ARTIFICIAL INTELLIGENCE

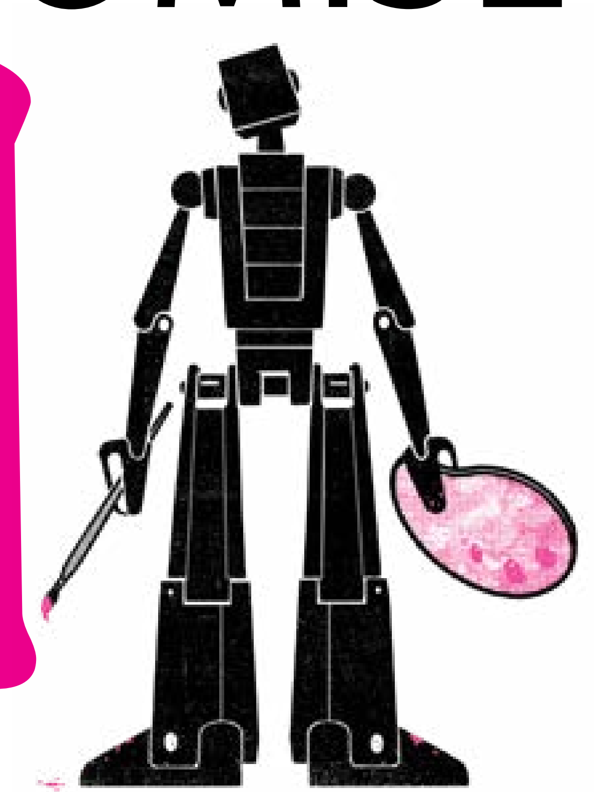
Making Good on the Promise of AI

A closer look at the issues companies face in deriving benefit from AI — while avoiding its downsides.

Brought to you by:

bold360 by LogMeIn

MAKING GOOD ON THE PROMISE OF AI



DESPITE THE POTENTIAL of artificial intelligence to transform processes, products, and customer relationships, few companies are well equipped to realize its promise. Most know they have to develop new skills and mindsets — and yet they don't know quite where to begin. This special report provides some insight.

In the lead article, Senén Barro and Thomas H. Davenport discuss what it takes to acquire and develop the human capital needed to innovate more deeply with AI. This is just one area where organizations can step up their game. They can reimagine the work they're doing and how they're doing it, not just the offerings they're creating.

Another area where AI efforts tend to fall short of their promise is strategy. As David Kiron and Michael Schrage explain, having a cogent strategy *for* AI is what most businesses aspire to do. But creating strategy *with* AI and machine learning — using these technologies to select the right key performance indicators and prioritize them appropriately — matters just as much, maybe more.

To capture value through AI-enhanced operations as well, organizations must develop a capacity for what's called enterprise cognitive computing. In their research, Monideepa Tarafdar, Cynthia M. Beath, and Jeanne W. Ross have identified several critical capabilities and key practices that can help businesses radically improve their processes with AI.

Of course, an important part of realizing AI's potential is managing its risks. As Julian Friedland points out, AI applications are allowing us to outsource more and more of our cognitive and emotional labor. As a consequence, he argues, our capacity for moral self-awareness and critical reflection is suffering. So he urges creators of these tools to restore some of the friction they've removed from our lives. It serves an essential purpose.

—The Editors

SPECIAL REPORT

DAN PAGE/THEISPOT.COM

People and Machines: Partners in Innovation

The greatest impact of intelligent technologies won't be from eliminating jobs but from changing what people do and driving innovation deeper into the business.

BY SENÉN BARRO AND THOMAS H. DAVENPORT

Thoughtful adoption of intelligent technologies will be essential to survival for many companies. But simply implementing the newest technologies and automation tools won't be enough. Success will depend on whether organizations use them to innovate in their operations and in their products and services — and whether they acquire and develop the human capital to do so.

In a recent Deloitte survey of 250 executives familiar with how their companies are thinking about intelligent technologies, nearly three out of four said that they expected AI to substantially transform their organizations within three years.[1] Of course, the workforce will be deeply affected by all this change. Yet even as AI eliminates some jobs in the coming decade (it most certainly will), it may create as many positions as it kills and open up vast new opportunities for collaborations between humans and machines. Earlier talk of large-scale job loss[2] has subsided somewhat. In the Deloitte survey, for example, reducing head count through automation was the lowest-ranked objective for AI — only 7% of the respondents selected that as their first priority. Indeed, many observers are shifting their

DAVIDE BONAZZI/THEISPOT.COM

expectations away from job loss to job change, as humans find ways to work closely with machines.

Given the likelihood that many jobs will change rather than disappear, organizations need to understand the new skills required. In a recent McKinsey survey of executives at companies with revenues of more than $100 million, 66% of respondents said "addressing potential skills gaps related to automation/digitization" within their workforce was a "top 10 priority"; 64% of the U.S. respondents and 70% of the European respondents said they needed to retrain or replace at least a quarter of their current workforce.[3] Significantly, just 16% of the business leaders responded that they were "very prepared" to address potential skills gaps, raising serious questions about their readiness to compete. Other recent surveys suggest that the high expectations executives have for intelligent technologies exceed their skills and experience in integrating such technologies into their companies.[4]

Although we have observed and worked with many large companies and startups on AI issues, we know of very few that have begun significant job redesign, re-skilling, or retraining programs. Moreover, most individuals aren't being adequately re-skilled or retrained for automation-enabled work.[5] (See "About the Analysis," p. 24.) Smart organizations will take steps not just to adopt intelligent technologies but also to recruit and retrain people for skilled roles, redesign tasks and jobs, and use AI as an enabler of innovation in products, processes, and business models in what we call *innovation based on intelligent automation*. This will be a job-by-job, task-by-task transformation, but we can already see places where major advances in technology are being undermined by insufficient attention to integration and human capital. Surgeons, for example, are increasingly using robotic technology to assist them in routine surgery; the new technology provides better vision, more-precise incisions, and neater sutures. However, few hospitals and medical schools have developed effective approaches for training surgical residents on the technology; surgeons in training get no hands-on experience.[6]

Although the potential for AI-enabled innovation exists in virtually every aspect of business and society, it is largely unrealized. A study of internal audit organizations, for example, found that less than one-third of the audit teams had road maps for incorporating new technologies.[7] Technology vendors are conceiving and producing innovations ranging from self-driving cars and trucks to the "self-driving enterprise,"[8] but very few would-be adopters have begun to envision how AI will change jobs in their companies and what new skills must be developed. Because many new AI technologies are appearing now or will be here in the near future, organizations have no time to waste in planning for them and creating work design innovations that parallel the technological innovations.

A Spectrum of Intelligent Automation

When intelligent technologies support individual workers, allowing them to do their jobs better or more efficiently, what we're really talking about are *tools* rather than automation. A good example is a taxi driver who uses GPS for driving directions. Automation goes a step further: It allows tasks or processes to be carried out without human assistance or participation, but humans may supervise the work or perform adjacent or complementary tasks. For instance, intelligent diagnostic systems can read X-ray images, but radiologists are still needed to define the imaging to be performed, relate imaging results to other medical records and tests, discuss findings with patients, and perform other activities.[9]

Although the earliest applications involved manual and systematic (structured and repeatable) cognitive tasks, we are moving toward nonsystematic cognitive tasks that include creativity and job variability, which until recently seemed beyond the scope of automation. And we are progressively adding greater autonomy to products and services. (See "AI at Work," p. 25.)

In fact, we are beginning to see autonomous systems that can perform tasks without any human involvement at all, using carefully prescribed guidelines. Consider automated financial trading. Because it depends entirely on algorithms, companies can complete transactions much faster with it than with systems relying on humans. In a similar fashion, robots are performing narrow tasks autonomously in manufacturing settings. In 2015, for example, Changying Precision Technology, a Chinese company involved in the production of mobile phones, replaced 90% of the workers in one of its plants with robots. In doing so, the company says it was able to

THE LEADING QUESTION
How can companies use intelligent technologies to innovate?

- **They can redesign tasks and jobs to facilitate human-machine cooperation at work.**
- **They can automate products, processes, and business models to support users' needs.**
- **They can integrate intelligent technologies into their organizations.**

ABOUT THE ANALYSIS

Over the past four years, the authors conducted research or consulting work at more than 50 companies pursuing initiatives in artificial intelligence. Both authors participate in several AI-related startups as cofounders or advisers. In addition, Thomas Davenport has been a coauthor of Deloitte's State of AI in the Enterprise surveys[i] and has worked with ServiceNow on its global worker survey on automation issues.[ii] Drawing on those sources of data and insight, this article looks at opportunities for companies to match the innovation in AI technology with innovations in work redesign and enhancement of employee capabilities.

more than double its output and slash defects by 80%.[10] More commonly, however, AI and robotics change jobs rather than eliminate them. Amazon, for example, has hired more than 300,000 people since its 2018 purchase of Kiva Systems, a maker of warehouse robots. One distribution center employee, who "babysits" several robots and ensures they have bins to load, commented on her job: "For me, it's the most mentally challenging thing we have here. It's not repetitive."[11]

While today most AI systems augment only existing workers, many people believe it's just a matter of time before complex systems will be able to operate by themselves in unstructured and dynamic environments. For example, in the next two or three years we will have self-driving vehicles capable of operating in limited spatial areas or under special circumstances (classified by the Society of Automotive Engineers as Level 4); by 2030 or so, many anticipate vehicles that operate without human intervention at all (classified as Level 5).[12]

Furthermore, there is the growing possibility that in the not-too-distant future we will have machines that can operate according to their *own* goals. An example that's as immediate as it is frightening is autonomous weaponry that will be able to decide where, when, and against whom it uses its capacity for destruction. This application represents the negative Mr. Hyde aspect of fully autonomous systems. But eventually, we can also expect to see a Dr. Jekyll side, with applications that have the potential to make life better.

Minds Working With Machines

Just as semiconductors enabled us to reduce the cost of calculations and apply arithmetic to new areas — first to scientific and military applications and later across all professional and social spheres — innovation based on AI will unleash an avalanche of both improved and entirely new products and services. The impact on the world of work will be unprecedented.

Human versus machine matchups in chess illustrate how humans will need to continuously change their roles relative to smart machines. Back in 1996 and 1997, IBM's Deep Blue competed against world champion Garry Kasparov and became the first computer ever to beat a world champion in a six-game match. As with other chess programs, Deep Blue's strategy blended computing power and strategic knowledge of the game provided by human experts. People could sharpen their skills by playing against it and studying its moves, but they wouldn't learn anything new, per se. But now the competence of chess programs has risen to the point where many chess masters use them to improve their own level of play. At the end of 2017, a new chess milestone was achieved when AlphaZero software, developed by Alphabet's DeepMind, learned how to play solely on the basis of its knowledge of the rules.[13] In less than one day of playing against itself, AlphaZero learned enough to crush Stockfish, which had previously been the leading chess program.[14] Among chess experts, one of the most surprising things about AlphaZero is that it has learned strategies that extend beyond how humans play. Humans taught Deep Blue to play chess, but AlphaZero developed its own approach — one that humans could learn from.

Such changes in the human-machine relationship will emerge in the workplace, too, as AI becomes increasingly intelligent. It will not be a spontaneous process but will be induced by designers and users of intelligent technologies and, of course, by companies that innovate on the basis of such technologies and have the right human resources in place to make it happen. However, major changes in jobs and skills don't coalesce overnight, even when the approach involves hiring new employees instead of retraining existing ones. Once companies identify the needed changes, implementing them will take time.

In the future, organizations will need to place both adoption of technology and human capital development at the center of their innovation strategies. As time goes on, how companies deploy technology and human capital will have a tremendous impact on their competitiveness and their very survival. We see four basic scenarios playing out in the organizations we have worked with:

1. Minimal investment in automation technology and people. For a variety of reasons — including cost and lack of vision or knowledge, especially among executives — some companies delay making the kinds of fundamental decisions and commitments that will make them viable AI innovators in the future. In this scenario, they underinvest in the necessary technologies and human capital. Such reluctance to enact changes will inevitably lead to a loss of competitiveness and an inability to maintain a sustainable business. These companies will have higher labor costs, fewer intelligent products and services, and lower levels of customer service than their competitors. In wealth management, for example, companies without intelligent robo-advisers are already losing business to competitors such as Vanguard and Charles Schwab that offer low- or no-cost advice.

2. Heavy investment in automation technology but little investment in human capital. Some companies we have worked with are willing to make major investments in automation but are prepared only to make incremental changes in job design and training, expecting that the technology itself will bring about organizational transformation largely through improvements in efficiency and productivity.

