Environmental Design Guidelines for Circular Building Components Based on LCA and MFA: Lessons From the Circular Kitchen and Renovation Façade," Journal of Cleaner Production 357 (July 10, 2022): 131375.
6. M. Braungart and W. McDonough, "Cradle to Cradle" (New York: North Point Press, 2002).
7. S. Hankammer, S. Brenk, H. Fabry, et al., "Towards Circular Business Models: Identifying Consumer Needs Based on the Jobs-to-Be-Done Theory," Journal of Cleaner Production 231 (Sept. 10, 2019): 341-358.
8. P. Cassell, I. Ellison, A. Pearson, et al., "Collaboration for a Closed-Loop Value Chain: Transferable Learning Points From the REALCAR Project," PDF file (Cambridge, United Kingdom: University of Cambridge Institute for Sustainability Leadership, Jaguar Land Rover, and Novelis, January 2016), https://www.cisl.cam.ac.uk.

Reprint 65415. For ordering information, see page 4.

How to Come Back Stronger From Organizational Trauma

Traumatic events are destabilizing. In their aftermath, leaders can help individuals and teams recover and grow.

By Payal Sharma

IT IS A SOBERING REALITY OF LIFE TODAY that many organizations across sectors and industries will face trauma. My institution, the Lee Business School at the University of Nevada, Las Vegas (UNLV), became one of them on Dec. 6, 2023, when a shooting on campus profoundly changed our community.

Trauma is extraordinary, uncontrollable, and overwhelming to those who experience it.[1] Its impact is devastating, and it leaves survivors with ongoing pain and loss that cannot be overstated. When we experience trauma, it shatters our belief that the world makes sense, and we consequently feel less safe, less in control, and more vulnerable.[2] However, psychology research has also found that as they recover from trauma, individual survivors can experience *post-traumatic growth* (PTG).[3] This process doesn't minimize the suffering or psychological challenges that survivors encounter but rather taps the "rich and remarkable resources, creativity, and success of the human spirit to adapt, cope, and survive," in the words of psychologist Ronnie Janoff-Bulman.

While research into PTG has focused on individuals, the possibility that organizations might experience similar effects after a traumatic event is intriguing. This

Carolyn Geason-Beissel/MIT SMR | Getty Images

article aims to provide an overview of current thinking about organizational trauma and explores the question: In the aftermath of trauma, how might leaders help their organization move forward to collectively survive — and even engage in learning and growth that surpasses its pretrauma state?

Understanding Organizational Trauma and PTG

Events that cause trauma for organizations are catastrophic, life-threatening, or life-altering, and disrupt core functions; their causes can be either internal or external.[4] They include incidents such as workplace violence, natural disasters, and terrorism. Organizational trauma is both distinct from and related to what we define as a crisis: Many traumatic events are crises, but not all crises are traumatic events. An organizational crisis — such as a financial scandal, major product recall, or consumer boycott — may challenge members' beliefs about the organization and its mission, but trauma is akin to an earthquake that displaces members' sense of the world and their collective place in it.

As referenced above, research over the past two decades has found an emergent pattern in survivors' stories: transformative psychological changes as a result of struggles in the aftermath of trauma.[5] PTG reflects the human capacity to adapt and think expansively following horrific experiences, rather than to regress and turn inward.[6] It is a long-term process of development and change over time. In contrast to trauma, which is fast and forced on those who experience it, PTG is slow and trauma survivors are in control.

PTG is distinct, though it is related to organizational resilience and crisis management. A resilient organization perseveres in the face of adversity, absorbing the trauma and then returning to, or restoring, the status quo before the trauma: It bounces back.[7] However, an organization that experiences PTG bounces *forward*; it follows "a trajectory of increased positive functioning," in the words of organizational psychologist Sally Maitlis.[8] Resilience may be an outcome of PTG. Compared with crisis management, a tactical reaction to acute, threatening circumstances, PTG involves constructing new positive assumptions about safety and stability, after trauma has demolished old core beliefs. Like *kintsugi*, the Japanese art of mending pottery with molten gold, it acknowledges our history and our strength to put ourselves back together.

How Organizational Trauma Affects Individuals

For leaders to understand how to facilitate organizational growth after trauma, they first need to understand the psychological impact of trauma. Research has long demonstrated that in the aftermath of trauma, individuals experience adverse effects on their thoughts, feelings, and behaviors. They become hypervigilant, and experience intrusive thoughts and high levels of emotional arousal, involving intense anger, anxiety, depression, and fear. These emotions may trigger involuntary fight, flight, or freeze responses. People have difficulty focusing on work and thinking deeply; they often cannot control outsize emotional reactions to everyday situations and interactions; and they may be exhausted as nightmares or insomnia rob them of sleep. They may become isolated by believing that no one external to the traumatic event can understand what they are going through. They may engage in catastrophic thinking and expect only the worst to happen. They may cope with new beliefs that they are helpless or powerless to change their present reality by dissociating from it.

Following a traumatic event in an organization, people are likely to feel overwhelmed with doubts about their safety on the job, to question the meaning of their work and their commitment to the organization, and to withdraw from others.[9] These responses are a function of how trauma destroys what Janoff-Bulman calls *assumptive worlds*, or three core assumptions we have about the world and ourselves:

1. **We are safe.** We believe the world is benevolent and a safe place, that good things happen more than bad things, and that most people have good intentions and can be trusted.
2. **We have control.** The world is meaningful and is ordered along just and logical principles. We can maximize control over outcomes and prevent misfortune.
3. **We are deserving.** We as individuals are worthy, fortunate, and deserve good things. We know bad things happen, but we don't expect these bad things to happen to us.

As long as nothing disrupts these assumptions, they are strongly held and confirmed by our experiences. Trauma has such a devastating impact on our psyches because it shatters these beliefs that kept us feeling safe, in control, and protected. At the collective level, the shattering of organizational members' assumptive worlds may manifest in their feeling distracted from the work and overwhelmed with thoughts about insecurity, risk, and instability in their work lives and identities. What then can leaders do to encourage their organization to recover, heal, and grow?

Because our assumptive worlds reflect simple and generalized views, growth after trauma entails moving away from such absolutist views and building new core

beliefs that are more nuanced.[10] These evolved understandings — termed *schema changes* — serve as cognitive toolkits to help us effectively navigate life and work. Thus, in the aftermath of trauma, leaders can facilitate their organization's growth through the collective construction of revised beliefs that are more complex and concrete about safety, control, and protection. The goal is to create a rebuilt assumptive world that integrates the trauma, including an accompanying collective sense of vulnerability and disillusionment — but allows the group to move forward and cease fully defining itself by the trauma.[11] This is also the path from seeing oneself as a victim to seeing oneself as a survivor.

How Can Leaders Facilitate PTG?

Research into PTG in individuals has found three pathways for growth.[12] The first, seeing strength through suffering, is recognizing that one has sufficient strength to tolerate and continue on despite pain and suffering. The second pathway, generating psychological preparedness, can be seen as building the capacity to face future traumatic events with equanimity. The third pathway, crafting greater meaning and purpose, arises as survivors of trauma often experience a change in perspective and reprioritize what they most value; they stop taking what's important for granted. Below, I'll discuss how each of these can be activated in an organizational context.

Define our organization by our collective strength. Just as individuals can realize they possess courage, confidence, new competencies, and abilities after trauma — all of which create pathways toward seeing new possibilities in their lives — so can organizations. At the individual level, the research describes tendencies people have initially to avoid thinking about a traumatic event, given its overwhelming nature. However, and it can take a long, long time, individuals can start to use what is called *approach oriented* coping. In doing so, they realize that the trauma was incredibly painful, and they can see their own strength as they navigate its aftermath and accept that it happened.

We lived this pathway at UNLV. For context, the campus shooting occurred six years after the city of Las Vegas experienced its own trauma when a gunman opened fire at a country music festival, killing 58 people and injuring more than 800 others. Within hours of what has been called the deadliest mass shooting by an individual in U.S. history, visitors and locals in Las Vegas started using the hashtag #vegasstrong on social media. According to online reports, the hashtag honored both the horrific tragedy for the city and the fact that people did not want the city to be defined by it.[13]

In a similar vein, within days of the 2023 shooting, UNLV utilized the hashtag #UNLVstrong in its public messaging. This reflected our capacity to be strong collectively as we processed the aftereffects of such a devastating event. In subsequent communication sent both internally and externally to the university community, the school's leadership consistently referred to our organizational strength.

What this does for organizational members is to help them integrate the trauma into their psyches *and* to recognize that they are not defined alone by the associated pain, but by their strength as well. Stated differently, when leaders take control of the narrative after a traumatizing experience for their communities and frame this in ways that do not speak to members as powerless or helpless, these members can start to heal and recover. We don't have a choice when trauma happens to us in our work settings, but we can choose to see our ability to survive in its aftermath.

Leaders can encourage discussion and reflection by sharing a series of open-ended questions, such as, "What abilities, if any, do we now realize we have as an organization?" (See "Engaging the Organization in Post-Traumatic Growth.") Open-ended questions can help support members as they consider new abilities and possibilities for the organization as it moves forward and heals. Further, on a day-to-day basis, leaders ideally can create time and space for employees to document and share how, as a collective, they are achieving healing-oriented small wins that make a big difference in recovering from trauma and feeling strong.

Generate organizational preparedness. A second process toward PTG involves moving from feeling stuck in a hypervigilant state to developing more resilient assumptive worlds. For leaders, this means acknowledging that future traumatic events are inevitable, they are part of our world, and we can survive them. Doing so helps to restore a sense of control that may have been obliterated by the trauma. We understand we are vulnerable, and we can preemptively feel prepared for what lies ahead.

This all speaks to the importance of leaders proactively managing expectations around the reality of

Building greater meaning and purpose can help survivors focus less on the question of why the trauma happened to them.

Engaging the Organization in Post-Traumatic Growth

As leaders work with organizational members to help the collective move forward from a traumatic event, they can draw on these questions and suggested activities.

DEFINE OUR ORGANIZATION BY OUR COLLECTIVE STRENGTH

Questions for discussion

- How, if at all, do we see our collective fortitude since the trauma?
- What abilities, if any, do we now realize we have as an organization?
- What are new possibilities, if any, that we see for us collectively as we continue to recover?

Practices to implement

- Create virtual platforms — such as Slack channels or electronic bulletin boards — to collect group narratives/anecdotes about sources of pride in the organization in the aftermath.
- Normalize and encourage regular gatherings such as weekly staff or town hall meetings to surface and celebrate small wins focused on healing and recovery. Encourage these discussions to capture day-to-day what employees are doing collectively to help them feel stronger at work.