Take chatbots, which many companies are using to handle relatively simple customer service tasks. Starbucks, for example, uses chatbots to notify customers when their orders are ready; Mastercard uses them to make it easy for customers to get information on their transactions. (For more complicated problems, human agents typically take over.) To the extent that such companies reconfigure jobs or processes and help workers learn how to work with the technology, the chatbots can provide synergies, or at least a better distribution of tasks. Unfortunately, automation doesn't always work this way. For example, in 2017, Tesla invested heavily in robots for manufacturing and underinvested in skill development for human workers. When it realized that the robots weren't doing enough to help the company meet ambitious production goals for its Model 3 cars, management backed away from its reliance on robots and hired and trained humans to perform the necessary tasks.[15] But for the final vehicle assembly, Tesla took a more nuanced, integrated approach, assigning humans to the complex tasks and using robots for specialized tasks such as moving goods around the factory, lifting heavy components, and testing seats. The result was, as one observer put it, "a delicate dance of human workers and robots on the production line."[16]

3. Incremental changes in jobs and skills with little investment in intelligent technologies. Many companies that prioritize incremental process improvement (for example, using Six Sigma or Lean programs) don't invest enough in new technology — in part because the methods don't include a role for technology. In addition, it can be difficult to adopt broad, cross-organizational changes in jobs and technologies at the same time because the impact of AI and other technologies on jobs tends to be specific to particular jobs. Although it's true that hiring and retraining skilled workers can generate short-term improvements, that approach alone won't lead to meaningful change. Indeed, we have found that unless companies are willing to commit resources to AI technologies, they risk falling behind competitors in both productivity and quality. Eventually, moreover, they hurt their ability to hire and retain quality knowledge workers, who may see better opportunities elsewhere. Of course, there are particular settings

AI AT WORK

Intelligent technologies vary in the amount of autonomy they provide.

EXAMPLE	LEVEL OF TASK AUTOMATION
GPS navigators	Support for human drivers
Medical diagnostic and fraud detection systems	Automation under human supervision
Automated financial trading and mobile robots in industrial environments	Autonomy with precise guidelines
Self-driving cars	Autonomy with general guidelines
Autonomous weapons	Capacity to set their own objectives

in which an emphasis on people-oriented strategies makes sense. High-end restaurants, for example, are less dependent on automation than are fast-food establishments. The same goes for fashion and other luxury businesses. But even in these cases, intelligent automation should have a growing presence in back-end functions and processes such as supply chain management and customer support.

4. Significant investment in both intelligent technology and human capital innovation. Organizations with a broad-based investment approach are best equipped to pursue innovation in both AI application and human capital development. Rather than simply looking at automation as a way to cut costs, these companies create innovative products, services, processes, and business models by implementing intelligent technologies, redesigning jobs, acquiring new skills through hiring, and training their existing workers. This approach is especially vital for companies that compete in markets dominated by global giants.

For example, GE — notwithstanding its current difficulties with its GE Power and Genworth Financial business units — is actively trying to use both AI applications and human capital to drive innovation. One way it is doing this is by studying the needs of different types of employee users, or personas, and then considering how they might be supported by technology. Personas are part of a widely used approach for understanding customer needs in marketing and product development, but they are rarely used for the development of internal systems and even less so to create AI systems.

One of the GE personas is made up of employees involved with buying or sourcing industrial materials. A key task for these employees is to ensure that the needed materials are available on the manufacturing line at the right time. Historically, they relied on their intuition to manage the delivery schedules, but machine learning models have the ability to learn from past deliveries and provide model-driven estimates. Users are being trained to understand how the models work and how they can be improved. Today, the models inform the sourcing manager, who makes the final decision about when to order. Eventually, GE expects the AI systems will be capable of making decisions on their own to optimize things like delivery schedules and in-process inventory. The role of humans will be to tweak the processes and address problems that occur.[17]

Despite the power of AI and other new technologies, the likelihood that they will replace managers and professionals in the near term is minimal. Rather, many observers, including Erik Brynjolfsson and Andrew McAfee, codirectors of the MIT Initiative on the Digital Economy, believe that the change will be more gradual — that those "who use AI will replace those who don't."[18] In our view, the challenge for companies is finding ways to ease intelligent technologies into their organizations, while simultaneously determining how to take advantage of what intelligent humans have to offer.

Think Before You Automate

There is no simple recipe for successful innovation based on automation. Different companies will have different opportunities to put intelligent technologies to work. However, in researching knowledge and technology transfer within companies and advising organizations on AI adoption, we have developed a set of guidelines:

Start with management education. The best starting point is to invest in training for the executives charged with making the strategy decisions about intelligent technologies. Based on our experience, executive ignorance often leads to two opposite but equally negative behaviors: If leaders underestimate the potential of these technologies, their companies will miss opportunities to benefit from them. On the other hand, if they overestimate it and initiate projects that are too ambitious and costly, they will waste resources and perhaps even generate a bias within the company against new projects, even those that are reasonable. To prepare leaders to make future decisions, a leading property and casualty insurance company, for example, held daylong sessions for top executives on what AI is, how best to manage it, and what it might mean for employees. Anthem Insurance Companies, a large health insurance corporation, and Bank of America have run similar sessions for their leaders and board members.

Develop a road map for future initiatives involving technology and people. As with any project, implementing an intelligent automation initiative requires having a road map that describes the objectives, the necessary resources, and the implementation

> At Amazon, CEO Jeff Bezos says that many of the company's investments in machine learning are focused on "quietly but meaningfully improving core operations."

schedule. A good road map should help the organization anticipate the potential benefits beyond the most obvious ones and should include a communication strategy, both internal and external, especially when intelligent automation projects might lead to a reduction in jobs. For example, Situm Technologies, a Spanish startup (of which one of us, Senén Barro, is a founder), developed technology that accurately tracks the location of people and assets via smartphones inside facilities such as hospitals, airports, and factories. The initial applications were fairly narrow — an early customer in the building-security business wanted to track the routes of its security guards. Eventually, however, the company developed a road map for using Situm's technology within facilities in other ways — for example, to manage people during emergency situations such as fires or assaults. This enabled the company to offer a set of solutions that aligns the benefit of optimizing human resources with safety.

Focus on immediately valuable projects and be wary of initiatives that are too ambitious. Companies that lack significant AI experience should focus initially on low-hanging-fruit projects that will enable them to gain experience. Highly ambitious projects to treat cancer, provide individual investors with detailed investment recommendations, or eliminate drivers from cars have all either failed or taken far longer than researchers expected. Even Amazon has had challenges with its Amazon Go stores, and its drone delivery project is taking a long time to emerge.

Combining several manageable projects in a single business area often has a better chance of yielding significant results than trying to pursue one big one. At Amazon, for example, CEO Jeff Bezos says that many of the company's investments in machine learning are focused on "quietly but meaningfully improving core operations."[19] If the company's strategic focus is on using AI to enhance customer relationships, for example, the component projects might include chatbots or intelligent agents to answer questions quickly 24-7, machine learning models to capture the "voice of the customer" from call center operations, recommendation engines to pitch promotions only to customers with high interest, and so forth. This incremental approach also creates more time to redesign work and re-skill workers, since each AI-supported task will typically require only incremental change in jobs. The objective should be clear; even in cases where the goal is automating tasks previously performed by workers, key workflows should be designed or redesigned, focusing on the division of labor between humans and smart machines. The aim throughout should be innovative and effective work design, not just cost reduction.

Invest in building internal staff capabilities. Identify the workers who will adopt the solution and train the staff in its use. Ideally, some people would be involved in the development of the AI system — serving, perhaps, as process or subject-matter experts. Given their expertise, they can be lead users of early versions of AI systems and provide feedback on what works and what doesn't. HR and corporate learning departments can partner with these individuals to structure training programs for other workers affected by the systems.

To innovate around intelligent automation, you should plan to develop or hire your own people as opposed to only borrowing them from consulting firms or vendors. For example, training chatbots requires a deep understanding of the business and current and evolving customer or internal user requirements, which are things that experienced employees inside the company can best provide.

Plan on making improvements over time. Obviously, whatever technologies you use should be suited for the projects at hand. However, intelligent technologies are improving quickly, which means that innovation based on automation needs to be continuous rather than episodic. For example, recent advances in natural language generation enable organizations to incorporate narrated reports into their business intelligence applications. This new capability may greatly increase the ability of nonexperts to understand technical and financial reports, which may decrease the need for human or AI-based customer service. Leading companies such as USAA, an insurance and financial services company, are working along multiple lines — chatbots, virtual assistants, and narrative generation — to facilitate better customer communications, and therefore they must constantly monitor the relationships among the various tools.

Managers need to recognize that intelligent technologies will find their way into more and more industry sectors and occupations in the coming years. Business solutions powered by AI will reduce costs and improve productivity. However, we expect that the greatest impact will be to drive innovation deeper into the business — and for that to happen, people and machines must be partners in the innovation process. Investing in intelligent technologies and in human resources capable of using them, cooperating with them, and innovating from them may be costly. But failure to do so will be much more costly.

***Senén Barro** (@senenbarro) is the scientific director of the University of Santiago de Compostela's Singular Research Center on Intelligent Technologies in Galicia, Spain, and a visiting scholar at Babson College in Wellesley, Massachusetts. **Thomas H. Davenport** (@tdav) is the President's Distinguished Professor of IT and Management at Babson, a fellow at the MIT Initiative on the Digital Economy, and a senior adviser to Deloitte's Analytics and Cognitive practices. His latest book is* The AI Advantage: How to Put the Artificial Intelligence Revolution to Work *(MIT Press, 2018). Comment on this article at http://sloanreview.mit.edu/x/60407.*

REFERENCES

1. T.H. Davenport, J. Loucks, and D. Schatsky, "Bullish on the Business Value of Cognitive: Leaders in Cognitive and AI Weigh In on What's Working and What's Next," Deloitte, 2017, www2.deloitte.com.

2. C.B. Frey and M.A. Osborne, "The Future of Employment: How Susceptible Are Jobs to Computerisation?" Technological Forecasting and Social Change 114 (January 2013): 254-280.

3. P. Illanes, S. Lund, M. Mourshed, et al., "Retraining and Reskilling Workers in the Age of Automation," McKinsey Global Institute, January 2018, www.mckinsey.com.

4. "Avoiding Setbacks in the Intelligent Automation Race," KPMG, accessed April 3, 2019, https://advisory.kpmg.us.

5. T.H. Davenport, "The Business Value of Digital Workflows," Workflow Quarterly (spring 2019), https://workflow.servicenow.com.

6. M. Beane, "Shadow Learning: Building Robotic Surgical Skill When Approved Means Fail," Administrative Science Quarterly 64, no. 1 (March 2019): 87-123.

7. M. Cohn, "For Internal Auditors, Innovation Is a Work in Progress," Accounting Today, March 12, 2019, www.accountingtoday.com.

8. S. Lauchlan, "After the Self-Driving Car, Welcome the Self-Driving Enterprise — and All Its Pyramid Organization Implications," Diginomica, May 2, 2018, https://diginomica.com.

9. T.H. Davenport and K.J. Dreyer, "AI Will Change Radiology, but It Won't Replace Radiologists," Harvard Business Review, March 27, 2018, https://hbr.org.

10. A. Prakash, "Forget the Markets, Robots Are China's New Worry," Forbes, Jan. 28, 2016, www.forbes.com.

11. N. Wingfield, "As Amazon Pushes Forward With Robots, Workers Find New Roles," The New York Times, Sept. 10, 2017.

12. "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles," SAE International, J3016_201806, revised June 15, 2018, www.sae.org.

13. D. Silver, T. Hubert, J. Schrittwieser, et al., "A General Reinforcement Learning Algorithm That Masters Chess, Shogi, and Go Through Self-Play," Science 362, no. 6419 (Dec. 7, 2018): 1140-1144.

14. AlphaZero learns through complex deep-learning algorithms, and it uses a so-called reinforcement learning approach similar to the way humans and other living beings learn. If a decision is made that is shown to be appropriate over time, a positive reinforcement is obtained that reaffirms that decision for the future; decisions that don't work are penalized.

15. H. Edwards and D. Edwards, "How Tesla 'Shot Itself in the Foot' by Trying to Hyper-Automate Its Factory," Quartz, May 1, 2018, https://qz.com.

16. S. Schrader, "This Time Lapse of a Tesla Model 3 Getting Built Is Weirdly Soothing," The Drive, Jan. 5, 2019, www.thedrive.com.