GENERATE ORGANIZATIONAL PREPAREDNESS

Questions for discussion

- How, if at all, can we as a community understand or accept that the trauma happened so that it is a reality?
- How, if at all, can this understanding help us be better prepared psychologically for a future trauma?
- How, if at all, can we move forward as an organization with these two sets of understanding in mind?

Practices to implement

- Provide members with mental health resources — for example, through employee assistance programs — to help them integrate the trauma into their psyches (rather than avoid it).
- Invite groups within the organization to share examples of how they have processed the shock of the trauma and share the strategies and toolkits they have either newly developed or are existing and have sharpened, on which they can draw when a future trauma happens — as a form of proactive coping.[i]

CRAFT GREATER ORGANIZATIONAL MEANING AND PURPOSE

Questions for discussion

- How, if at all, might our organization change because of what happened?
- How, if at all, does our organization feel connected to broader society and humanity because of what happened?[ii]
- Where, if at all, do we perceive we have control or freedom with reshaping our vision or mission as an organization?

Practices to implement

- Encourage departments to revisit their pretrauma vision and/or mission, and update them to reflect members' insights about their work and its meaning or contributions to society.
- Facilitate a leadership retreat for executives to brainstorm questions such as, "How can our organization's story about our traumatic past inform our better future?" and "What are ways that our organization can serve other organizations that may be facing similar pain and losses?"

trauma at work. While individuals may be most afraid of a catastrophic event occurring, I believe leaders can leverage that fear to help members focus on how they might feel prepared and not entirely caught off guard. Although we cannot predict when and in what form, we know we can expect trauma in our organizations; and such an understanding and acceptance enables us to build psychological "armor" — so that our assumptive worlds feel less destroyed by the trauma in the aftermath.

At UNLV, after the shooting happened, the leadership repeatedly acknowledged that workplace violence can and does happen, and we are not the last community that will be targeted. As sobering as these words were, they allowed us to integrate into our revised assumptive world the acceptance that other traumatic events are likely to happen in the future. Importantly, the purpose of this observation was not to keep our community in a fearful state. Rather, stating this reality allowed us collectively to rebuild our core beliefs about safety, control, and protection with more nuanced and less absolutist understandings about the world and ourselves.

Craft greater organizational meaning and purpose. According to Janoff-Bulman, building greater meaning and purpose — the third pathway of PTG — can help survivors focus less on the question of why the trauma happened to them. With an increased focus on their purpose moving forward, the "why us?" question starts to soften and no longer has control over people's

Leaders can consider how the narrative that emerges around a traumatic event informs their future vision.

thinking. Within management research, it's well understood that individuals and teams thrive when they feel that their work is meaningful and relevant, and they can influence key outcomes.

The goal here is to widen the aperture of our post-trauma understanding to include that we are connected to society at large and to others within our local community. Thus, our healing and recovery can be aided by nurturing and developing these connections in a variety of ways, such as to engage in service to others who similarly have experienced trauma.

After the 2012 shooting at Sandy Hook Elementary School in Newtown, Connecticut, a nonprofit organization called Sandy Hook Promise was formed by family members who lost loved ones in the devastating tragedy. As their website explains, "Our intent is to honor all victims of gun violence by turning our tragedy into a moment of transformation." The organization seeks to educate and inform students and parents about preventing gun violence and identifying early preventive measures.

Leaders can likewise encourage this pathway with members in their organization. They can take time apart as a management team to think more broadly about how the narrative that emerges around a traumatic event informs their future vision. As part of that, considering how the organization might support and help others through similar traumatic experiences can, like Sandy Hook Promise, help rebuild a sense of purpose.[14]

NO ONE IS EVER TRULY PREPARED FOR TRAUmatic events, but by understanding how they impact individuals and organizations, leaders will be better positioned to facilitate healing, growth, and a new understanding of how to move forward. Processing the reality of the trauma while at the same time looking to the future as an organization powerfully enables more complex and concrete perceptions of safety, control, and protection. Then, we can focus actively on where we do have agency: creating renewed meaning and commitment in our work lives. ■

Payal Sharma *is an assistant professor of management at Lee Business School at the University of Nevada, Las Vegas.*

REFERENCES

1. R.M. Vogel and M.C. Bolino, "Recurring Nightmares and Silver Linings: Understanding How Past Abusive Supervision May Lead to Post-Traumatic Stress and Post-Traumatic Growth," Academy of Management Review 45, no. 3 (July 2020): 549-569.
2. R. Janoff-Bulman, "Assumptive Worlds and the Stress of Traumatic Events: Applications of the Schema Construct," Social Cognition 7, no. 2 (June 1989): 113-136.
3. R.G. Tedeschi and L.G. Calhoun, "Post-Traumatic Growth: Conceptual Foundations and Empirical Evidence," Psychological Inquiry 15, no. 1 (2004): 1-18.
4. M.R. Kramer, L. Page, and G. Klemic, "Post-Traumatic Growth in Organizations: Leadership's Role in Deploying Organizational Energy Beyond Survival," Organization Development Review 54, no. 3 (2022): 18-26; and B.N. Alexander, B.E. Greenbaum, A.B. (Rami) Shani, et al., "Organizational Post-Traumatic Growth: Thriving After Adversity," Journal of Applied Behavioral Science 57, no. 1 (March 2021): 30-56.
5. A. Ehlers, R.A. Mayou, and B. Bryant, "Psychological Predictors of Chronic Post-Traumatic Stress Disorder After Motor Vehicle Accidents," Journal of Abnormal Psychology 107, no. 3 (August 1998): 508-519; R.F. Hanson, D.G. Kilpatrick, J.R. Freedy, et al., "Los Angeles County After the 1992 Civil Disturbances: Degree of Exposure and Impact on Mental Health," Journal of Consulting and Clinical Psychology 63, no. 6 (December 1995): 987-996; H.S. Resnick, D.G. Kilpatrick, B.S. Dansky, et al., "Prevalence of Civilian Trauma and Post-Traumatic Stress Disorder in a Representative National Sample of Women," Journal of Consulting and Clinical Psychology 61, no. 6 (December 1993): 984-991; P.B. Sutker, J.M. Davis, M. Uddo, et al., "War Zone Stress, Personal Resources, and PTSD in Persian Gulf War Returnees," Journal of Abnormal Psychology 104, no. 3 (August 1995): 444-452; and R.G. Tedeschi, J. Shakespeare-Finch, K. Taku, et al., "Post-Traumatic Growth: Theory, Research, and Applications" (New York: Routledge, 2018).
6. R. Janoff-Bulman, "Schema-Change Perspectives on Post-Traumatic Growth," in "Handbook of Post-Traumatic Growth: Research & Practice," eds. L.G. Calhoun and R.G. Tedeschi (Mahwah, New Jersey: Lawrence Erlbaum Associates Publishers, 2006): 81-99.
7. T.A. Williams, D.A. Gruber, K.M. Sutcliffe, et al., "Organizational Response to Adversity: Fusing Crisis Management and Resilience Research Streams," Academy of Management Annals 11, no. 2 (June 2017): 733-769; and A.D. Meyer, "Adapting to Environmental Jolts," Administrative Science Quarterly 27, no. 4 (December 1982): 515-537.
8. Tedeschi and Calhoun, "Post-Traumatic Growth," 1-18; and S. Maitlis, "Post-Traumatic Growth at Work," Annual Review of Organizational Psychology and Organizational Behavior 7 (2020): 395-419.
9. P.N. Sharma and M.J. Pearsall, "Leading Under Adversity: Interactive Effects of Acute Stressors and Upper-Level Supportive Leadership Climate on Lower-Level Supportive Leadership Climate," Leadership Quarterly 27, no. 6 (December 2016): 856-868.
10. Sharma and Pearsall, "Leading Under Adversity," 856-868.
11. R. Janoff-Bulman, "Post-Traumatic Growth: Three Explanatory Models," Psychological Inquiry 15, no. 1 (2004): 30-34.
12. R.G. Tedeschi, "Growth After Trauma," Harvard Business Review 98 no. 4 (July-August 2020).
13. C. Morell, "Reflecting on 'Vegas Strong' One Year Later," Nevada Public Radio, Oct. 1, 2018, https://knpr.org.
14. Tedeschi, "Growth After Trauma."
i. P.N. Sharma, J.M. Silvas, and M. Guadagnoli, "Attention Leaders: Are You Losing the Battle With Stress? Arm Yourself With Proactive Coping," Organizational Dynamics 51, no. 2 (April-June 2022).
ii. K. Olson, T. Shanafelt, and S. Southwick, "Pandemic-Driven Post-Traumatic Growth for Organizations and Individuals," Journal of the American Medical Association 324, no. 18 (2020): 1829-1830.

Engineer Your Own Luck

Companies that modularize and externalize their best capabilities are in a strong position to seize unexpected opportunities.

By Mark J. Greeven, Howard Yu, and Jialu Shan

PREDICTION IS HARD. THE LONGER the time range — six months, three years, a decade — the less reliable any forecasting becomes. And yet, some companies prosper through extreme uncertainties.

Consider cosmetics giant L'Oréal. When COVID-19 unexpectedly struck the world, consumers rapidly shifted to online platforms. Many companies were unprepared, but L'Oréal quickly adjusted, ramping up digital efforts to offer beauty products and services online. Use of its virtual try-on technology increased by a factor of five, and conversion rates from virtual try-ons tripled. This pivot not only supported consumers but also aided retailers such as Amazon and Sephora, which adopted the technology for their sales of L'Oreal products. In addition, the digital team added functionality that allowed beauty influencers to host livestreaming events on L'Oreal's own website.

L'Oréal's swift response can be traced back to its 2018 purchase of ModiFace, which develops advanced augmented reality and virtual reality technology for the beauty sector. During the pandemic, it worked with 27 key strategic partners representing 30 countries to make ModiFace available on their websites for applicable L'Oréal brands, facilitating integration via a software development kit. Thanks to its modularized capabilities, the company was able to move quickly to capitalize on unexpected trends in the midst of a highly uncertain environment.

Our ongoing research shows that increased unpredictability and rapidly changing conditions demand that executives strategize at a higher level. Rather than focus solely on products and markets, leaders must develop the ability to discover and act on a company's options — what is known as *optionality* — and to scale new initiatives quickly. That will position them to rapidly adjust to unforeseen market shifts, as L'Oreal did.

Neil Webb / theispot.com

We see this formula for success in the face of uncertainty playing out across industries — Mastercard in finance, Haier in home appliances, Nvidia and TSMC in chipset design and manufacturing, and Xiaomi in consumer electronics. All have pursued a particular path to growth. They have modularized internal capabilities into plug-and-play, mix-and-match services that partners and

clients can build on to grow new market segments. Even Nvidia's recent success in AI can be traced back to its 2006 launch of a computing platform and related application programming interfaces (APIs) that made it feasible to use Nvidia's powerful graphics processing units for general-purpose computing. That move opened new markets for the company's chip designs in a variety of high-performance computing applications.

Put another way, the top-performing companies, ranked by a composite score we call the *future readiness indicator*, grow by sharing their capabilities with partners that will develop new markets. (See "Defining Future-Ready.") Consequently, these companies can easily participate in multiple ecosystems, are resilient during difficult times, and are well positioned to capture unexpected upswings and opportunities. By treating capabilities as modular digital services, they can maximize their options for growth and scale quickly. That is how they engineer their own luck.

Steps to Architect Optionality

If attaining optionality and speed is important, why don't more companies do it? Because it requires grappling with four distinct challenges: (1) identifying core capabilities, (2) codifying relevant process knowledge, (3) breaking down complex, monolithic systems into loosely coupled modules, and (4) exposing interfaces to these modules so that others can harness them to develop new products and services. Overcoming these interrelated challenges requires leaders to be willing to make a long-term commitment and investment.

That commitment, in turn, requires courage, because the early stage of these initiatives are time consuming, can be resource intensive, and may not yield obvious payback aside from some internal productivity gains. But as the medieval theologian Thomas à Kempis observed, "The loftier the building, the deeper the foundation must be laid." In this article, we'll explain how to modularize your company's core capabilities and show how leading companies have succeeded in doing this work, creating both optionality and agility for themselves.

Identify Core Strategic Capabilities

It starts with being clear about your most important strategic capabilities. For a financial services company, this may include the capability to enable secure and smooth payment transactions for clients, which directly ties to its core mission. For consumer packaged goods manufacturers, like L'Oréal, it may include the capability to effectively engage consumers or manage suppliers. In our research, we look at three dimensions of identifying such strategic capability: relevance, frequency, and uniqueness.

Relevance is defined by how important a capability is for continuing to create value for customers. Frequency refers to how often the capability is deployed; after all, if it's deployed only once per year, there's little use in providing it to partners. Uniqueness means that the capability should reflect what the company does best — how it gives the organization a unique advantage in the marketplace compared with competitors. With the capabilities to externalize identified, the next step is codification.

Codify Process Knowledge

It may sound straightforward, but taking steps to break down and codify the process knowledge underpinning strategic capabilities is the crucial prerequisite for pulling apart a complex, often monolithic process into simpler components and facilitating modularization.

To begin, don't start with your most secret and critical process. Your goal is to make a handful of internal processes available to customers and partners to innovate faster. Focus on processes that are less sensitive. Maybe it's the supplier onboarding process, or maybe it's the process by which you seek out customers' input during product development. Prioritize so you can start to codify.

A German chemical company that one of us has worked with is known for its expertise in molecular design and delivery of specialty chemicals for coatings. Perfecting the formulas for the desired outcome is crucial, whether that's flame resistance, a smooth texture, scratch resistance, or water repellency. The company's clients include car manufacturers, furniture makers, smartphone companies, and paint producers, all of which have diverse needs and unique requirements that change rapidly based on consumer demands.

Given stringent safety regulations in its sector, the company was already aware that it needed to begin to codify relevant process knowledge that tended to reside only in the brains of individual experts. The planned launch of a digital tool that would allow customers to virtually experiment with chemical recipes provided the impetus to get serious about documenting the company's core capabilities. This required exposing critical process knowledge via a comprehensive, searchable, and user-friendly platform; and for that, the team needed to first codify that knowledge. They interviewed internal experts and pored through documents to detail each molecule, recipe, effect, and cost estimate. The resulting tool lets users see how various combinations of molecules can lead to new products and even estimates the cost of these new recipes. Capturing information on what users are investigating with the tool in turn provides insight to the R&D team about customers' emerging needs.

Defining Future-Ready

AT IMD, WE'VE COMPILED A FUTURE-READINESS INDICATOR to measure a company's preparedness, focusing on the top players across six industries: automotive, consumer packaged goods, fashion, finance, pharmaceutical, and technology. This comprehensive ranking, based on seven equally weighted factors and encompassing 36 variables, evaluates the financial fundamentals crucial for future investment, including a healthy cash flow, cash on hand, and debt. Additionally, it assesses growth prospects through investors' expectations and the intensity of investment in startups or new ventures.

Our analysis, which began in 2021, extends to the diversity within executive teams, highlighting the importance of gender, nationality, and industry diversity in promoting broadened perspectives beyond day-to-day operations. Productivity is gauged through metrics like operating revenue per employee, and the trajectory of new product rollouts is monitored to assess openness to new ideas and early innovation results. The balanced composite score derives from hard data sources, including financial reports, investors' calls, LinkedIn profiles of the management team, CrunchBase, Factiva, and other publicly available reports, ensuring a robust assessment of each company's future readiness.

Key to the success of this codification project was its narrow, well-defined scope. By focusing first on knowledge related to key clients' most frequent needs, the team turned the company's process knowledge from tacit and closely guarded into explicit and sharable. In contrast, we have seen large companies bogged down when they set the more ambitious agenda of codifying all process knowledge, as one CTO we know tried to do in support of a broad digital transformation.

Break Down Complex, Monolithic Systems Into Modular Capabilities

Once your company has identified the core capabilities that it can leverage with partners and customers and codified knowledge about related processes, the next step is to break those processes or systems into their component elements. Activating those specific, atomized capabilities involves calling on the workhorses of an agile digital architecture: APIs. These enable modularity by providing a standard way for software applications or components to communicate and exchange data or functionality. That modularity allows for fast changes, such as adding new functionality to a software suite without affecting other components.

Mastercard Labs, the global payment company's R&D unit based in Dublin, Ireland, realized how critical APIs could be via some hard lessons. It built a cloud-based payment gateway that made it easier for small businesses to take credit cards, called Simplify Commerce, but found that scaling the product was impossible because the company's core technology infrastructure was rigid and inflexible. Mastercard CEO Michael Miebach, who served as chief product officer from 2016 to 2020, observed that for a new product to thrive, the company would need to do some re-architecting of the core operation.

By a lucky coincidence, Mastercard Labs had also helped launch Mastercard Developers, which made APIs and other digital tools available to external developers so that they could integrate their innovations into the company's network. "The way we constructed products was API-first. It was a lot more transparent. It was a lot more modular," Miebach said. "So our ability to take something and plug it in somewhere else was much easier than before." As the chief product officer, he spent years rewiring Mastercard internally so that APIs became the standard interfaces across departments and processes.

What Mastercard went through was nearly a complete overhaul. Operations such as fraud scores, security, token translations, and identity verification were all historically bundled into "The Transaction." And while this core operation was tightly coupled, it also gave Mastercard its reputation for great stability, maintaining uptimes at 99.999% (which was not a metaphor but an actual metric, known throughout the company as "the five nines"). Yet it was the rigidity of these tightly coupled operations that prevented Mastercard from scaling new offerings like Simplify. In re-architecting the existing core, individual capabilities such as fraud detection, balance checking, and loyalty programs could be configured as mix-and-match components depending on new use cases. That now makes it easier for Mastercard managers to come up with modified offerings wherever and whenever they are needed. The result is true combinatory flexibility thanks to the modularized capabilities.

Mastercard's challenges in modularizing systems are not unique. Most companies' legacy systems, such as the classic version of SAP, are tightly coupled monoliths.

Breaking up a monolith can be long, slow work. It requires patience and commitment. Miebach recalls sitting in management committee meetings where, he said, "You feel like you're having to explain and defend why the APIs you are so bullish on don't always seem to be up to par." Nonetheless, with regular, meticulous assessments

and full transparency with leadership, both the quantity and quality of API-first offerings were inching upward. Mastercard eventually established what it calls the Single Front Door, a one-stop shop for all modular services. The benefits? Flexible scaling on demand.

With a modular setup, startups can mix and match Mastercard's existing capabilities to meet the global regulatory environment. This self-serving approach is critical for Mastercard to cater to the long tail of requirements from the fintech world. It's how Revolut, a fintech player based in the U.K., leveraged Mastercard's modular capabilities to create disposable virtual cards, which change their number with every online transaction for added fraud protection. And it is how Grab, the Uber equivalent in Singapore, launched its GrabPay Card, accepted anywhere Mastercard is, allowing Grab app users to make transactions both online and offline globally as they travel.

In November 2023, Mastercard received authorization to have its cards linked to China's biggest payment apps after years of being limited to cross-border services. The "pay like a local" feature was quick to implement. There was no more technology development at grand scale. Mastercard was ready whenever the opportunity materialized.

Open Your System to the External Ecosystem

As Mastercard's example illustrates, a modular approach can speed internal innovation. After establishing a modular internal architecture through APIs, the next step to increase optionality is to extend the open architecture beyond your company's boundary. This will enable your business to participate in multiple ecosystems at speed and scale.

China-based Xiaomi, a $50 billion consumer electronics company operating globally, has pursued a growth strategy built on partnerships. Its offerings span numerous categories, including smart phones, home appliances, and personal care items, with close to half of its revenues (44.9% in 2023) generated outside of China.

Xiaomi's advantage in getting into new markets lies in its digital exchange platform. This is built around MiUI, the user interface, or "skin," for its Android-based mobile phones, which are the center of the Xiaomi product ecosystem. Leveraging MiUI as a kind of universal remote control, the company began launching products that could be linked to and controlled by its phones, such as TVs, air conditioners, air purifiers, and smart lamps. But instead of scaling up R&D and operations to build new products, Xiaomi scouted for hundreds of partners that could build out this new range of offerings. What Xiaomi essentially did was open its internal architecture to third parties.

A small team led by one of the cofounders, Liu De, makes final decisions to sign new partners. But these decisions are informed by feedback from the active and engaged MiUI user community, which has been suggesting new features since the platform was merely a phone interface. By analyzing this large-scale feedback, Xiaomi can understand what products or features customers want. To facilitate product integration into the Xiaomi ecosystem, the company provides partners with a software development kit and an open API suite.

This strategy was informed by Xiaomi's deep belief in competing as a "bamboo forest" rather than as "the largest single tree." MiUI became the intricately connected underground root system that enabled new bamboo shoots — the new product categories — to sprout. Today, Xiaomi boasts an impressive connected ecosystem, orchestrating over 330 companies with 2,700 SKUs.

Similarly, L'Oréal's ModiFace has expanded into a full suite of API services for third-party deployment. It started with internal capabilities, morphed into a plugin for Amazon.com and Sephora, and was then made broadly available.