17. T.H. Davenport interview of D. Burns, chief information officer of GE Aviation, Feb. 5, 2018.

18. E. Brynjolfsson and A. McAfee, "The Business of Artificial Intelligence: What It Can — and Cannot — Do for Your Organization," Harvard Business Review, July 7, 2017, https://hbr.org.

19. K. Leswing, "Jeff Bezos Just Perfectly Summed Up What You Need to Know About Artificial Intelligence," Business Insider, April 12, 2017, www.businessinsider.in.

i. J. Loucks, D. Schatsky, and T. Davenport, "State of AI in the Enterprise, 2nd Edition: Early Adopters Combine Bullish Enthusiasm With Strategic Investments," Deloitte Insights, Oct. 22, 2018, www2.deloitte.com.

ii. Davenport, "The Business Value of Digital Workflows."

Reprint 60407.

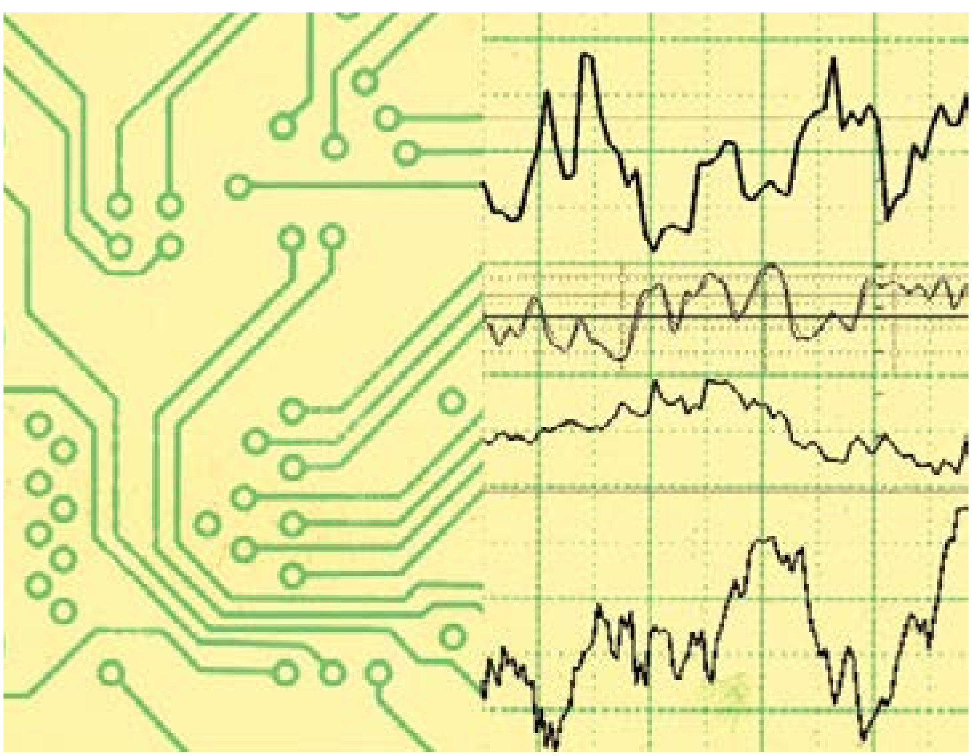
Strategy For and With AI

A company's strategy is defined by its key performance indicators. Artificial intelligence can help determine which outcomes to measure, how to measure them, and how to prioritize them.

BY DAVID KIRON AND MICHAEL SCHRAGE

Many executives, intent on understanding and exploiting AI for their companies, travel to Silicon Valley to acquaint themselves with the technology and its many promises. These pilgrimages have grown so common that tours now exist to facilitate inside peeks at innovative startups. Buoyed by hype and smatterings of algorithmic knowledge, returning executives share a common goal: determining what products, services, and processes AI can enhance or inspire to sharpen competitive edges. They believe a comprehensive strategy for AI is essential for success.

That well-intentioned belief is off the mark. A strategy *for* AI is not enough. Creating strategy *with* AI matters as much — or even more — in terms of exploring and exploiting strategic opportunity. This distinction is not semantic gamesmanship; it's at the core of how algorithmic innovation truly works in organizations. Real-world success requires making these strategies both complementary and interdependent. Strategies *for* novel capabilities demand different managerial skills and emphases than strategies *with* them.



Machine learning pioneers — Amazon, Google, Alibaba, and Netflix come to mind — have learned that separating strategies *for* developing disruptive capabilities from strategies deployed *with* those capabilities invariably leads to diminished returns and misalignments. Not incidentally, these organizations are intensely data- and analytics-driven. Their leaders rely heavily on metrics to define, communicate, and drive strategy. This reliance on quantitative measures has increased right along with their growing investment in AI capabilities.

Our research strongly suggests that in a machine learning era, enterprise strategy is defined by the key performance indicators (KPIs) leaders choose to optimize. (See "About the Analysis," p. 32.) These KPIs can be customer centric or cost driven, process specific or investor oriented. These are the measures organizations use to create value, accountability, and competitive advantage. Bluntly: Leadership teams that can't

RICHARD MIA/THEISPOT.COM

clearly identify and justify their strategic KPI portfolios have no strategy.

In data-rich, digitally instrumented, and algorithmically informed markets, AI plays a critical role in determining what KPIs are measured, how they are measured, and how best to optimize them. Optimizing carefully selected KPIs becomes AI's strategic purpose. Understanding the business value of optimization is key to aligning and integrating strategies for and with AI and machine learning. KPIs create accountability for optimizing strategic aspirations. Strategic KPIs are what smart machines learn to optimize. We see this with Amazon, Alibaba, Facebook, Uber, and assorted legacy enterprises seeking to transform themselves.

These principles have sweeping and disruptive implications. As "accountable optimization" becomes an AI-enabled business norm, there is no escaping analytically enhanced oversight. Boards of directors and members of the C-suite will have a greater fiduciary responsibility to articulate which KPIs matter most — and why — to shareholders and stakeholders alike. Transformative capabilities transform responsibilities. You are what your KPIs say you are.

Distinct Complements

Historical context and precedent are important: Blending strategy *for* and strategy *with* is hardly unique to AI and machine learning. John D. Rockefeller's Standard Oil, for example, dominated the petroleum market not just because the company had an effective strategy for capitalizing on the nascent railroad industry's emerging capabilities but also because it allowed those capabilities — logistical powers of transport and delivery — to shape its broader strategy. By ruthlessly exploiting scale and acquiring and designing fuel tank cars, Standard Oil consistently reaped disproportionate returns from a rapidly expanding physical network.[1]

More recently, incumbents grasped that they urgently needed a strategy for the internet to compete with disruptive born-digital startups. But those organizations discovered — sooner or later — that their strategies for the internet were contingent upon the success of their strategies with the internet. Retailers, for example, commonly use internet-based omnichannel strategies to compete on customer experience. They might start by building strong relationships with shoppers online, for example, but when those same customers go to physical store locations, geofencing apps alert the company to their imminent arrival. Staff is then primed to help facilitate customer pickups. These seamless experiences blend strategy with and for the internet.

Creating an enterprise strategy for developing or applying a capability is not organizationally, culturally, or operationally the same as cultivating a strategy *with* that capability. These activities are complements. A strategy for sustainability (such as lowering one's carbon footprint or reducing waste) should not be divorced from having a sustainable overall strategy enabling the business to operate in thriving communities. Similarly, a strategy for AI shouldn't be viewed as a substitute for creating a strategy with AI.

Where Opportunity Lies

What, then, does strategy *with* AI pragmatically mean? Like any corporate strategy, it expresses what enterprise leaders deliberately seek to emphasize and prioritize over a given time frame. Strategies articulate how and why an organization expects to succeed in its chosen market. These aspirations might involve, for example, superior customer experience and satisfaction, increased growth or profitability, greater market share, or agile fast-followership when rivals out-innovate the company.

Whatever the specific strategy, virtually all organizations create corresponding measures to characterize and communicate desirable strategic outcomes. Those metrics — be they KPIs, objectives and key results (OKRs), or a Balanced Scorecard — are how organizations hold humans and algorithms accountable. For public companies, strategic KPIs typically respect and reflect investor concerns; for private equity, strategic KPIs might be calibrated to maximize a sale price or facilitate an IPO. Data-driven systems, enhanced by machine learning, convert these aspirations into computation. World-class organizations can no longer meaningfully discuss optimizing strategic KPIs without embracing machine learning (ML) capabilities.

Uber, for example, runs hundreds of ML models to optimize its ride-sharing platform and food-delivery

THE LEADING QUESTION
What does creating strategy with AI involve?

- **Organizations must first realize that their key performance indicator (KPI) portfolio represents their strategy.**
- **They can then use machine learning applications to choose, measure, and optimize their KPIs.**
- **They must manage their data as an asset in order to enhance their KPIs and help their machines learn.**

ABOUT THE ANALYSIS

This article draws on results from a 2018 survey of 3,225 business executives, managers, and analysts from companies based in 107 countries and 20 industries. To complement our survey analysis, we conducted 30- to 60-minute interviews with 17 executives and academics about the role of KPIs as a leadership tool. Some related findings were published in the 2018 *MIT SMR* report "Leading With Next-Generation Key Performance Indicators." This article extends that discussion by drawing out the implications of machine learning and AI for both identifying and optimizing strategic metrics.

business. Uber has made enormous investments in its machine learning capabilities and implementations. Whether it enjoys an abundance of available cars on call or relies on relatively few, its ability to estimate accurate arrival times for customer and driver alike is essential to how it competes in the marketplace.

"Accurate ETAs are critical to a positive user experience," observes Jeremy Hermann, who heads Uber's machine learning platform, "and these metrics are fed into myriad other internal systems to help determine pricing and routing. However, ETAs are notoriously difficult to get right."[2]

Yet, so many critical outcomes are dependent on robust ETA analytics — rider and driver expectations, fares, food pickup and delivery — that ETA is a core Uber metric. Hermann notes, "Uber's Map Services team developed a sophisticated segment-by-segment routing system that is used to calculate base ETA values. These base ETAs have consistent patterns of errors. The Map Services team discovered they could use a machine learning model to predict these errors and then use the predicted error to make a correction. As this model was rolled out city by city (and then globally …), we have seen a dramatic increase in the accuracy of the ETAs, *in some cases reducing average ETA error by more than 50%*."[3] [emphasis added]

Simply celebrating effective and globally scalable machine learning models misses the larger point. Uber cannot deliver on operational or strategic aspirations without reliably delivering on its ETA KPI. Chaotic ETA outcomes would prevent Uber from being a "low cost" or "best value" provider of mobility/delivery services. Technical, organizational, or operational changes that might threaten ETA outcomes are counterproductive. Uber must marginalize or minimize KPIs that might conflict or compete with effective ETA prediction.

Clarifying those constraints is crucial. In the words of Harvard Business School's Michael Porter, "The essence of strategy is choosing what not to do."[4] Once those guardrails are established, identifying and minimizing unwelcome consequences becomes as important as promoting the outcomes you want. The essential takeaway here is that prioritizing KPIs — ranking them according to what matters most and what the organization must learn the best — is essential to enterprise strategy. In an always-on big data world, your system of measurement is your strategy.

Determining the optimal "metrics mix" for key enterprise stakeholders becomes an executive imperative. Are customer-centric strategies, for example, better optimized via customer lifetime value (CLV) or balanced blends of earnings before interest, taxes, depreciation, and amortization (EBITDA) and net promoter score? For what customer segments should profitability be privileged over satisfaction or loyalty? As algorithms get smarter, leaders must have the courage to explore how best to answer these questions. AI makes that feasible, affordable, and desirable.[5]

This optimization imperative, our research suggests, demands a rigorous rethinking of the metrics chosen to define desirable (and undesirable) strategic outcomes. When machine learning measures management and manages measurement, metrics don't just reflect strategy but drive it. Achieving KPI outcomes (and suggesting new KPIs) is what smart machines need to do — and *need to learn to do*.

AI is not just about building products, services, or processes. Leaders need to recognize that AI must be primarily about enhancing the formulation and execution of strategy. To the extent that KPIs are essential to formulating and communicating strategy, strategy is quintessentially a system of measurement. Our research shows that AI transforms the strategist's choices about which KPIs to optimize and how to optimize them. Strategy is about optimizing KPIs with AI/ML.

Looking Forward and Backward

Machine learning profoundly changes how to approach optimizing leading and lagging KPIs. McDonald's has a multipart growth plan explicitly combining the two types of indicators. A key strategic aspiration is to once again be a family destination that appeals to parents. A lagging indicator is more visits by families with kids under the age of 13. A leading indicator is any evidence of becoming "a place I'm happy to bring my children," says McDonald's global chief marketing officer Silvia Lagnado.