THRIVING IN TODAY'S AMBIGUITY DEMANDS companies move away from a rigid, manual-based structure to a flexible, modular setup. However, you must get your own house in order before you can fully leverage the benefits of opening your company's ecosystem. As a first step, identify a small number of customers or partners. Or, work with a small but avid fanbase so that the early crowd of adopters will buy into your vision of growing together.

Pioneers have showed us the way. These companies swapped out their organizational silos for standardized knowledge bases. They tore down their once-monolithic structures to create modular setups. They flung their doors wide open, sharing their tools with others. Using digital interfaces like open APIs, they slashed the cost of collaboration. These trailblazers offer a key path to staying forward: Treat capabilities as services. That is how to maximize optionality while scaling at extreme speed.

Being prepared for the unexpected will help you engineer your own luck. ■

Mark J. Greeven *is a professor of innovation and strategy at IMD Business School.* **Howard Yu** *is the Lego Professor of Management and Innovation at IMD Business School and leads IMD's Center for Future Readiness.* **Jialu Shan** *is an associate research director at the Center for Future Readiness at IMD Business School.*

Reprint 65434. For ordering information, see page 4.

Serve More Customers With Inclusive Product Design

Use these questions to empower teams to design products for more diverse populations.

By Vanessa M. Patrick and Jeffrey D. Shulman

Sjoerd van Leeuwen / theispot.com

IMAGINE A PRODUCT THAT MILLIONS OF PEOPLE find frustrating to use because of a design choice that could have easily been avoided. If it were your product, wouldn't you want to know who was frustrated by it and why, how to fix the problem, and how your organization could avoid making similar mistakes in the future?

Obviously, letting a customer down is bad for business. Yet some products do this every day, unintentionally.

Consider, for instance, how marketers and designers use color to create a distinct visual identity for their products and brands. Decisions about color are often based on the expectation that all users can see the colors that are chosen. But such an assumption might unintentionally exclude colorblind users, who account for 8% of the global population. Recognizing the importance of designing for accessibility, Adobe Color provides accessibility tools that enable designers to select hues directly from the color wheel at the start of the design process, helping them to avoid issues later.

Businesses have an opportunity — and a responsibility — to widen their lenses when considering whom their products are designed for. Inclusive product design aims to facilitate a match between the product and a diverse set of users, which can cultivate wins for the business, customers, and society in the process. Even small acts of attentiveness can yield significant results by making products usable by a larger population.

ADDRESSING User Diversity

Inclusive product design starts by understanding and considering the many dimensions of identity. Recent research, by Vanessa

Patrick and Candice Hollenbeck, published in the *Journal of Consumer Psychology*, has introduced psychologist Pamela Hays's ADDRESSING framework of sociocultural identity to the field of marketing. The factors on which the acronym is based — Age, Developmental and acquired Disabilities, Religion/spirituality, Ethnic/racial identity, Socioeconomic status, Sexual orientation, Indigenous heritage, National origin, and Gender — can help product teams discern underrepresented groups across the spectrum of identities.

Even companies that recognize the importance of inclusive product design might find themselves unsure about how to approach this work when faced with constraints and competing priorities. Using the ADDRESSING framework, we highlight four sets of questions that you can pose to expand the lens through which to view your product's target market, build an inclusive product development process, and develop inclusive innovations.

Questions to ask to include edge users in your target market. *How well are we as an organization meeting the whole spectrum of needs of our target market? Are we unintentionally overlooking user groups whose needs we can (profitably) meet? Do we know whom we are excluding?*

Apple found that speech recognition technology performed much worse when used by people who stutter and thus that population used such technology at a much lower rate than the general population. Using recorded speech data from users with speech impediments, Apple improved how its products worked for them.

Another example of meeting a previously unrecognized customer need involves transgender and nonbinary individuals, who often encounter excessive scrutiny and potential embarrassment when using credit or debit cards that don't reflect their identity. In 2019, Mastercard introduced the True Name feature, which allows consumers to display their chosen name on their card, even if they haven't legally changed their name.

Modern product development processes often focus on addressing the needs of the typical or average user, which can be too narrow a view. Companies can use surveys, customer reviews, and usage and performance data to identify differences in how well their products address the needs of specific populations.

Questions to ask to embed inclusive processes. *Do we include underrepresented and historically marginalized individuals throughout our product development and design process? Can we adapt our existing products, or do we need to build new products to support inclusion?*

When the needs of edge consumers can't be adequately met by existing products, companies can create new, inclusive offerings that center their needs. For

Inclusive product design is not just the responsibility of product designers.

example, in 2018, after having released several iterations of the Xbox controller, Microsoft's Xbox team developed a new model for gamers with limited mobility. It worked in close consultation with the disability community by testing prototypes with occupational therapists and listening to people in the gaming community who had accessibility needs. What resulted is a controller with a modular design that can be used with customizable accessories (joysticks, mounts, and buttons) and enables all users, regardless of physical ability, to experience the joy of gaming.

Centering on the most marginalized users in the product development process can create benefits for all customers, who often exceed the expectations of the remainder of the target market. Consider Nike's GO FlyEase hands-free shoes, which were inspired by a customer with cerebral palsy and designed in consultation with amputees to be easy to put on. The slip-on sneakers also benefit customers experiencing a temporary disability, such as a broken arm, and people whose hands are situationally unavailable, whether holding groceries, carrying a baby, or clutching a dog leash. Solving for mismatches created by a disability can also solve for other mismatches. For example, sidewalk curb cuts enhance the mobility of not only wheelchair users but also people pushing strollers, pulling grocery carts, or riding scooters.

To consider another kind of mismatch, products developed with artificial intelligence have suffered due to the underrepresentation of Native Americans in the training data. For instance, image recognition software has been shown to misidentify elements of Native American culture, and a popular AI image generator has been found to underrepresent Indigenous people. These examples demonstrate the potential downsides when historically marginalized communities are not included in data sets upon which algorithms are trained or in the evaluation of the algorithm's performance. In an effort to build algorithms with and for a broader set of users, Google's responsible AI practices involve assessing fairness in data sets and AI outcomes across different subgroups.

Bolting on accessibility or inclusion features to an

existing product late in its life cycle can lead to a suboptimal user experience, increased development and marketing costs, and missed opportunities to delight customers. A user who struggles with a product on their first try might never return to discover the accessibility features that were added in a later iteration.

Questions to ask to build a diverse product team. *Do we have hiring processes in place to ensure that we reach and recruit diverse talent that represents the market? Do we have a culture of belonging and psychological safety so that diverse perspectives are voiced and respected?*

Asking whether your team reflects the diversity of the market can help you embed diverse perspectives throughout the development process.

Companies can facilitate a more inclusive product development process when they hire employees across a range of identities. Sumaira (Sam) Latif, who is blind, has been with Procter & Gamble for more than two decades. As a special consultant for inclusive design, she worked with the Herbal Essence team to create a new packaging design that features tactile indentations to help vision-impaired individuals differentiate the shampoo bottle from the conditioner bottle. This not only benefits those with vision impairments but all who struggle with identifying the product they need to use in the shower.

Building inclusive products requires that organizations have not only windows into the needs of diverse users but also mirrors in which the teams themselves reflect the diversity of the market. Checking your "windows" and "mirrors" allows you to invent on behalf of a broader set of customers and cocreate meaningful innovations that impact their lives. For instance, Snap collects and reports not only gender, race, and ethnicity of its employees but also how many employees are neurodivergent, live with a physical disability, are members of the LGBTQ+ community, are first-generation college graduates, or have served in the military. Having a product design team with the lived experience of marginalization has resulted in Snap launching new product features like the Inclusive Camera, which better captures skin tones. That benefits everyone, particularly users with darker skin tones, who have historically been poorly served by camera-based apps.

Questions to ask to prioritize inclusion. *Which inclusion features should we prioritize? Are our estimates of the impact of inclusive products biased? Can we test hypotheses about the ROI of prioritizing inclusion?*

A product team will pursue development of features that broaden the customer base only if the return on such an investment is expected to be high. As a result, features that improve accessibility and inclusivity are often seen as nice-to-haves rather than must-haves during the development process.

This deprioritization of accessibility and inclusion might not be consistent with what is best for a company. Recent research, by Jeffrey Shulman and Zheyin (Jane) Gu, published in *Marketing Science*, has shown how research bias can lead companies to systematically under- or overinvest in developing inclusive products. Doubts about the impact of inclusion can also be self-reinforcing: Users from certain populations might not use a product because it was not originally designed for them, and therefore they might not respond to iterative improvements that are designed with them in mind. To interrupt this cycle, it is important to validate the potential impact of inclusive products by conducting a more comprehensive test.

Leaders can overcome resistance to prioritizing inclusive product design by applying standard process improvement methodologies, such as Six Sigma's DMAIC, to identify and demonstrate the impact of inclusion initiatives on key success metrics. Well-defined pilot projects can help a team analyze how intentionally engaging diverse communities of users will affect product development, including time and financial costs and product performance.

TO PROVIDE EXPERTISE AND GUIDANCE FOR product teams seeking to build more inclusive products, several companies are creating specific roles dedicated to product accessibility or inclusion. For instance, ADP has hired a chief of product inclusion and Microsoft a chief inclusion officer. Product equity teams are in place at Airbnb, Google, LinkedIn, Meta, and Snap.

As a leader, you have an opportunity to utilize the four sets of questions outlined above to uncover lucrative business opportunities that can have a positive impact on people and society. Inclusive product design is not just the responsibility of product designers, user experience researchers, engineers, or product managers. Leaders can nurture change by creating organizational goals around product inclusion and encouraging each team to establish success metrics and initiatives that support these goals. ■

Vanessa M. Patrick *is the Bauer Professor of Marketing at the University of Houston's Bauer College of Business. She has published more than 40 peer-reviewed articles and is the author of the book* The Power of Saying No: The New Science of How to Say No That Puts You in Charge of Your Life (*Sourcebooks, 2023*). **Jeffrey D. Shulman** *is a podcaster, filmmaker, and catalyst for change who helps organizations achieve the outcomes that matter to them. He is the Marion B. Ingersoll Professor of Marketing at the University of Washington Foster School of Business.*

The CEO's Cyber Resilience Playbook

What do CEOs who led through a serious cyberattack regret? Use this guide to learn from their experiences and take smarter actions before, during, and after an attack.

By Manuel Hepfer, Rashmy Chatterjee, and Michael Smets

ON MAY 7, 2021, EXECUTIVES AT Colonial Pipeline discovered that cybercriminals had launched a ransomware attack on its IT systems. To prevent the malware from spreading further, the company took its computer systems offline, disabling 5,500 miles of pipeline that supplied 45% of the fuel consumed on the U.S. East Coast. The disruption lasted nearly a week, resulting in panic buying and fuel shortages. In a controversial decision, Colonial Pipeline paid a ransom of nearly $4.4 million in exchange for the decryption keys to get its systems back online. One month later, with recovery efforts and investigations ongoing, Colonial Pipeline CEO Joseph Blount defended that decision before the U.S. Senate, testifying, "We were in a harrowing situation and had to make difficult choices that no company ever wants to face."

Blount's testimony echoes the experiences of many of the CEOs we have interviewed as part of our research into how leaders manage cybersecurity risk and attacks.[1] These CEOs shared with us similarly painful accounts of having to make existential decisions based on imperfect information, under enormous pressure, in an area where they had relatively little expertise. Serious cyberattacks thrust CEOs into the public eye, scrutinized by the media, shareholders, regulators, and other stakeholders.

We conducted 37 in-depth interviews with the chief executives of large enterprises (with average revenues of $12 billion) in the United States, Europe, and Asia. Nine of them had led their company through a serious cyberattack, which allowed us to compare their battle-tested views with those of CEOs who had not yet suffered such an attack. This article outlines strategies, based on their lessons, to help your organization stop over-relying on cybersecurity and start building cyber resilience as a strategic opportunity.

What CEOs Regret After a Serious Cyberattack

The CEOs who had lived through cyberattacks on their organizations spoke candidly (and anonymously) about their experiences, evaluating their preparation strategies and the actions they had taken. They also shared their regrets based on lessons learned from their experiences.

They focused too narrowly on prevention. It would be a struggle to find a company that does not currently list cyber risk near the top of its enterprise risk register. Cybersecurity has become an inescapable priority for chief executives. But too often, leaders believe that

it is possible to protect the confidentiality, integrity, and availability of information systems and data by preventing attacks. This has possibly never been true — and it most certainly isn't today. As cybercrime gains sophistication and is increasingly weaponized by nation states, even the most technologically advanced organizations will be breached — and they need to plan for that inevitability.

By far the most common regret we heard from CEOs was that they overemphasized cybersecurity to the neglect of cyber resilience. Only after the attack did they come to understand that trying to prevent cyberattacks is a losing game. These executives had focused on *whether* they would get attacked instead of on *when* they would get attacked and how they would respond when it happened. Although the CEOs poured significant resources into technological defenses, their companies often lacked basic organizational cyber resilience practices.

Cyber resilience describes an organization's ability to anticipate, withstand, respond to, and adapt to cyberattacks. The goal is not to simply avoid an attack but rather to minimize its impact, recover quickly, and emerge stronger. By shifting an organization's focus away from prevention and onto preparation, leaders prioritize developing adaptive capabilities, which should happen across all departments.

They felt merely accountable for cybersecurity. All of the CEOs we interviewed insisted that they are accountable for everything in their business, including cybersecurity — and yet 72% declared that they were not comfortable making decisions in that area. Notably, CEOs who had been through an attack regretted feeling *merely* accountable — that is, taking ownership after something bad had happened. Being responsible, in contrast, involves ongoing, proactive engagement before things go wrong.

As one CEO put it, they needed to be "co-responsible" for cyber resilience alongside their chief information security officer (CISO). That's a development that many of the 40 CISOs we also spoke with as part of our research would welcome, given that they saw their CEOs as insufficiently accountable, let alone responsible, for cyber risk. One CEO in our study reported a tangible commitment to this co-responsibility: He spent 10 full days with his cybersecurity team after the company had recovered from a devastating attack, to show that he would be responsible alongside them going forward.

They blindly trusted cybersecurity and technology teams. CEOs trust their teams all the time. Cybersecurity and technology teams should be no different — but they are. Most CEOs have some degree of expertise in functions such as marketing or finance that helps them evaluate and challenge advice from others. In contrast, very few chief executives have a background in IT, let alone cybersecurity. After a cyberattack, many of the CEOs found themselves leaving the fate of the company in the hands of experts whose counsel they couldn't fully comprehend.

As a result, CEOs who had been through an attack sought to develop more informed trust in their cybersecurity teams. Executives must understand enough to be able to ask the right questions, challenge advice, and make decisions. This requires both curiosity and the humility to continuously learn as cybersecurity and the threat landscape constantly evolve. The more CEOs understand about cyber resilience, the more they become concerned; the more concerned CEOs are, the more engaged they become.

They felt well prepared for a cyberattack. With the notable exception of those who had endured a cyberattack, the majority of CEOs in our study considered their organizations to be well prepared for such an event. That mindset can breed complacency that undermines efforts to continually shore up defenses and resilience planning. The CEOs with cyberattack experience admitted that only when attackers struck did they realize how ill-prepared they were and that the situation was far more dire than they had envisioned.

Crisis plans and playbooks can fuel the illusion of preparedness. While they are necessary, they aren't sufficient preparation for a threat that is likely to unfold unpredictably. Because attackers continually adapt and develop new tactics, organizations are likely to face threats that were previously unknown: Tomorrow's cyberattack is unlikely to look anything like yesterday's. And while playbooks lay out procedures for managing the crisis, an attack may compromise the very infrastructure needed to do that, such as key communication channels.

CEOs who have been through an attack do not believe that organizations can be truly prepared. They adopt a mindset of constant underpreparedness to encourage their teams to continually test and evolve the organization's ability to respond to an attack. In other words, they never feel prepared but are always ready for a cyberattack.

They reacted rather than reassured stakeholders. During a cyberattack, CEOs face tremendous amounts of pressure from all sides. Shareholders worry about the financial impact, the board wants evidence of business recovery, regulators want answers, customers worry about their data, and business partners want to know whether their systems are at risk of contagion. Faced with this barrage of competing demands for rapid responses, CEOs may default to a reactive mode in which they largely transmit information without first carefully

Michael Glenwood Gibbs / theispot.com

evaluating it, such as communicating the IT team's unrealistic recovery schedule to external stakeholders. In retrospect, many CEOs regretted not engaging various stakeholders more proactively — reassuring them rather than merely reacting.

CEOs can express reassurance in three key ways. As an *amplifier*, the CEO reinforces external pressures to create a sense of urgency. This is especially useful for reassuring external stakeholders that the organization is not succumbing to complacency throughout day-to-day operations and when new risks emerge. As a *filter*, the CEO judges what kinds of pressures to absorb or transmit. This is especially helpful during and after a cyberattack, when external pressures for accountability, and even blame, may distract people at the heart of the crisis response from recovery efforts. Finally, as an *absorber*, the CEO does not pass on any pressure but absorbs it and focuses on reassuring all stakeholders about the company's resilience.

How CEOs Can Build Cyber Resilience

Drawing on our research, we developed a playbook of best practices for CEOs to help their organizations build greater cyber resilience and gain confidence in their own ability to manage all stakeholders when cyberattacks inevitably occur.

Invite an outsider CEO with cyberattack experience to speak to the CEO, leadership team, and board of directors. Hearing firsthand accounts of attacks can serve as an important wakeup call to the complacent. Executives who have managed a serious attack can share powerful personal stories and valuable experiences that are relevant to any company. This amplifies external pressure for greater cyber resilience and reassures stakeholders that the issue is being taken seriously. Hearing from executives rather than technology experts provides a more relatable perspective on how executives can anticipate cyberattacks and prepare to weather them. And, of course, CEOs themselves can gain direct insight into how their peers have contributed to limiting the damage and developed a more informed perspective.

Set up a cyber resilience learning forum. Some CEOs in our study regularly bring together the board, management team, business unit leaders, and security and IT teams for an open exchange about the most pressing cybersecurity challenges and current business priorities. The objective of these sessions is to build shared understanding — a more integrated, end-to-end perspective on business resilience that business leaders and cyber experts can draw on for both cyber planning and attack response. The CEO should chair the forum and establish the norm that the most naive questions are welcome: They help move individuals toward more-informed trust.

Cyber resilience is not just about avoiding loss — it can also lead to value creation.

Consider holding a forum at least annually, though quarterly may be most appropriate for this fast-moving topic. The forum sits outside formal risk governance processes and emphasizes learning. It does not replace cyber risk governance processes and committees.

Commission an independent cyber audit. Consider the value gained from financial audits. Similarly, CEOs should commission independent cyber resilience audits once per year. These specialists will report any findings directly to the chief executive and offer advice on addressing any issues they uncover. The purpose is not to get certified to some industry standard but for the CEO to gain a better understanding of the status quo, build more-informed trust with the technology team, and discover blind spots in the organization. To achieve all this, CEOs should work closely with their cybersecurity lead to analyze the results, which might range from technical to organizational and strategic depending on the focus of the audit. Unfortunately, we often find that CISOs hesitate to encourage the CEO to commission cyber audits because the CISO fears being put on the spot about the findings. But in reality, an audit often builds trust between the CEO and the cybersecurity team and forges co-responsibility for cyber resilience efforts.

Identify critical processes and priorities in case of attack. In the absence of a comprehensive response plan, when all systems are disabled, every department will insist that its business processes are critical and must be restored first. To avoid scrambling and infighting — and mount a coordinated and swift response — CEOs should sit down with their board and management team and identify the two to five business processes that are most critical to keeping the organization running. The components of each process should be mapped to the applications and servers that support them so that teams not only understand priorities for action but can make each critical process more resilient. A manufacturing company's critical processes might be taking customer orders and keeping certain production lines running; for a law firm, they might be email and time-tracking.

Agreeing on such priorities aligns the organization and informs decentralized decision-making when systems are down. However, even when organizations have carefully crafted plans and agreed-upon business

priorities, the truism that "no plan survives contact with the enemy" still applies. Leaders must be ready to make bold decisions on priorities in the moment. But CEOs should not become the single point of decision-making. The CEOs we interviewed emphasized the importance of a shared set of guiding principles for the organization. This could be as simple as communicating that employee safety comes first, clients second, and shareholders third.

Seek expert assistance in advance of a crisis. Despite well-laid plans, organizations rarely have all of the up-to-date capabilities needed to deal with a serious cyberattack entirely on their own. Leaders should maintain a panel of trusted advisers to call on in the event of an attack. These might include law firms to help comply with regulations and interact with government agencies, communication specialists to protect the company's reputation, forensic investigators to discover how the attack happened, and even ransom negotiators to stall for time or, in the worst case, reduce a ransom demand.