Reliably measuring "happy place to bring my children" is methodologically challenging. Customer surveys are limited to those who fill them out, a source of selection bias. Machine learning-based sentiment analysis improves on this approach: It can classify large volumes of geotagged Twitter data and other data sets to correlate neighborhood-level well-being with comments about fast-food locations. A group of University of Utah academics developed a blueprint for this type of ML application.[6] Such machine learning mashups are becoming standard practice in academic and business research.

With machine learning, McDonald's can more effectively pursue high-priority KPIs. Marketers exploring in-store promotions with family-oriented advertising and menu options might improve family traffic but will fail if those promotions produce store conditions that annoy parents. Maximizing sales or revenues cannot come at that cost. Striking a productive balance between those measures is what optimization means. That's what McDonald's machines need to learn to serve up.

Not coincidentally, in March 2019, McDonald's announced its $300 million acquisition of Israel-based Dynamic Yield, which uses machine learning and big data to make personalized recommendations. McDonald's says it intends to use the company's tools to customize the drive-thru experience by creating dynamic digital menu boards that recommend menu items based on local demographics, previous orders, weather, and time of day, among other factors.

GoDaddy, the multibillion-dollar web-hosting and internet registry innovator, is also embracing leading as well as lagging data-driven KPIs. Since 2016, the Scottsdale, Arizona-based company's market value has grown more than 2.5X in no small part due to its dual commitment to strategic KPIs and machine learning. "We're very excited about the prospect of using the large data sets that we have," observes GoDaddy COO Andrew Low Ah Kee, "[to] train a model to solve and optimize against [customer] lifetime value as opposed to solving for transactional period revenue."[7]

Low Ah Kee's essential insight is that leaders have the duty and responsibility to pick which time horizons and "objective functions" to optimize. GoDaddy's emphasis on customer lifetime value (which anticipates future revenues, costs, and loyalty in addition to capturing past purchase behavior) reduces short-termism and threats to customer experience quality, he asserts. "We see in our customer base, when we help our customers succeed, the lifetime value it brings to us is significantly higher than for people whom we approach with just a transactional view," he notes. "As you start to extend the time horizon, I think the degree of [organizational] misalignment tends to go down." It's easier to miss long-term goals if the focus is on short-term tactics.

Making Smarter Trade-Offs

We argue that strategy is best understood and experienced as how the business invests in, manages, and prioritizes its KPI portfolio. KPIs and the relationships between them are the critical strategic units of

LEARNING WHAT TO OPTIMIZE

Optimizing known KPIs is important but not strategically sufficient. When appropriately trained, machine learning models can learn to identify and recommend novel or emergent KPIs. That is, machines can "learn to discover" enterprise KPIs on their own, without expert guidance. This is the difference between supervised and unsupervised learning. GE Healthcare CMO Glenn Thomas explains that his data science teams are "actually boiling out the KPIs from the data rather than setting the KPIs to be measured."

While Thomas declines to disclose emergent KPIs produced this way, an important irony cannot be overlooked: Thomas and his marketing/data team increasingly use machine learning to find KPIs they might never have discovered on their own. In marketing, promotion, interaction, and engagement domains, technology can go beyond "learning to optimize" to suggest what can and should be optimized.

analysis. Strategic success means the company's machines learn to optimize KPI portfolio returns.

To be clear, optimization in this context does not mean maximization. On the contrary, it means computationally learning to advance toward desired strategic outcomes through carefully calculated and calibrated KPI trade-offs. Understanding trade-offs among and between competing — and complementary — KPIs is essential. Simply optimizing individual KPIs by priority or rank ignores their inherent interdependence. For any KPI portfolio, identifying and calculating how best to weight and balance individual KPIs becomes the strategic optimization challenge. (See "Key Performance Indicators and Ethical Strategy.")

Even as "yield management" machine learning models for airlines, hotels, and other travel-related businesses algorithmically improve, strategic challenges sharpen: How can revenue-enhancement KPIs be optimized in the context of customer satisfaction and net promoter score KPIs? Do loyal customers deserve preferential rates or service bundles relative to typical customers? Learning to optimize for "best customers" draws on different data sets and expectations than learning to optimize for typical or average customers. What does an optimal balance between loyal customers and asset monetization margins look like? Smart machines can learn to strike that balance, but preemptively minimizing human insight and oversight seems foolish.

Similarly, high-frequency algorithmic traders may seek to maximize the frequency of profitable trades and/or maximize hourly, daily, or weekly profits. Yet, at the same time, they may wish to avoid or minimize the risk of regulatory intervention. One KPI maximizes profit (or "profits per trade" or "profits per trading strategy") while another signals that the company's trading patterns are unlikely to trigger an external review. Again, smart machines can learn to strike that balance. What is the risk appetite, not for particular trades but for particular regulators?

Every organization confronts this clash and conflict of strategic prioritization. No right answer exists. That said, some KPIs deliver disproportionate value and insight into helping company leaders better — or more optimally — achieve their strategic aspirations. Weighting these measures and metrics lends itself to machine learning applications. They facilitate alignment between local optima and the desired global optimum. Consequently, there can be no meaningful discussion about "optimal" strategic trade-offs in a KPI portfolio without a machine learning/AI capability.

The Essential Role of Data

There is no enterprise strategy for or with AI without an enterprise strategy for — and with — data. It is the essential ingredient for machine learning and dynamic optimization. As the Uber, McDonald's, and GoDaddy examples affirm, optimizing strategic KPIs — ETAs, happy families, CLV — is contingent upon data volume, velocity, variety, and quality.

That makes data governance key. Organizations must invest in recognizing which data might enhance or elevate their KPIs — and which data will help their machines learn. Digital processes and platforms that combine and analyze data, siloed and scattered, empower the company's artificial intelligentsia.

Technology titans and a growing number of legacy companies embrace comprehensive data strategies and practices. They explicitly, ruthlessly, and relentlessly manage data as an asset. This, as much as their technical prowess, sets them apart operationally and culturally. They employ chief data officers, data scientists, and data wranglers,

KEY PERFORMANCE INDICATORS AND ETHICAL STRATEGY

Google's YouTube division introduced two new internal metrics in the past two years for gauging how well videos are performing, according to people familiar with the company's plans. One tracks the total time people spend on YouTube, including comments they post and read (not just the clips they watch). The other is a measurement called "quality watch time," a squishier statistic with a noble goal: to spot content that achieves something more constructive than just keeping users glued to their phones.

The changes are supposed to reward videos that are more palatable to advertisers and the broader public, and to help YouTube ward off criticism that its service is addictive and socially corrosive. Creating the right metric for success could help marginalize videos that are inappropriate or popular among small but active communities with extreme views. It could also help YouTube make up for previous failures in curbing the spread of toxic content.

> Technology titans and a growing number of legacy companies explicitly and relentlessly manage data as an asset. This, as much as their technical prowess, sets them apart operationally and culturally.

holding people and processes accountable for getting value from data. Increasingly, much of that value comes from how quickly, accurately, and reliably that data trains machines.

Unfortunately, crisp and clear alignment between enterprise data governance and strategic AI initiatives remains elusive. A recent *Forbes Insights* CXO survey on AI and machine learning revealed that three out of four top executives declared AI a core component of their digital transformation plans. However, only 11% of the surveyed executives said their companies have begun implementing an enterprisewide data strategy, and only 2% said they have a serious "data governance" process in place.[8]

These findings, unhappily consistent with our own, suggest that successful and sustainable implementations of AI/ML-enabled optimization strategies are unlikely until data is explicitly treated as an asset. Organizations need effective data platforms and processes to enable effective machine learning platforms and processes. Ironically (even perversely), many companies have enormous amounts of timely, relevant, and valuable data for strategic AI efforts but lack the commitment and competence to harness it. Their data doesn't inform their KPIs or their strategy. An unwillingness or inability to use strategic KPIs to prioritize or align data assets with strategic outcomes further undermines their AI aspirations. These gaps render strategies for/with AI impotent.

LIKE ROCKEFELLER'S RAILROADS and the internet, artificial intelligence and machine learning represent enormously powerful strategic capabilities. They computationally transform the economics of optimization for business. Appropriately developed and deployed, they can literally learn how to create more value for more customers at lower cost and with greater speed. A strategy *for* AI matters less than clearly articulating the strategic aspirations, goals, and outcomes that leaders wish to optimize. Machine learning, like transportation and communication, is a means to an end. What needs to be transported? What needs to be communicated? What needs to be optimized? Artificial intelligence and machine learning can, in principle and practice, offer actionable answers to these questions. The true strategic opportunity and impact of these technologies is the chance to rethink and redefine how the enterprise optimizes value for itself and its customers.

David Kiron *(@davidkiron1) is the executive editor of* MIT Sloan Management Review*.* ***Michael Schrage*** *is a research fellow at the MIT Sloan School of Management's Initiative on the Digital Economy. Comment on this article at http://sloanreview.mit.edu/x/60416.*

REFERENCES

1. In fact, Rockefeller's ability to obtain railroad rebates was a significant source of competitive advantage. These rebates were eventually deemed unfair to competitors and contributed to the breakup of Standard Oil. See D.A. Crane, "Were Standard Oil's Railroad Rebates and Drawbacks Cost Justified?" Southern California Law Review 85, no. 3 (March 2012): 559-572.

2. J. Hermann and M. Del Balso, "Scaling Machine Learning at Uber With Michelangelo," Uber Engineering, Nov. 2, 2018, https://eng.uber.com.

3. Ibid.

4. M.E. Porter, "What Is Strategy?" Harvard Business Review 74, no. 6 (November-December, 1996): 61-78.

5. See, for instance, A. Agrawal, J. Gans, and A. Goldfarb, Prediction Machines: The Simple Economics of Artificial Intelligence (Boston: Harvard Business Review Press, 2018).

6. Q.C. Nguyen, D. Li, H.W. Meng, et al., "Building a National Neighborhood Dataset From Geotagged Twitter Data for Indicators of Happiness, Diet, and Physical Activity," JMIR Public Health Surveillance, no. 2 (Oct. 17, 2016): e158.

7. For more details on how McDonald's, GoDaddy, and others use machine learning to optimize KPIs, see M. Schrage and D. Kiron, "Leading With Next-Generation Key Performance Indicators," www.sloanreview.mit.edu, June 26, 2018.

8. "Closing the Corporate Gap on AI," Forbes Insights, Sept. 21, 2018.

Reprint 60416.

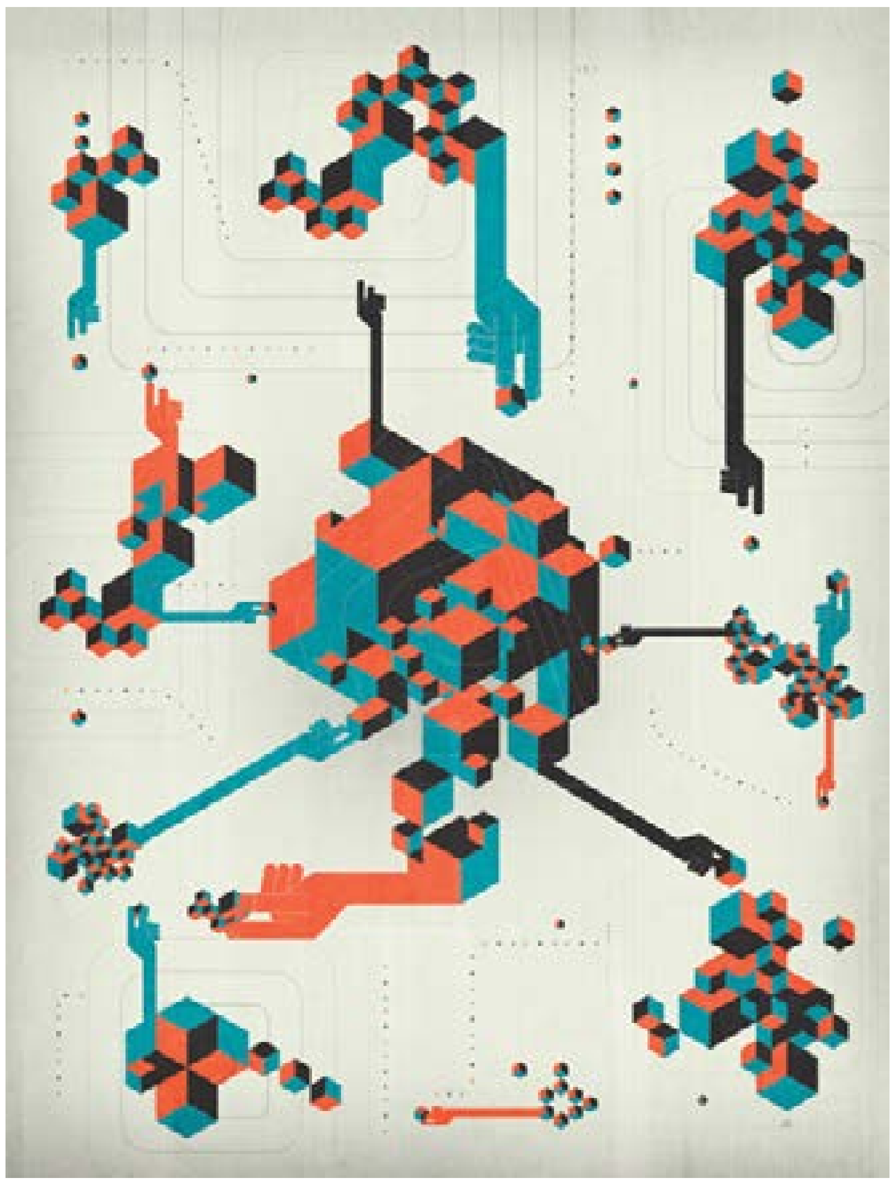
Using AI to Enhance Business Operations

How organizations can improve processes and capture value through enterprise cognitive computing.

BY MONIDEEPA TARAFDAR, CYNTHIA M. BEATH, AND JEANNE W. ROSS

Artificial intelligence invariably conjures up visions of self-driving vehicles, obliging personal assistants, and intelligent robots. But AI's effect on how companies operate is no less transformational than its impact on such products.

Enterprise cognitive computing — the use of AI to enhance business operations — involves embedding algorithms into applications that support organizational processes.[1] ECC applications can automate repetitive, formulaic tasks and, in doing so, deliver orders-of-magnitude improvements in the speed of information analysis and in the reliability and accuracy of outputs. For example, ECC call center applications can answer customer calls within 5 seconds on a 24-7-365 basis, accurately address their issues on the first call 90% of the time, and transfer complex issues to employees, with less than half of the customers knowing that they are interacting with a machine.[2] The power of ECC applications stems from their ability to reduce search time and process more data to inform decisions. That's how they enhance productivity and free employees to perform higher-level work — specifically, work that requires human adaptability and creativity. Ultimately, ECC applications can enhance operational excellence, customer satisfaction, and employee experience.[3]



JOHN HERSEY/THEISPOT.COM

THE LEADING QUESTION
How can companies develop their ability to use AI to transform business operations?

- **CEOs recognize the potential of AI to improve operations and capture value but are struggling to realize its promise.**
- **Business domain proficiency — a deep understanding of the tasks, workflows, and logic of existing processes — establishes the essential link between data science and business value.**
- **ECC applications must be managed throughout their life cycles because ever-changing conditions ensure that AI algorithms become a less accurate reflection of reality over time.**

ECC applications come in many flavors. For instance, in addition to call center applications, they include banking applications for processing loan requests and identifying potential fraud, legal applications for identifying relevant case precedents, investment applications for developing buy/sell predictions and recommendations, manufacturing applications for scheduling equipment maintenance, and pharmaceutical R&D applications for predicting the success of drugs under development.

Not surprisingly, most business and technology leaders are optimistic about ECC's value-creating potential. In a 2017 survey of 3,000 senior executives across industries, company sizes, and countries, 63% said that ECC applications would have a large effect on their organization's offerings within five years.[4] However, the actual rate of adoption is low, and benefits have proved elusive for most organizations. In 2017, when we conducted our own survey of senior executives at 106 companies, half of the respondents reported that their company had no ECC applications in place. Moreover, only half of the respondents whose companies had applications believed they had produced measurable business outcomes. Other studies report similar results.[5]

This suggests that generating value from ECC applications is not easy — and that reality has caught many business leaders off guard. Indeed, we found that some of the excitement around ECC resulted from unrealistic expectations about the powers of "intelligent machines." In addition, we observed that many companies that hoped to benefit from ECC but failed to do so had not developed the necessary organizational capabilities. To help address that problem, we undertook a program of research aimed at identifying the foundations of ECC competence. We found five capabilities and four practices that companies need to splice the ECC gene into their organization's DNA.

Five Crucial Capabilities

We found that companies that successfully create value (that is, radically improve business processes to reduce costs and/or generate new revenues) using ECC applications possess five capabilities: data science competence, business domain proficiency, enterprise architecture expertise, an operational IT backbone, and digital inquisitiveness.

Data science competence. Data science competence encompasses a wide range of skills essential to ECC. It involves ensuring the availability and usefulness of massive amounts of data: collecting, cleaning, curating, tagging, and analyzing internal and external data from multiple sources. Such competence also entails identifying and describing relationships between data, as well as developing AI algorithms that have learned from data how to identify patterns and probabilities.

Top-notch data scientists have extensive knowledge in areas such as natural language processing, statistical inference, knowledge representation, and learning algorithms. Wipro, the Indian IT services company, includes these areas among the pillars of its data science expertise. Its data scientists deploy their skills and a variety of tools to create AI algorithms that can be inserted into enterprise applications.

For organizations that cannot develop the talent internally, obtaining data science competence is expensive and can require multiple hires from, for example, software development companies, technology consulting companies, AI startups, or university graduate programs in related fields. At a financial services company we studied — we call it OneBankAssure — the CEO hired a new direct report who was a technically accomplished data science academic and consultant. This person, in turn, hired the 20 data scientists who became the core ECC development team. Companies that are serious about ECC spend the money to hire the right data science talent. To raise the money, one pharmaceutical company we studied reduced its operational IT costs (by eliminating duplication in systems and standardizing processes across its business units) and redirected the savings to the acquisition of data science skills.

Business domain proficiency. Domain proficiency is needed to understand the tasks, workflows, and logic of existing business processes, as well as to imagine how ECC applications could improve them. As many organizations have learned the hard way, it's possible — even easy — to develop an elegant AI algorithm that uses massive amounts of data to learn how to predict or categorize something but doesn't improve the business. Having the right technical skills isn't enough. Domain proficiency links data science competence to business value.

ABOUT THE RESEARCH

The research activities on which this article is based were undertaken between January 2016 and December 2017, and covered companies across industries in North America, Europe, Asia, and Australia. We interviewed senior executives in IT and innovation units in 33 companies, as well as industry and technical experts in eight enterprise cognitive computing developer/vendor organizations, regarding ECC uptake in a range of organizations and industries, and ECC challenges and opportunities. We studied 51 ECC use cases (37% deployed; 48% in ideation or design stages; and 15% abandoned prior to development). We surveyed senior IT and technology leaders in 106 companies about ECC applications in place, application development and management issues, and outcomes. Finally, we researched and prepared three in-depth case studies on three organizations, for which we interviewed 35 people: C-level officers; functional leaders in IT, marketing, sales, and strategy; and data science and domain/process experts.[i]

For example, the ability of data scientists to effectively curate, tag, and analyze data depends on a clear understanding of the relationships among the data from a process and business point of view. Domain proficiency provides clarity around those relationships, referred to as ontologies. Data ontologies can become quite complex and even counterintuitive. Here's how a domain expert at a pharmaceutical company described some of the complexities he encountered capturing the data ontologies needed to support the company's research on diabetes: "A big part of diabetes is being overweight. Should there be an obesity dimension in our ontology of diabetes? Or is diabetes an attribute of obesity? Oh, and people who are overweight often have joint replacement issues. If they're overweight and their joints hurt and they have diabetes, the incidence of depression is very high, and dealing with depression is an important part of generating outcomes. Do I train the algorithm on depression?"

Domain proficiency is also important for creating the business rules that shape how the outputs from the algorithm are handled by the ECC application. For example, an ECC application that helps banks predict which customers are most likely to repay loans on time must include business rules for how the algorithm's prediction will be applied, such as: Will some loans be granted automatically? If so, under what conditions? With whom will the predictions be shared? Under what circumstances can a prediction be overridden?

For any given ECC application, domain proficiency is needed in all the functional areas that have a bearing on — or are stakeholders in — the operations of the focal process. For example, a team at a U.S. bank that developed an ECC application to detect financial fraud needed proficiency not only in fraud identification and prevention but also in the related areas of regulatory compliance and banking law.

People with domain proficiency have deep process knowledge. They may be process owners, although they are often people with a regular hands-on role. Some companies seek to hire data scientists with domain expertise. Indeed, such individuals can partner well with business domain experts, but they cannot substitute for them when an ECC application is being developed. That's because they usually lack enterprise-specific knowledge about processes, policies, and practices currently in play.

Enterprise architecture expertise. Implementations of enterprise systems have a history of disappointing leaders who underestimated the organizational changes needed to capture their value. Too many leaders are reliving this disappointment with ECC applications. ECC applications do not deliver value by simply processing data and delivering outputs. They deliver value when the organization changes its behavior — that is, when it changes processes, policies, and practices — to gain and apply the insights from those outputs. Experts in enterprise architecture design the new organization needed to create business value from ECC applications, and they help manage the transition from the old organization to the new one.

The most ambitious ECC applications usually affect several, often fundamentally different business processes. In such cases, enterprise architects are needed to orchestrate the redesign of the systems, processes, and roles across organizational units. The more ambitious the ECC application, the more likely it will require far-reaching organizational changes.

Organization design and change issues can surface for seemingly small-scale applications. One medical drug distributor failed to recoup its investment in an ECC application that could accurately predict whether an online customer's insurance would cover a claim 90% of the time because the accounts payable department balked at making costly process changes required to support the application. If an enterprise architect had been engaged in the project from the outset, this loss might have been avoided.

The organizational changes needed to unlock the potential of an ECC application can be complex and intertwined. Enterprise architects are familiar with the organizational roadblocks that drive up costs or limit impact. At Wipro, enterprise architecture expertise helped smooth the way for a new help desk ECC application by first merging the company's existing help desk applications, reducing the types of fault tickets from 3,000 to 2,200, and eliminating redundancies in support tasks. By simplifying and standardizing the help desk process prior to the development of the ECC application, the company reduced and simplified the work of getting the data needed to train the AI algorithm, developing it, and ultimately automating the process, thus unlocking additional value.

Enterprise architects also recognize when ECC applications require changes in employees' jobs. They may see the need for upskilling, re-skilling, or the creation of entirely new roles. When a seemingly simple sales-lead-generating ECC application required its agents to do more cold calling and make more targeted pitches, OneBankAssure's enterprise architects designed a new coaching role to help agents that proved essential to generating benefits from the application.

Given the breadth of skills that enterprise architects draw on, this expertise can be difficult to develop. It often resides in people who are steeped in organization design and change management such as business leaders with experience managing technology-driven transformations or other reorganizations. Human resource professionals with exposure to a broad range of organizational roles can be a good source of architectural expertise in role design and redesign, as well as skills training. IT professionals with exposure to many different business processes, who can help streamline processes and establish the proper division of work between ECC applications and employees, also can be tapped.

Operational IT backbone. A company's existing technology and data foundation — its operational IT backbone — and the people responsible for it support the development and running of ECC applications. They supply the IT capabilities needed to store and access critical data, integrate ECC applications with other applications, provide reliable operations, and ensure privacy and security.

As noted earlier, for an AI algorithm to learn from data, a company must make available massive amounts of high-quality data that is cleaned and tagged. The lack of high-quality data is the most pernicious and least anticipated obstacle in the development of AI algorithms. OneBankAssure overcame this obstacle and accelerated its adoption of ECC by separating responsibilities for developing AI algorithms from responsibilities for providing the data. Since the IT unit already maintained the underlying operational infrastructure and good-quality operational data, it was able to support the algorithm developers by providing them with access to a data lake containing operational and external data. Responsibility for structuring the data for developing algorithms rested with the data scientists.

Almost no new enterprise application can operate in isolation from other enterprise applications. ECC is no exception. If an application is not

“Enterprise architects recognize when ECC applications require changes in employees' jobs. They may see the need for upskilling, re-skilling, or the creation of entirely new roles.

> Often, users of ECC applications must apply human judgment to predictions made by algorithms. They need to possess digital inquisitiveness — the inclination to question and evaluate the data before them.

properly integrated, it will be hard to use and possibly ignored. That's why the IT unit at OneBankAssure embedded the company's new sales lead system into its customer relationship manager system, which was part of its operational IT backbone. The CRM linked up-to-date contact information and customer history data to the sales leads. It also provided a set of processes within which the ECC sales leads could be seamlessly presented to users. Being part of the IT backbone also meant that the sales lead system would be scalable, reliable, and secure.