Prepare to communicate proactively. One of the most important early decisions CEOs face during a cyberattack is what and how to communicate with employees, stakeholders, and the market. Keeping it all under wraps is not an option: It is nearly impossible to control the narrative without an open and transparent media strategy. Markets and regulators are far less forgiving of cyberattacks if the company appears to have been caught flat-footed or, worse, appears to be hiding something. CEOs should have a crisis communications plan in place so that they are ready to proactively reassure stakeholders about how the company will minimize impact and disruption.

Conduct a postmortem if the worst happens. When an organization has just survived a cyberattack, its leaders must capture insights quickly. While many in the organization are likely to be exhausted, CEOs should initiate an incident review with stakeholders from multiple business units. This discussion should focus on more than just the technical aspects of the attack and what can be done to avoid similar ones in the future. Participants must also consider elements of organizational resilience that might be improved, such as communication, crisis management capabilities, and business continuity processes. What's more, postmortem workshops can expose outdated business processes and previously unnoticed inefficiencies that impeded resilience during the cyberattack.

Look out for opportunities. Cyber resilience is not just about avoiding loss — it can also lead to value creation.[2] One CEO we spoke with told us that the attack exposed significant inefficiencies in the organization's technology setup. Efforts to shore up cyber resilience by consolidating systems produced efficiency gains. The CEO allocated the resulting savings to building cyber resilience — a win-win.

But opportunities also exist in wider organizational structures, beyond optimizing the technology stack. A cyberattack can create a new appreciation for the criticality of otherwise hidden business processes, greater awareness of mutual dependencies, and greater recognition of leaders who stepped up at a time of crisis. This is the moment to pull the entire organization more closely together, redesign organizational structures, and reshuffle leadership teams for enhanced resilience.

These considerations should not stop at organizational boundaries. A major cyberattack often ripples across industries and pulls together otherwise competitive rivals. Capture opportunities to take more of an ecosystem approach to developing cyber resilience together with competitors, suppliers, and other stakeholders.

But most importantly, know that these opportunities will not just happen on their own. CEOs must actively look for them, treating the post-attack period as an opportunity to strengthen the very foundation of their company for the digital era.

THE EXPERIENCE OF A SERIOUS CYBERATTACK is one that many CEOs would like to forget. While they typically find this difficult personally, they note that organizationally, memories tend to fade once normalcy returns. Therefore, while it is a painful trauma to revisit, it is important that they keep the memory alive.

To build cyber resilience beyond the boundaries of their organizations, CEOs should also share their experience with other leaders. While top executives naturally hold their cards close as they seek to outcompete and outsmart their rivals, cyber resilience thrives on mutual support and sharing information. Cyber should be seen as a noncompetitive domain where companies — even those in the same industry — work together to achieve greater levels of resilience. By sharing their stories and what they have learned with other industry peers, executives contribute to the resilience of the entire ecosystem. ■

Manuel Hepfer *is the head of knowledge and insights at cybersecurity firm Istari and a research affiliate at Oxford University's Saïd Business School.* **Rashmy Chatterjee** *is the CEO of Istari.* **Michael Smets** *is a professor of management at Saïd Business School.*

REFERENCES

1. M. Hepfer, R. Chatterjee, and M. Smets, "The CEO Report on Cyber Resilience," PDF file (London: Istari, 2023), https://istari-global.com.
2. M. Hepfer and T.C. Powell, "Make Cybersecurity a Strategic Asset," MIT Sloan Management Review 62, no. 1 (fall 2020): 40-45.

Return-to-Office Mandates: How to Lose Your Best Performers

Your organization's highest-performing employees want executives to focus on outcomes and accountability, not office badge swipes.

By Brian Elliott

RECENT RETURN-TO-OFFICE (RTO) mandates like those at UPS and Boeing have a simple message: Come back to the office five days a week. CEOs cite productivity as a core reason for these proclamations, even in the face of employee resistance. Many executives simply don't trust that employees are as effective as possible when managers can't see them at their desks.

But in a world of globally distributed teams, falling back on management-through-monitoring is falling back on the weakest form of management — and one that drives down employee engagement. There is mounting evidence that mandates don't improve financial performance. Instead, they damage employee engagement and increase attrition, especially among high-performing employees and particularly those with caregiving responsibilities.

There is a better way forward, but it requires culture work at the top as well as deep within organizations, along with a significant upgrade in management philosophy. Too many organizational cultures use face time at the office as their metric for productivity. That's not the best benchmark. Instead, focusing on outcomes while providing trust and flexibility about where and when to get work done allows individuals and organizations to thrive.

What's Behind the RTO Drive

We're four years past the start of the pandemic-driven shift toward flexible work. While the peak of remote work has passed for now, office utilization in the U.S. has been effectively flat. For the past year, it's been hovering at 50% of pre-pandemic norms.

RTO pronouncements have come loud and clear from Amazon, Google, IBM, JPMorgan Chase, and more. Some CEOs talk about the solidarity that office workers need to have with front-line workers, although Gallup research shows that a strong majority of front-line workers aren't bothered by office worker flexibility. In fact, they want flexibility themselves, including a choice of which days to work, the option for four 10-hour days, and some flextime — in other words, equity, not equality. CEO concerns also seem a bit selective. As one chief human resources officer put it to me, "We don't seem to be offering our factory workers access to the corporate jet."

Wall Street pressures on CEOs are often closely tied to RTO pronouncements. In 2022 and 2023, you could draw correlations between pressure from activist investors and subsequent announcements of RTO mandates, particularly in the technology sector. While activist investors targeting companies such as Amazon, Disney,

Carolyn Geason-Beissel / MIT SMR / Getty Images

Google, and Salesforce all pointed to the need to drive higher returns, the timing of their pressure also coincided with the appearance of RTO mandates shortly thereafter. There is a strong feeling of wanting things to go back to the way they worked in the old days.

But there's no clear evidence that these mandates improve financial performance. A recent study of S&P 500 companies that was conducted by University of Pittsburgh researchers found that executives are "using RTO mandates to reassert control over employees and blame employees as a scapegoat for bad firm performance." Those policies result in "significant declines in employees' job satisfaction but no significant changes in financial performance or firm values," they concluded.

I still get asked, almost weekly, "What about those headlines last summer saying that studies showed that remote work was less productive?" Much of the data was acquired during the throes of the pandemic, often under horrible circumstances and in the absence of support in terms of management oversight. Productivity among the IT workers in India who were thrown into remote work in 2020 and whose performance is cited in those studies was, indeed, demonstrably worse, thanks to the horrific impact of the pandemic in that country.

Perhaps most instructive is evidence that most executives don't think the mandates they've already imposed have helped. Among executives who have instituted a return-to-office mandate, only 1 in 3 thinks it had "even a slight positive impact on productivity." That's not surprising to anyone who's been inside an organization during these drives. They create internal churn and distract employees and their managers from focusing on what matters more: their customers.

The lingering uncertainty I've heard about in many conversations with senior executives comes down to two simple questions: Are people really working when they're at home? And would they work more and get more done if they were in the office? Some of this skepticism is fed through a CEO echo chamber and by board conversations, where anecdotes about bad experiences can easily start to feel like trends, and where dissent by CEOs sympathetic to hybrid work paints them as outsiders. In the words of one chief people officer, "We dread every time CEOs get together — the peer pressure is a real challenge."

Performance Monitoring Leads to the Doom Loop

The "Are they really working?" mystery would be easily overcome if organizations had good ways of measuring productivity. But over the past 50 years, efforts to evaluate the productivity of knowledge workers have largely been futile; as Peter Drucker pointed out in 1969, once we automate rote tasks, what's left is creativity, decision-making, and complex problem-solving. Putting a yardstick on knowledge work is hard — and it's usually unsatisfying for both the employee and the organization attempting to make the evaluation.

But because productivity remains at the top of the stack of executive concerns, according to recent research by Atlassian, The Conference Board, and Slack, executives are attempting to measure it through activity. According to the Slack survey, 70% of executives use visible activity (such as what time people show up and how many hours they log) as a primary measure of productivity. Notably, the same survey found that, on average, employees spend 32% of their time "on performative work that gives the appearance of productivity."

Far too often, this combination of executive anxiety, monitoring, a lack of trust, and having employees perform to the metric of being in the office all contribute to lower performance, which I call a *doom loop*. Employees who know they are being monitored react in a variety of ways. To start, 49% of them report high levels of anxiety, compared with 7% of people not being monitored, according to a study from the Centre for Transformative Work Design. Not a recipe for getting the best out of employees.

More to the point, monitoring does not work. It's easy enough for employees to perform to the system by scheduling when their emails go out, keeping their status active online even when they're not working, or buying mouse jigglers. While this might seem like cheating the system, employees are reacting to the core issue, which is a lack of trust.

At the individual level, a lack of confidence that employees will do the right thing is corrosive to employee engagement. In a Slack survey, people who said they felt trusted by their employer reported twice the productivity of those who didn't feel trusted, and they were 30% more likely to say they put in extra effort at work. Companies that don't exhibit that trust also diminish passion.

The Real Impact of Return-to-Office Mandates

RTO mandates do have an impact — and it's pretty much all negative. The same study that found that RTO mandates have had no impact on the financial performance of the S&P 500 did find one enduring effect: significant declines in employees' job satisfaction.

Forcing people who've been effective working from home three or four days a week to commute to the office instead is a recipe for grief. Some 75% of people who aren't happy with their level of flexibility said they will

look for new jobs, according to Future Forum research.

Who's most likely to leave? Women, caregivers, and other historically underrepresented groups at work, to start. The same Future Forum research found that 59% of working mothers want to work from home three or more days a week — as do 47% of working fathers. Flexible work is feminist, as the headline on a *Fortune* article by Erin Grau so eloquently put it. We've also seen for four years that Black, Latine, and Asian American office workers value flexibility more than their White colleagues do — with the ability to dial in and out offering a respite from code-switching. Forcing the march to the office undermines the ability to build inclusive, not just diverse, organizations.

If none of that captures your attention as a leader, how about this: A Gartner survey of more than 2,000 office workers found that the group *most* likely to leave are high performers, where "intent to stay" in the face of a mandated RTO dropped 16% — compared with 8% for the average employee. For top talent, the signal "We don't trust you" stings deeply. These employees are already going above and beyond, but capricious policies tell them that appearances matter more than results.

Building the Boom Loop: The Outcomes-Driven Organization

The impact is clear: Mandates and monitoring lead to not only decreased employee satisfaction, retention issues, and a loss of trust but also to lower performance on what matters most to the C-suite — business results.

Conversely, companies that build trust outperform their peers financially. Institute for Corporate Productivity research showed that employees at high-performance companies (organizations that outperformed industry peers on factors like revenue growth, profitability, and customer satisfaction) were *11 times more likely* to say they had senior leaders and managers who trusted them to do their jobs.

Trust and performance come from building alignment on how to measure results and drive accountability. When companies offer trust, flexibility, and a focus on outcomes, they see greater engagement and productivity, which is what I call the *boom loop*.

Organizations that have figured out how to do this — or that are hoping to — know that clarity around the most important outcomes the organization is trying to drive is the starting point. Getting there requires hard discussions and decision-making at the executive level, and clear communication up and down the organization.

There are many facets to getting to this kind of clarity. The long-term purpose of the organization needs to lead to near-term objectives and short-term ways of measuring progress. Those objectives and outcomes need to be prioritized — what's most important for the organization to accomplish — with a well-defined understanding of who is accountable for performance.

Done right, exercises to get clarity around outcomes and priorities result in companies saying no (or at least "not now") to a wide variety of activities. This will feel painful to at least part of the organization. (Google, which had a well-known objectives and key results process in its early days, had a set of OKRs when I was there that was so broad that nothing was off the table.)

Anyone who's worked in an organization of more than a few hundred people knows that what I've just described can be a massive shift that requires effort, leadership commitment, and hard decisions. But once you're there, magic things can happen.

Teams can be given the flexibility to find the right path to the outcomes they need to achieve and ways to ask for help when they need it. Flexibility at the team level leads to individual flexibility — and flexibility leads to trust.

A company driven by measuring outcomes instead of monitoring employees will allow for flexibility about where and even when people work. Most people want some time together with their teams and some time at home. Teams that figure out a regular rhythm that works for them perform best. Many executives are aware of this, which explains why, in a Conference Board survey, 27% of U.S. CEOs, 30% of Europe CEOs, and 13% of Latin America CEOs said that "maintain hybrid work" is a human capital priority in 2024, whereas only 4% to 6% of CEOs across those regions wanted to drive a full-time, five-day-a-week return to the office.

The bottom line is that when trust is balanced with accountability, people and organizations will thrive. The time employees save by not having to commute can be put toward their jobs or their personal lives. Judging employees on the outcomes they drive, not whether they're showing up to an office after a grueling hour in traffic or on public transportation, reinforces two key elements: first, that the company takes accountability for performance seriously, and second, that people will be rewarded for their performance. Instead of focusing on internal debates about policies, mandates, and monitoring, employees can focus on delivering outcomes.

That's the boom loop. It's time to go build it. ■

Brian Elliott *is an executive adviser and speaker. He is coauthor of* How the Future Works: Leading Flexible Teams to Do the Best Work of Their Lives (*Wiley, 2022*).

A Tale of Two Hot Sauces: Spicing Up Diversification

The contrasting paths of two hot sauce manufacturers show that managing exposure on multiple fronts is essential.

By Achal Bassamboo and James G. Conley

IT WAS THE BEST OF TIMES, IT WAS THE worst of times. The divergent fates of two rival condiment businesses — the best of times for McIlhenny Company, maker of well-known Tabasco-brand hot sauces, and the worst of times for Huy Fong Foods, originator of the U.S. version of the popular Sriracha pepper sauce — highlight the power and perils of diversification. While McIlhenny took strategic steps to diversify with future performance and risk in mind, Huy Fong failed to capitalize on its wild initial success, making missteps that ultimately rendered the former market leader an also-ran in hot sauce, even in the subcategory Huy Fong itself created.

The cases of McIlhenny and Huy Fong illustrate the right and wrong ways to diversify.

150 Years of Hot Sauce

Hot sauces, used to spice up a wide variety of main dishes, sides, and snacks, are simple products containing a limited set of ingredients such as pepper mash, vinegar, salt, garlic, and sugar. With the exception of the peppers, these are typically commodity inputs. Thus profit in this industry is realized by selling a unique product at scale with strict control over input costs.

McIlhenny, now in its fifth generation of family ownership, has been producing food products since the late 1800s, starting with its best-known "original" line of hot sauce. The business carefully manages the particular kind of peppers its contract farmers must grow and source from multiple suppliers on multiple continents, as we'll discuss below. In 2014, McIlhenny started producing

Carolyn Geason-Beissel / MIT SMR / Getty Images

its own brand of Sriracha sauce based on the growing demand for these products.

Huy Fong, a relatively new entrant to the category, was founded by Vietnamese immigrant David Tran in Los Angeles in 1980 and rose to prominence with its Sriracha sauce, a complex combination of spicy, garlicky, and sweet flavor notes — a condiment Tran introduced to the U.S. Also called "rooster sauce," based on the Huy Fong logo (a symbol of Tran's zodiac sign), sales grew rapidly as the product's hot-sweet taste resonated with contemporary palates. By 2020, sales were over $130 million, and the business was valued at $1 billion. Huy Fong helped meet rising demand with a long-term arrangement with Underwood Ranches, its single, local source of the jalapeño peppers essential to the product — one of multiple critical errors the company would make.

A Story of Diversification

Given Huy Fong's rapid rise on the back of its Sriracha sauce, it may be surprising to hear that the product disappeared off shelves in the early 2020s (with ongoing uncertainty about its availability), essentially ceding the Sriracha market to rivals including McIlhenny, the current leader. The simple explanation is diversification strategy — or the smart moves McIlhenny and others made in the face of Huy Fong's misguided efforts.

These key dimensions of diversification determined these businesses' highly different fates.

Supply Chain. In 2016, Huy Fong sought to diversify its pepper sources, largely due to a disagreement with and lawsuit against Underwood, its sole supplier until then, regarding overpayment for peppers. With Underwood no longer a partner, the business contracted with pepper suppliers in California, New Mexico, and Mexico. But over the period from 2020 to 2022, the new suppliers faced climate challenges that severely limited their ability to produce.

These supply chain issues compromised Huy Fong's production capacity, as outlined in the company's communications with key customers. Meanwhile, multiple parties with better access to peppers, including Underwood, entered the hot sauce market with a Sriracha product. Not only was the formulation unprotected, but Huy Fong's branding focus on the name Sriracha, a geographic term, limited the trademark protection, inadvertently creating opportunities for others in the hot-sauce supply chain.

When Huy Fong ceased production of its Sriracha sauce, it began to lose all associated market share, yet McIlhenny's multi-sourcing strategy positioned it well to fill the market shortage. The business strategically sourced peppers and other inputs from multiple continents to avoid the very challenges that Huy Fong had overlooked — such as the vagaries of local climates. By October 2023, Tabasco-branded Sriracha was the market leader.

Products, Customers, and Channel. Huy Fong put all its proverbial eggs in one basket by relying almost solely on its flagship Sriracha sauce for revenue, running the risk of losing business if and when customer tastes shifted.

McIlhenny, meanwhile, steadily built a portfolio of hot sauce products, extending the brand to nine sauces including the original. "We need to branch out with the brand," said Harold Osborn, CEO and fifth-generation family member. That conferred key advantages to the business: The company could cater to multiple segments of customers with different taste preferences, such as with jalapeño- and chipotle-pepper-based sauces. Unlike Huy Fong, McIlhenny was less susceptible to ingredient shortages and could ramp up production of alternative products in the face of supply chain issues. Moreover, if customers couldn't find a given Tabasco product on the shelf, they could try a different one, potentially expanding growth and profits.

Importantly, McIlhenny diversified to selling wholesale products as inputs as well, driving more value while reducing risk. McIlhenny made the smart decision to diversify from B2C to B2B products, selling peppers and other inputs to other food makers. Huy Fong, too, provided Sriracha to fast-food and other restaurants, but it wasn't sufficient to offset its supply chain challenges and limited product selection.

Diversification on key dimensions drives value while reducing risk. Supply chain, products, customers, and channel are elements of diversification that led to McIllhenny's rise alongside Huy Fong's decline. Diversification is critical, but it has to be done right. Huy Fong limited its own success by primarily focusing on one product and demonstrating a "too little, too late" approach by waiting too long to diversify its supply chain. Lack of diversification on these fronts failed to protect Huy Fung against the possibility of a supplier becoming a competitor.

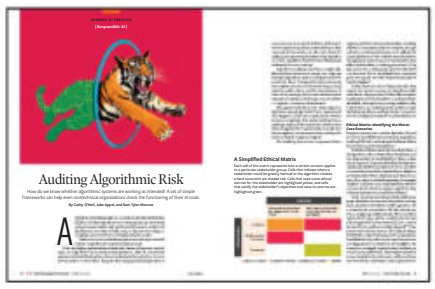
The bottom line: If you're aiming for long-term success, diversification is a must. We hope the ideas here help you strategically diversify your way to the best of times for your business. ■

Achal Bassamboo *is the Charles E. Morrison Professor of Decision Sciences at Northwestern University's Kellogg School of Management.* **James G. Conley** *is a clinical professor of operations at Kellogg.* *The authors acknowledge the generous support of the Center for Research in Technology & Innovation at Kellogg.*

Reprint 65409. For ordering information, see page 4.

Executive Briefings

SUMMER 2024 | VOLUME 65, NO. 4



Auditing Algorithmic Risk

Cathy O'Neil, Jake Appel, and Sam Tyner-Monroe pp. 30-37

Key Insight: A set of frameworks can help organizations identify potential harms posed by algorithms, AI tools, or large language models (LLMs).

Top Takeaways: As business uses of artificial intelligence, LLMs, and other algorithmic applications expand, so, too, does the possibility of unintended negative consequences for users and other stakeholders. The authors introduce two auditing frameworks — Ethical Matrix and Explainable Fairness — that can help organizations identify these potential risks and address critical questions about who could be harmed by algorithmic systems and how. They also discuss applying red teaming and benchmarking to difficult-to-audit LLMs, before applying all four of the approaches to a real-life example to demonstrate how an algorithmic audit could have prevented a very public chatbot failure.

REPRINT 65435

Avoid ML Failures by Asking the Right Questions

Dusan Popovic, Shreyas Lakhtakia, Will Landecker, and Melissa Valentine pp. 38-41

Key Insight: Checking assumptions and mapping out work processes can help ensure that machine learning solutions fit the job to be done.

Top Takeaways: A significant reason why machine learning projects fail to deliver business value is data scientists' failure to adequately understand the business context. Development teams can avoid mistakes when they put aside any reticence to ask basic questions and engage with colleagues on the business side. The authors advise gaining input from all involved stakeholders and suggest some specific types of queries that might help machine learning developers get to the heart of the problem at hand.

REPRINT 65402

How Generative AI Can Support Advanced Analytics Practice

Pedro Amorim and João Alves pp. 42-46

Key Insight: The natural language capabilities of large language models (LLMs) can augment the predictive powers of advanced analytics when well-designed prompts are applied.