The existing IT staff is the logical source for operational IT backbone expertise. At OneBankAssure, the IT function set the standards for ECC plug-ins and adapted applications to the company's production environment by refactoring and retesting the code. It also managed disaster recovery and security for installed ECC applications.

Digital inquisitiveness. The AI algorithms in ECC applications do not produce definitive answers. Rather, they produce predictions based on probabilities: the probability that a customer will buy a product, that a patient has a disease, that a loan will be repaid. Often, application users must consider these predictions and apply human judgment to arrive at decisions about how and where to promote offerings, what treatments to prescribe, or what loans to approve. To do this effectively, they need to possess digital inquisitiveness — a habitual inclination to question and evaluate the data before them. They must use that skill to better understand the options provided by ECC applications and continually improve outcomes.

The development of this capability requires a broad-based effort. A number of companies we studied instituted mechanisms to cultivate digital inquisitiveness. OneBankAssure's corporate university delivered a training program that introduced executives to the idea of using data effectively in making decisions. One exercise incorporated a strategy game, in which participants vied to develop the highest-value ECC solution to a business problem. At different phases of the game, they had to deal with poor-quality data (in fact, real company data), build decision trees, teach an algorithm to detect patterns, and develop a model to solve a problem. Wipro created an e-learning platform on which employees were able to take courses to understand what AI was, how it could be used in business processes, and how to work effectively in ECC-enabled processes. The company also trained hundreds of domain experts to act as AI champions throughout the organization.

Four Key Practices

Developing the five capabilities equips organizations to derive value from ECC applications, but then companies must apply those capabilities. We've found that four practices in particular help them do that, creating the conditions for a given application — and its underlying AI algorithm — to deliver on its promise.

Develop clear, realistic use cases. A use case provides a clear definition of what an ECC application will do and illustrates how its AI algorithms will enhance the execution and outcomes of a business process or set of processes. It shows how work will be divided between an application and a user. In doing so, a use case establishes the need for process changes and provides initial insights into any new capabilities users will need (as well as any skills that will no longer be needed). A well-designed use case also facilitates the estimation of the costs and benefits of the ECC application.

Consider an ECC application in a call center: Its use case might include a simple version of an AI algorithm that matches customer queries to resolutions. It would show what the algorithm would do and what automated resolutions the

> Because of their business focus, domain experts are often more persuasive ECC champions than are data scientists and IT professionals, who may be perceived as overly enamored with AI.

application could provide to customers. It would also show that some queries would be passed along to call center representatives. Required work and capability changes for call center employees could be inferred as well. All that information would allow a domain expert to roughly gauge the challenges of adoption and estimate intended benefits in terms of reduced response time, reduced labor, fewer follow-up calls, greater customer satisfaction, or a combination of outcomes.

Developing an ECC use case that is grounded in reality is a team activity. It is primarily the responsibility of domain experts and data scientists, who specify how an AI algorithm will enhance organizational outcomes and what data is needed to create it. But enterprise architects weigh in, too, identifying any new structures, roles, and systems required by a proposed ECC application, especially those affected indirectly by the new application. IT experts assess the need for integration with other applications and identify any additional IT support the application might require.

Properly developed use cases can help companies avoid sloppy or ill-considered ECC implementations that waste resources and may limit enthusiasm for — and effective implementation of — ECC. In fact, if the use cases for early ECC applications highlight quick wins for high-profile issues, they can be a powerful driver of organizational uptake of ECC. Bench scientists at one pharmaceutical company suggested developing an ECC application that could mine patent data for a specific disease knowing that if it was successful, the application itself would serve as a use case for similar applications for other diseases — and it did. Sometimes the algorithms themselves can be substantially reused. Wipro developed a use case for new-customer verification in the financial services sector, in which the AI algorithm automated the extraction and interpretation of information from customers' financial documents. This use case gave rise to an ECC application in the engineering sector that extracted and interpreted information from digitized blueprints.

Manage ECC application learning. AI algorithms in products such as smartphones use the data they process to improve themselves without human intervention. In contrast, ECC applications have a much more complex feedback loop. Business conditions and demands change constantly. As a result, the data used to create an AI algorithm becomes a less accurate reflection of reality over time — the algorithm *drifts*. It thus becomes necessary to manage the learning of the ECC applications throughout their life cycles.

Algorithm drift may occur quickly, as in predicting the sales of fashion apparel, or slowly, as in predicting the presence of a disease. To manage drift and keep ECC applications up-to-date, companies usually rely on a combination of IT backbone capabilities, data science competence, and domain proficiency. They build reporting mechanisms into ECC applications that generate alerts if the business results derived from the application's outputs are no longer aligned with the organization's goals, the algorithm's recommendations aren't within preestablished error ranges, or the application isn't running properly.

When deviations occur, AI algorithms need to be retrained and ECC applications relaunched. Domain experts and data scientists need to work together to identify, access, clean, tag, and architect new sources of data to improve the accuracy of AI algorithms and the utility of ECC applications. In addition, as the performance of the algorithm is better understood or as users become more proficient with the application, new business rules or processes that can enhance the value of the application may be required.

At OneBankAssure, domain experts and data scientists identified new external sources of data that could help identify productive sales leads, so they retrained their AI algorithm. They also learned that agent experience affected sales success, so they developed more elaborate rules to govern how the ECC application presented leads. The new data and business rules led to a richer, more complex ECC application that OneBankAssure continues to enhance.

Cocreate throughout the application life cycle. A data scientist or business domain expert cannot develop and sustain an ECC application in isolation. Interviewees in companies that effectively exploited AI repeatedly told us that they had, at first, badly underestimated the intense level of interdisciplinary cocreation needed to achieve success with ECC. They said they began to make progress only when they realized that ECC applications require people from disparate specialties and disciplines to work as a single team, not just during initial development and implementation but also in ongoing development throughout the application life cycle.

One reason cocreation is important for ECC is because business experts do not yet understand what AI can and can't do. During the development of ECC, close and sustained collaborative relationships across diverse areas of expertise can ameliorate this problem. At OneBankAssure, as a matter of hard-learned policy, every ECC application is created by a team of process owners and users with domain expertise, enterprise architects, and data scientists, with added assistance from the IT function. There are few handoffs within the team. No team member ever works completely alone, and in the end, no one team member is responsible for success or failure. The interaction of team members results in a shared vocabulary about the business need and potential solutions, enabling them to better visualize and make sense of how people will actually use the application.

During implementation, owners of the IT operational backbone get involved not with a single handoff but rather by working with the ECC application team to cocreate a solution for integrating the application with the backbone upon production. After implementation, responsibility for maintaining and sustaining ECC applications continues to be highly interdependent in nature, as described above.

Think "cognitive." Companies that successfully develop and use ECC applications champion the uptake of AI and create positive buzz and excitement around its use. They encourage employees to generate ideas for new ECC applications that can improve their own work.

The employee response to ECC varies widely. Some people do not see the potential of ECC at first. Others have exaggerated expectations, thinking that ECC applications will automatically solve difficult business problems. Still others do not trust AI and see risks to ECC-enabled business processes, such as rogue behavior in AI algorithms and capability or job losses.

Domain experts who have seen what AI can do are the best stewards of realistic and credible conversations about ECC within their companies. Because of their business focus, they are more likely to be able to create a positive buzz around ECC than are data scientists and IT professionals, who

LEARNING THROUGH DOING

As companies apply their enterprise cognitive computing capabilities through the four key practices, they're also enriching their capabilities. Practices are, after all, opportunities to practice.

The pharmaceutical company we studied offers a good example. Recognizing that data science and ECC applications would become increasingly important to curing and preventing disease, the company hired data scientists to conduct workshops that would help senior staff (mainly business domain experts and enterprise architects) imagine the possibilities. They worked with business leaders to identify information-processing bottlenecks that created backlogs in drug discovery, clinical trials, manufacturing, and commercialization. The bottlenecks highlight opportunities for AI applications that could solve problems for small groups of analysts and decision makers in the organization.

These early efforts generated incremental business value, but the business leaders were far more focused on building capabilities than on building game-changing applications. They carefully chose use cases to meet the needs of people who naturally think "cognitive" and then engaged all the needed expertise — data scientists, domain experts, and IT specialists — to cocreate and manage the applications. In those pockets of the company, people deepened their understanding of organizational impacts and developed the capabilities to identify and pursue more ambitious ECC applications. What's more, the gains they made in efficiency and productivity inspired others in the company to seek out their own use cases and build their own capabilities. Creating this virtuous cycle of continuous organizational learning has mitigated the risks of the company's AI investments and positioned the company to make ECC a competitive advantage.

may be perceived as overly enamored with AI. Indeed, at Wipro, domain experts were enlisted as AI champions — conducting "walkabouts" in their various departments, evangelizing ECC, and listening to ideas put forth by their colleagues.

The most likely sources of ideas for new ECC applications are people with domain proficiency or data science competence (or both). At OneBankAssure, operational managers spent several months in discussions with data science professionals to envision how their business might be affected by AI in the future, to develop ideas for new ECC applications, and to draft road maps for how their ideas could be developed and commercialized.

Proactive data science leaders also can be effective idea generators. At a pharmaceutical company we studied, one ECC project got its start at a lunch in which a business leader told a data scientist about a business problem, and the data scientist proposed a simple solution leveraging an already developed AI algorithm. In another company, the head of the data science unit organized seminars for functional and business leaders to identify areas in which ECC applications could best serve them.

The digital inquisitiveness of the entire workforce should be harnessed, too. Wipro, for example, crowdsources ideas from employees. It encourages them to envision and suggest new ECC applications, evaluating the ideas for their potential contribution to top-line growth, bottom-line profits, customer satisfaction, or employee satisfaction.

BUSINESS APPLICATIONS OF AI may not create the same buzz as a self-driving car, but they can generate handsome returns — dramatic improvements in performance, profitability, revenues, and customer satisfaction. By cultivating the five capabilities and applying the four practices described in this article, business leaders can splice the ECC gene into their organizational DNA and set themselves up to reap those rewards.

It's a virtuous cycle: The capabilities enable employees to execute the practices, and the practices themselves exercise and strengthen the capabilities. This cycle helps companies become evermore adept at developing and using ECC applications that improve operations and create business value.

***Monideepa Tarafdar** is a professor of information systems and codirector of the Centre for Technological Futures at Lancaster University Management School in the United Kingdom. Her research focuses on how digital technologies transform organizations and societies. **Cynthia M. Beath** is professor emerita at the University of Texas McCombs School of Business in Austin. Her research focuses on how organizations get value from investments in IT. **Jeanne W. Ross** is a principal research scientist at MIT Sloan's Center for Information Systems Research in Cambridge, Massachusetts. She writes a quarterly column for* MIT SMR *about digital business management issues. Comment on this article at http://sloanreview.mit.edu/x/60402.*

REFERENCES

1. ECC applications are distinct from other kinds of enterprise software in that AI tools, rather than human deduction, are used to figure out what logic will optimize business outcomes. AI software tools apply computational and analytical techniques, such as neural network analysis, machine learning, and Bayesian statistics, to large sets of structured and unstructured data to create AI algorithms that will classify, cluster, predict, and match patterns. These algorithms become part of the logic of the ECC application.

2. J. Bughin and E. Hazan, "Five Management Strategies for Getting the Most From AI," MIT Sloan Management Review, Sept. 19, 2017, www.sloanreview.mit.edu.

3. M. Tarafdar, C.M. Beath, and J.W. Ross, "Enterprise Cognitive Computing Applications: Opportunities and Challenges," IEEE IT Professional 19, no. 4 (August 2017): 21-27.

4. S. Ransbotham, D. Kiron, P. Gerbert, et al., "Reshaping Business With Artificial Intelligence," MIT Sloan Management Review research report, Sept. 6, 2017.

5. S. Norton, "Machine Learning at Scale Remains Elusive for Many Firms," The Wall Street Journal, April 27, 2018; J. Bughin and E. Hazan, "Five Management Strategies"; and Ransbotham, et al., "Reshaping Business."