Top Takeaways: Advanced analytics, such as predictive and prescriptive models to support business decisions, remain the primary drivers of data science value in the enterprise. How might the flashy, fluent, but not entirely reliable generative AI LLMs contribute to traditional analytics practice? The authors describe some experimental prompts that show potential for labeling data and explaining model predictions, and share guidance on monitoring and verifying that output.

REPRINT 65436

Managing Data Privacy Risk in Advanced Analytics

Gregory Vial, Julien Crowe, and Patrick Mesana pp. 47-51

Key Insight: Gaining value from data assets that include customers' personal information requires the use of privacy techniques that balance data utility with data protection.

Top Takeaways: Many companies have large stores of customer data that can be tapped for valuable insights via analytics. At the same time, the cybersecurity tactics for protecting personal information within that data can render it less useful for analysis. Data science practices will increasingly require that teams collaborate with IT on each use case to identify which techniques will maximize data privacy while still exposing useful information in the data set for analysis. Organizations can achieve better balance between data utility and data security by including data privacy in data literacy programs, treating data privacy as a business issue, and formalizing their approach to addressing the issue.

REPRINT 65433


[Customer Relationships]

Acing Value-Based Sales

To get the best returns on innovative products, collaborate with customers to define and share the commercial opportunity.

By Marco Bertini, Oded Koenigsberg, and Todd Snelgrove

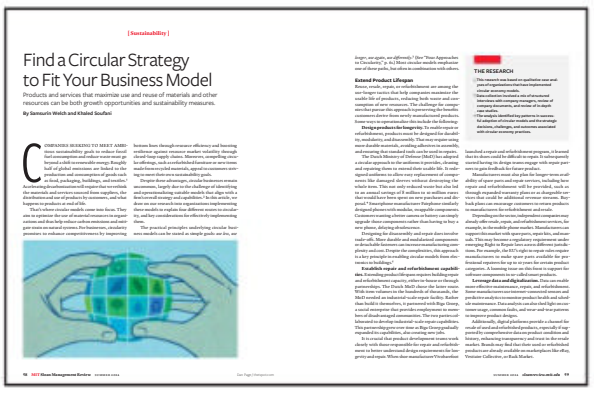
Acing Value-Based Sales

Marco Bertini, Oded Koenigsberg, and Todd Snelgrove pp. 52-57

Key Insight: Selling value requires more than quantifying benefits. It demands a deep commitment to collaborating with customers.

Top Takeaways: Value-based selling efforts often fizzle after an initial push because companies fail to see beyond the numbers when calculating the economic impact of product or service benefits. While quantifying evidence of benefits is at the heart of value-based selling, it's not enough. Companies can follow a five-step process that establishes a basis for joint value creation with customers and enables them to establish a deep mutual understanding of how both the buyer and seller can benefit from and share in that value.

REPRINT 65425


[Sustainability]

Find a Circular Strategy to Fit Your Business Model

Products and services that maximize use and reuse of materials and other resources can be both growth opportunities and sustainability measures.

By Samsurin Welch and Khaled Soufani

Find a Circular Strategy to Fit Your Business Model

Samsurin Welch and Khaled Soufani pp. 58-63

Key Insight: Analyzing your products' full life cycles can enable your organization to extract more value from materials and resources, and improve both sustainability and profitability.

Top Takeaways: Circular business models consider the full lifespan of a product, taking into account how materials are sourced; how the product is manufactured, distributed, and used; and what happens to the product at its end of life. When companies operate in this way, they stand to boost their own efficiency and competitiveness while better meeting their sustainability goals. The authors describe four types of circular business models — extending product lifetimes, tapping idle or wasted capacity, reclaiming material resources, and using resources more frugally — and discuss the capabilities required to execute each approach.

REPRINT 65415


[Growth & Development]

How to Come Back Stronger From Organizational Trauma

Traumatic events are destabilizing. In their aftermath, leaders can help individuals and teams recover and grow.

By Payal Sharma

How to Come Back Stronger From Organizational Trauma

Payal Sharma pp. 64-68

Key Insight: Leaders play a key role in promoting the growth and development that can emerge after trauma.

Top Takeaways: Traumatic events in the workplace, whether through acts of violence or natural phenomena, leave individuals and organizations reeling and destabilized. However, psychology research has identified a phenomenon called *post-traumatic growth* (PTG), whereby such experiences can enable survivors to develop new capabilities and a more nuanced understanding of the world. In this article, the author looks at how leaders can apply the concepts behind PTG to help their organizations move forward after a traumatic incident and help employees rebuild their sense of safety, control, protection, and purpose.

REPRINT 65420

Engineer Your Own Luck

Mark J. Greeven, Howard Yu, and Jialu Shan
pp. 69-72

Key Insight: No one can predict the future, but modularizing core capabilities can position companies to be ready for the unexpected.

Top Takeaways: Businesses that can move quickly to capture new opportunities don't have greater foresight than the rest. What they do have, increasingly, is *optionality*: the ability to discover and act on their options. One way to maximize an organization's options for growth is by treating internal capabilities as plug-and-play, mix-and-match modular digital services. This approach can enable businesses to scale quickly, and help their partners and clients grow new market segments.

REPRINT 65434

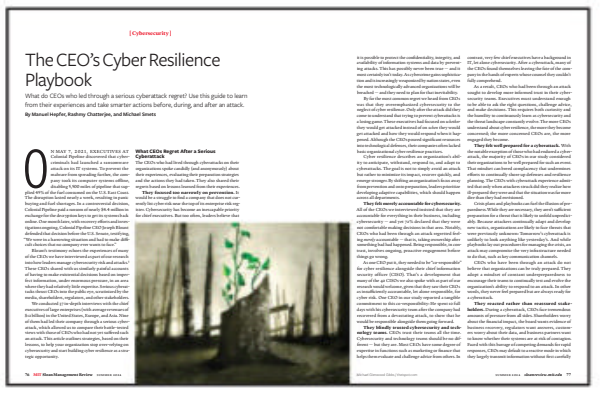
Serve More Customers With Inclusive Product Design

Vanessa M. Patrick and Jeffrey D. Shulman
pp. 73-75

Key Insight: Centering on the most marginalized users in the product development process can result in features that benefit all customers.

Top Takeaways: Products designed for the average user can unintentionally exclude those with special needs or from marginalized populations. Considering diversity along multiple dimensions — including age, disability, religion, and nationality — can help product teams widen their view of their target market. Even small increases in attentiveness to underserved users' needs can yield significant results by making products usable by a larger population. Pilot projects can help a team analyze how this kind of attentive engagement will affect product development, including time and financial costs and product performance.

REPRINT 65403



The CEO's Cyber Resilience Playbook

Manuel Hepfer, Rashmy Chatterjee, and Michael Smets pp. 76-79

Key Insight: CEOs are critical to orchestrating a shift from simply playing cybersecurity defense to developing an action plan for handling a cyberattack when — not if — it occurs.

Top Takeaways: Many organizations over-rely on cybersecurity defenses at the expense of building cyber resilience. CEOs whose companies have been through a malware attack or hacking incident stress the importance of accepting that cyberattacks are inevitable and making business continuity and recovery plans long before they're needed. Drawing on their conversations with such CEOs, the authors share best practices for building an organization's cyber resilience and managing stakeholders during these incidents.

REPRINT 65406

MIT Sloan Management Review (ISSN 1532-9194) is published quarterly by the Massachusetts Institute of Technology at the MIT Sloan School of Management. Periodicals postage paid at Boston, MA, and additional mailing offices. POSTMASTER: Send address changes to MIT Sloan Management Review, 1 Main Street, 9th floor, E90-9200, Cambridge, MA 02142, USA. Canadian GST No. 12761 8635 RT0001.

Ask Sanyin:

What's the Right Way to Carry Out Layoffs?

We are going through an organizational transformation, which will involve a workforce reduction. How can we do this quickly and efficiently?

IT'S TEMPTING TO DEAL WITH A difficult situation by getting it over with and moving on. But leaders making decisions that have a significant impact on employees' lives should aim to be thoughtful and compassionate. Layoffs can shatter trust and leave remaining employees feeling insecure. Tending to emotions and caring about the ramifications for those who are let go as well as those who remain is critical.

The best example of how to do this comes from a leader I know who oversaw a reorganization that eliminated several thousand jobs. This leader was deeply invested in the whole process, and the compassion she demonstrated bolstered support for her leadership. I took the following lessons from her experience.

1. Sweat the hardest question: Who is on the list? It can be easy to delegate the tough work of deciding who will go. "Here's the number you need to hit. Figure out how you're going to do it, and talk to me when it's done," say many leaders. It's tempting to avoid the human factor because it's messy.

The leader I'm thinking of cleared her schedule for several days and worked with the senior team to determine whom the layoffs would hit. She considered what type of work the company would need to stop doing or would be unable to take on in the absence of specific employees. She considered each worker for who they were: real people. She wanted to know the name of every person potentially affected. She also encouraged teams to roll out the plan as soon as they knew who would be affected, to mitigate the anxiety of uncertainty.

2. Prepare for the difficult conversations. The leader then considered how to support the people who were going to break the news to individuals being laid off. She made sure managers had all the tools they needed to handle these difficult conversations with respect and compassion for the people affected.

Sanyin Siang *is a CEO coach and leads the Fuqua/Coach K Center on Leadership & Ethics (COLE) at Duke University. Need advice? Send an email to Sanyin at* asksanyin@mit.edu.

The leader made sure she was attentive and engaged in helping managers with conversation prep, and she set the tone of care and humanity. By keeping the human rather than the role at the center, she made every effort to ensure that the relationships with those being let go continued to be positive. Because she leaned into an uncomfortable part of the work, the team leaned in as well. Managers were able to lead with compassion for those being laid off, for those remaining, and for themselves.

3. Invest in those who remain and mind their emotional energy. In addition to considering the experience of those being laid off, this leader showed concern for how her teams would feel afterward. As managers met with those affected by the layoffs, she continued to check in with them and their teams and offer support. As they approached the end of the day, she wanted to know when they were done so that she could send out a communication to the whole organization recognizing how difficult the day had been.

How a leader handles workforce reductions can reveal — and also strengthen — what values matter in your culture. By taking a human-centric approach rather than prioritizing efficiency, leaders can set an example for other leaders and create stronger teams in the long run. ■

Reprint 65439. For ordering information, see page 4.

Hanane Kai

BCG X | BCG

Artificial Intelligence

Human Potential

Unlock value
with GenAI

How can AI transform your business?

AI Strategy and Implementation

Integrate AI and analytics with your digital strategy for accelerated impact.

IMD / Real learning Real impact