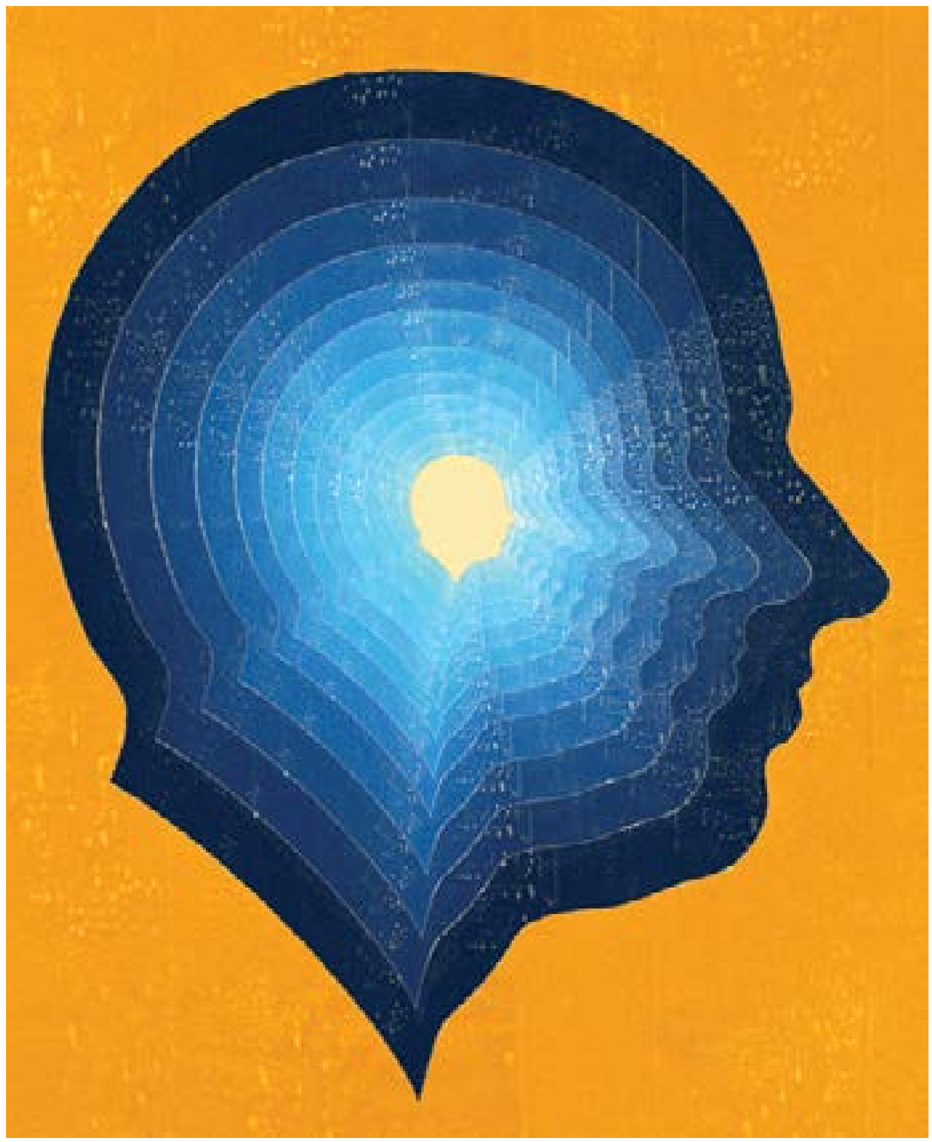
i. C.M. Beath, M. Tarafdar, and J.W. Ross, "OneBankAssure: Customer Intimacy Through Machine Learning," working paper, MIT Center for Information Systems Research, Cambridge, MA, March 12, 2018; M. Tarafdar and C.M. Beath, "Wipro Limited: Developing a Cognitive DNA," working paper, MIT Center for Information Systems Research, Cambridge, MA, April 27, 2018; and J.W. Ross, K. Moloney, and C.M. Beath, "Pharmco: Becoming a Data-Science Driven Company," working paper, MIT Center for Information Systems Research, Cambridge, MA, Feb. 21, 2019.

Reprint 60402.

AI Can Help Us Live More Deliberately

We need a little friction in our lives to trigger reflection, self-awareness, and responsible behavior.

BY JULIAN FRIEDLAND



As I search online for a present for my mother, considering the throw pillows with sewn-in sayings, plush bathrobes, and other options, and eventually narrowing in on one choice over the others, who exactly has done the deciding? Me? Or the algorithm designed to provide me with the most "thoughtful" options based on a wealth of data I could never process myself? And if Mom ends up hating the embroidered floral weekender bag I end up "choosing," is it my fault? It's becoming increasingly difficult to tell, because letting AI think for us saves us the trouble of doing it ourselves and owning the consequences.

AI is an immensely powerful tool that can help us live and work better by summoning vast amounts of information. It spares us from having to undergo many mundane, time-consuming, nerve-wracking annoyances. The problem is that such annoyances also play a key adaptive function: They help us learn to adjust our conduct in relation to one another and the world around us. Engaging directly with a grocery bagger, for instance, forces us to confront his or her humanity, and the interaction (ideally) reminds us not to get testy just because the line isn't moving as quickly as we'd like. Through the give-and-take of such encounters, we learn to temper our impulses by exercising compassion and self-control. Our interactions serve as a constantly evolving moral-checking mechanism.

Similarly, our interactions within the wider world of physical objects forces us to adapt to new environments. Walking, bicycling, or driving in a crowded city teaches us how to compensate for unforeseen obstacles such as varying road and weather conditions. On countless occasions every day, each of us seeks

DAN BEJAR/THEISPOT.COM

THE LEADING QUESTION
How do we design machines smart enough to keep us from becoming like machines ourselves?

- As we "outsource" myriad tasks to AI-assisted platforms, we may become less reflective and feel less responsible for outcomes.
- Moral self-awareness is a powerful motivating force that can help restore critical self-reflection, agency, and a sense of accountability.
- AI developers can incorporate prompts that promote moral self-awareness in areas ranging from health and well-being to media and civic engagement.

out an optimal compromise between shaping ourselves to fit the world and shaping the world to fit ourselves.[1] This kind of adaptation has led us to become self-reflective, capable of ethical considerations and aspirations.

Our rapidly increasing reliance on AI takes such interactions out of our days. The frictionless communication AI tends to propagate may increase cognitive and emotional distance, thereby letting our adaptive resilience slacken and our ethical virtues atrophy from disuse.[2] Relying on AI to preselect gifts for friends and family, for example, spares us the emotional labor of considering their needs and wants in our ordinary interactions with them to select a genuinely thoughtful gift. Many trends already well underway involve the offloading of cognitive, emotional, and ethical labor to AI software in myriad social, civil, personal, and professional contexts.[3] Gradually, we may lose the inclination and capacity to engage in critically reflective thought, making us more cognitively and emotionally vulnerable and thus more anxious[4] and prone to manipulation from false news, deceptive advertising, and political rhetoric.

In this article, I consider the overarching features of this problem and provide a framework to help AI designers tackle it through system enhancements in smartphones and other products and services in the burgeoning internet of things (IoT) marketplace. The framework is informed by two ideas: psychologist Daniel Kahneman's cognitive dual process theory[5] and moral self-awareness theory, a four-level model of moral identity that I developed with Benjamin M. Cole, a professor at Fordham University's Gabelli School of Business.[6] (See "Theories of Mind in an AI World.")

When Convenience Leads To Disengagement

The most immediately attractive feature of AI technology is its promise to handle the mundane aspects of life, thereby increasing the amount of time and attention each of us can devote to activities we consider more rewarding. Of course, every time this kind of outsourcing occurs, we cede a degree of control. Getting comfortable with these trade-offs reinforces new habitual behaviors that entail a measure of disengagement: from one another, the physical world, and even ourselves. This is because every time we delegate a degree of control to the AI system, we also invest a degree of trust into that system. In so doing, we will often shift from relying on what Kahneman calls our reflective mind (and its deliberative decision-making) to our autonomous mind (and its automatic reactions that guide decisions). This makes it easy to complete a routine task. But repeating this process creates a risk that our actions become increasingly automatic and less reflective overall, leading to six forms of disengagement:[7]

1. Increased passivity. As we accept assistance to complete a task, we require less effort to carry it out. We may become spectators rather than active participants. The AI systems that Netflix, Amazon Prime, and Facebook use to preselect entertainment and news options are examples. When we let these systems determine our options, we rarely confront perspectives that might challenge our preconceptions and biases. Gradually, we may become less prepared to expend the effort needed to think deeply and critically, thereby disengaging long-term memory.[8]

2. Emotional detachment. Diminished participation leads to emotional disengagement. Consequently, our actions can become insincere or deceptive. Think of a customer call center, where an AI system in a help desk or sales context aggressively coaches agents in real time as they respond to customers' emotional cues.[9] Such software, ideally designed to train operators to become more sensitive to customers' concerns, could have the reverse effect, making us increasingly inured to emotional cues because we will have less practice picking up these cues ourselves and have less interest in doing so.

3. Decreased agency. Disengagement reduces our power to make our own decisions by lessening our awareness of actions we might take. Consider an automated vehicle preprogrammed to weigh competing ethical priorities during a crash, such as whether to hit a pedestrian or another vehicle. Auto insurance rates might be adjusted according to the degree to which we set the automated driving system to integrate others' interests into the calculus.[10] And we would relinquish the agency to make our own choice as the crash takes place.

4. Decreased responsibility. In ceding control over a decision-making process, we can become less accountable for results — whether they are good or bad — because responsibility is diffused across the entire AI-based system, from design to delivery. Imagine a dieting app that orders prepared foods to be delivered to you according to a weight-loss plan set up by AI. If you lose weight, who deserves the credit? And if you don't, whose fault is it?[11]

5. Increased ignorance. AI translates our wants into algorithmic shorthand or mechanical processes that may end up functioning differently than we would ourselves. Of course, that can make up for deficiencies in our knowledge — but it can also reinforce those deficiencies. Virtual navigational apps like those offered by Waze, Garmin, and others do not require you to acknowledge your surroundings. You might, for instance, keep circling an incorrect location that the mapping app has not yet updated, out of preferential bias for the AI system, instead of returning back to your own direct perceptions and judgments.[12] At your intended destination, you might have no idea what route you took to get there nor how to get back to where you started without AI assistance.

6. De-skilling. Depending on an intermediary for completing routine tasks can dull many of the trained skills we rely on to interact with the physical world around us. We may forget how to perform basic tasks or become less proficient at doing them unaided. Using only navigation apps lulls us into forgetting how to use a conventional map or, in a future era of autonomous vehicles, even how to drive without the apps. We may also lose motivation to acquire new skills, opting instead for evermore outsourcing solutions.

Together, these trends present an ethical challenge: Because they multiply the instances in which we go through life while operating on autopilot, they have the potential to loosen our social bonds, exacerbate conflicts, and hamper moral progress by stifling self-critical thought. To mitigate these threats, designers of AI systems should build in features and interfaces that periodically re-trigger our reflective minds.

It Takes More Than "Nudges" to Make Us Think

In their influential book, *Nudge*, behavioral economist Richard Thaler and legal scholar Cass Sunstein have argued that cognitive nudges can spur us to action by using triggers that evoke emotions like empathy or self-interest.[13] Unfortunately, such nudges have limited power in practice because they prompt only behavioral impulses and do not engage critical reflection. This is the case even when pressing health risks are concerned. In a study of 1,509 patients who had heart attacks, efforts to prompt people to adhere to medication prescriptions (including electronic pill bottles and the chance for $5 or $50 rewards for enlisting the support of a friend or family member) did not significantly improve the likelihood that people would take their medicine.[14]

Triggering the reflective mind is more likely to solve the problem of disengagement and mitigate the risks of losing skills in the age of AI. By creating what we can call cognitive speed bumps that force us to reflect on decisions worthy of greater reflection, developers of AI systems can reintroduce *interactive friction* into the experiences they host. So as Mom's birthday approaches, instead of suggesting purchases, our AI system might instead suggest a good time to call or pay Mom a visit — an opportunity to enhance the personal relationship and even help come up with a thoughtful (and desired) gift.

THEORIES OF MIND IN AN AI WORLD

Cognitive dual process theory describes two overarching decision-making processes: (1) the *autonomous mind*, which automatically reacts to stimuli, and (2) the *reflective mind*, which responds consciously in a deliberate and reasoned fashion.[i]

Most AI-assisted platforms function to free up the attention of the conscious reflective mind for any activities that immediately suit a person's interests or grab his or her attention. Ideally, each new outsourced task is accomplished more effectively than via direct unassisted interaction. Thus, AI allows us to conveniently increase the levels at which we may productively process incoming information from the external physical and social worlds.

AI systems typically guide users with visceral notices, which researchers have divided into three general categories:[ii]

- **Familiarity notices** use familiarity with one technology to inform users about another. Example: camera-clicking sounds and dial tones on smartphones.
- **Psychological reaction notices** use common psychological reactions to shape a consumer's conception of the product or service. Example: casual interface designs like friendly avatars that signal greater honesty and openness.
- **Showing notices** promote self-awareness by showing users the results of their activities. Example: screen-time data embedded in the iPhone iOS 12.

Familiarity notices and psychological reaction notices are designed to trigger only the autonomous mind, but showing notices introduce communicative friction designed to trigger the reflective mind. Screen-time software embedded in the iPhone operating system shows people how often they use social networking, entertainment, and productivity apps. This allows them to better understand and take control of their own behavior.

The ramifications are profound. Perhaps the most seductive aspect of AI-assisted platforms is that they promote what technology ethicist Shannon Vallor describes as "frictionless interactions that deftly evade the boredom, awkwardness, conflict, fear, misunderstanding, exasperation, and uncomfortable intimacies that often arise from traditional communications, especially face-to-face encounters in physical space."[15] Here, Vallor is referring mainly to the avoidance of live conversations, through social media. But she may as well be talking about evasion of all the practical drudgeries of life, from reading a map, driving a car, and minding one's surroundings to making a grocery list, shopping, and cooking. And though most of us still have such frictional experiences, AI-assisted platforms promise to guide our attention in whatever directions we are likely to find most immediately satisfying, thereby reducing the chances that we will have to experience *unpleasant* friction. As a result, our moral attention — the ability to redirect our focus, delay gratification, temper our emotional urges, and restrain our unthinking reactions — erodes.

We need something to counteract this tendency: an AI choice architecture designed to preserve healthy measures of interactive friction between ourselves and the wider world.

How Friction Fosters Moral Self-Awareness

There is value to a world of friction-filled interactions. For instance, new research on childhood self-control suggests that one's cultural[16] and socioeconomic[17] environments may play a far greater role than genetic factors in developing grit and perseverance, which are highly correlated with professional success later in life.[18] It is only by learning how to navigate interactions that are not set up for our comfort that we are able to fully develop executive control over our own consciousness.[19]

Such interactions also foster moral self-awareness. As we experience friction again and again, the ways we react to various stimuli change, and moral identity evolves: We begin to think and feel differently about what our actions say about ourselves.[20]

The social psychological literature has established a clear relationship between what's called the self-importance of moral identity and moral thought and action,[21] and the wider literature on civic-mindedness indicates that pride is the most effective moral motivator of civic behavior.[22] There is also evidence that ethical consumers are happier and have stronger repurchase intentions when motivated by their moral self-image than when motivated by emotions such as guilt and empathy.[23]

What does all this have to do with AI? Designers of AI systems can use the four levels of moral self-awareness described below as a guide for developing applications that encourage reflective behavior. By incorporating triggers for interactive friction, they can prompt users to consider how their actions reflect their personal values and help them ascend to higher levels of awareness.

LEVEL 1: Social reflection. At this level, people rely chiefly upon negative feedback they receive from observers to guilt or shame them into changing their behavior. Researchers have demonstrated the power of negative feedback to inhibit a person's selfish behavior. For example, participants primed in a tragedy of the commons experiment to be self-interested gradually learned to temper their self-interest after being shamed by other subjects left with fewer resources.[24] Eventually, all subjects showed a preference for lowered individual returns in favor of equitable and sustainable longer-term outcomes.

LEVEL 2: Self-reflection. At this level, rather than relying on others' complaints to acknowledge the negative impacts of their actions, actors start to serve as their own source of feedback. This happens when they see the outcomes of others' behavior or

> It is only by learning how to navigate friction-filled interactions that we are able to fully develop executive control over our own consciousness.

> By providing ‘showing notices’ — snapshots of users’ behavior — AI applications can encourage people to move toward the highest level of moral self-awareness, where positive aspirations drive individual behavior.

when they consider the immediate ramifications of their own actions. For example, a person who notices a room containing swept litter is 2.5 times less likely to toss trash on the floor than in a litter-strewn room.[25] Observing the neatened-up litter increases the observer’s propensity to keep the room clean.

LEVEL 3: Anticipatory self-reflection. At this level, people start to anticipate potential negative consequences of their actions and do so independently from others’ signals. This behavior often comes after self-reflection on prior behavior has led to an internal sense of guilt or shame. At a crucial turning point in the tragedy of the commons experiment mentioned above,[26] one participant asked aloud, “Are we bad people?” This question was not so much an effort to shame other group members as an attempt to reconcile the inconsistency between one’s prior action (to serve self-interest) and one’s aspirational moral self-image. Such a reflective moment represents a crucial step, one that reveals the moral obligations of individuals to shape themselves to fit the world and their own aspirations within it.

LEVEL 4: Proactive self-reflection. At the highest level, people become increasingly forward-looking, considering both negative and positive impacts. They purposely engage in appropriate actions to realize positive outcomes. They internalize the self-image of potential hero rather than potential villain.[27] At best, these decisions are habit-forming, bringing people closer to becoming whom they aspire to be. This state of mind is linked with achieving greater happiness based on an individual’s self-conception.[28]

Triggering the Reflective Mind

In traditional face-to-face interactions, the external physical or social world provides the friction necessary to trigger the reflective mind into modifying one’s behavior for the better. As AI removes opportunities for those interactions, developers need a tool for tapping into users’ moral self-awareness. *Showing notices*, a type of visceral notice that AI systems can incorporate to shape users’ decision-making, can serve as that tool and compensate for the loss of give-and-take interactions in the social and physical world. (See “Theories of Mind in an AI World,” p. 47, for more detail about visceral notices.)

Showing notices provide users with snapshots of their behavior (the number of steps taken in a day, for example, or the amount of time spent online). They can enhance AI applications by encouraging users to move from the first, second, and third levels of moral self-awareness, in which negative feelings like guilt and shame primarily drive individual behaviors, toward level 4, in which positive aspirations encourage people to act, conscious that their choices can make a difference for themselves and society. Enabling users to share their progress on a given issue with others in a social group further enhances an app’s potential.

Considering that by current projections, global IoT spending could reach $1.4 trillion by 2021, such functionality presents rich opportunities for research and development.[29] Five lifestyle categories in particular have significant potential for this type of innovation: health and well-being, social responsibility, media and civic engagement, skill maintenance, and personal edification. We’ll consider each one here.

Health and well-being. There is already significant movement in providing showing notices in health and wellness apps — from those that facilitate personal fitness, mindfulness, or sleep management to those that allow us to set screen-time limits on our cellphones. Smart refrigerators are another frontier. For example, adding showing notices that illustrate patterns of consumption of highly processed, high-sugar, canned, frozen, and fresh food, along with daily calorie consumption data, could help users improve

> Media-quality applications could alert people to misleading or biased news sources. New tools could gradually introduce alternate points of view, encouraging users to break out of ideological echo chambers.

nutrition. Combined with data from grocery delivery services, such notices could guide users to order groceries according to healthier recipes and locally or sustainably sourced foods.

Social responsibility. Another area with potential is in helping people make thoughtful brand and investment choices that align with their social values. A few apps now highlight possible ethical concerns in financial portfolios, flagging sectors that users may wish to avoid in light of stated preferences (such as alcohol, petroleum, and tobacco) and providing finer-grained notices about any ethical quandaries companies may be involved in. Smart refrigerators could provide notices about the carbon footprint of groceries purchased (where consumers have access to carbon labels). Such notices could extend to other areas, alerting users to factors such as air and water pollution, resource depletion, and green packaging.

Media and civic engagement. Media-quality applications could use showing notices to alert people to misleading or biased news sources, both on a case-by-case basis and in their overall news consumption. New tools could gradually introduce alternate points of view, encouraging users to break out of ideological echo chambers. Smart citizen phone apps now allow users to develop localized crowdsourced maps revealing problem areas for litter, broken streetlights and windows, vandalism, potholes, and so on. Aptly designed visceral notices could track users' interventions and encourage citizens to increase their levels of civic awareness and engagement on local, national, and international levels, prompting them to take action where help is needed.

Skill maintenance. Our willingness to outsource tedious physical engagements with the external world may lead to a significant loss of everyday skills. GPS mapping and automated driving systems are cases in point. When following the visual or voice directions today's systems offer, users don't need to pay attention to landmarks and therefore may not be able to recall routes taken. Visceral notices offer a potential corrective. An AI-enabled system could include a setting that would mimic the way a person on the street might give directions but enhanced by 3D images of key landmarks and points of reference where turns must be made. This would give users the option of orienting themselves to their surroundings and relying on their own memory to reach their destinations instead of mechanically following voice commands as they are given. Other designs could encourage drivers to stay alert and to maintain their driving skills instead of becoming overly reliant on automated driving systems.

Personal edification. Ultimately, what aptly designed visceral notice environments can provide are AI systems that act less like objects and more like friends that help users develop to their fullest potential. Consider the capacity of AI systems to encourage greater discernment in domains such as the arts, cuisine, fashion, and entertainment. Instead of exposing people to whatever products they may react most impulsively to, as recommendation engines often do, they could show alternatives with high-quality ratings based not merely on popularity but also on a blend of expert opinion and personal and shared social preferences. Some services such as Netflix already provide such distinctions, but without a feature showing how the user's overall viewing choices and screen time map to the quality ratings.

AI-ASSISTED PLATFORMS provide consumers with extraordinarily powerful tools for controlling and managing their daily lives, activities, and interactions. Such technology, if designed carefully and conscientiously, also holds the power to alter human behavior for the better on a massive scale. But if designed shortsightedly, with few if any features for counteracting its own negative habit-forming effects, it could instead foster passivity, dependency, ignorance, and vulnerability. Applied to millions, these forces undermine the systems of liberal democracy and capitalism.

It is essential that companies working in this area formulate clear and cogent design strategies to allow customers to make informed choices regarding their own patterns of online behavior. The ones that do will establish stronger relationships with their customers while playing a key role in optimizing collective well-being by safeguarding personal agency.

***Julian Friedland** is an assistant professor of ethics at Trinity Business School at Trinity College Dublin. Comment on this article at http://sloanreview.mit.edu/x/60401.*

REFERENCES

1. R. Wollheim, The Thread of Life (New Haven, CT: Yale University Press, 1984).

2. M.J. Sandel, "The Case Against Perfection: What's Wrong With Designer Children, Bionic Athletes, and Genetic Engineering," Atlantic Monthly, April 2004, 1-11; and S. Vallor, Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting (Oxford, U.K.: Oxford University Press, 2016).

3. B. Frischmann and E. Selinger, Re-Engineering Humanity (Cambridge, U.K.: Cambridge University Press, 2018).

4. G. Lukianoff and J. Haidt, The Coddling of the American Mind: How Good Intentions and Bad Ideas Are Setting Up a Generation for Failure (New York: Penguin Random House, 2018).

5. D. Kahneman, Thinking, Fast and Slow (New York: Farrar, Straus and Giroux, 2011).

6. J. Friedland and B.M. Cole, "From Homo-Economicus to Homo-Virtus: A System-Theoretic Model for Raising Moral Self-Awareness," Journal of Business Ethics 155, no. 1 (March 2019): 191-205.

7. Brett Frischmann and Evan Selinger describe the six forms of disengagement in Re-Engineering Humanity.

8. N. Carr, The Glass Cage: How Our Computers Are Changing Us (New York: W.W. Norton & Co., 2015).

9. W. Knight, "Socially Sensitive AI Software Coaches Call-Center Workers," MIT Technology Review, 2017, www.technologyreview.com.

10. P. Lin, "Why Ethics Matters for Autonomous Cars," in M. Maurer, J.C. Gerdes, B. Lenz, et al., eds., Autonomous Driving: Technical, Legal, and Social Aspects (Berlin: Springer Open, 2016), 69-85.

11. Sandel, "The Case Against Perfection."

12. Carr, The Glass Cage.

13. R.H. Thaler and C.R. Sunstein, Nudge: Improving Decisions About Health, Wealth, and Happiness (New Haven, CT: Yale University Press, 2008).

14. K.G. Volpp, A.B. Troxel, S.J. Mehta, et al., "Effect of Electronic Reminders, Financial Incentives, and Social Support on Outcomes After Myocardial Infarction: The HeartStrong Randomized Clinical Trial," JAMA Internal Medicine 177, no. 8 (August 2017): 1093-1101.

15. Vallor, Technology and the Virtues, 161.

16. B. Lamm, H. Keller, J. Teiser, et al., "Waiting for the Second Treat: Developing Culture-Specific Modes of Self-Regulation," Child Development 89, no. 3 (June 2018): e261-e277.

17. T.W. Watts, G.J. Duncan, and H. Quan, "Revisiting the Marshmallow Test: A Conceptual Replication Investigating Links Between Early Delay of Gratification and Later Outcomes," Psychological Science 29, no. 7 (May 2018): 1159-1177.

18. A. Duckworth and J.J. Gross, "Self-Control and Grit: Related but Separable Determinants of Success," Current Directions in Psychological Science 23, no. 5 (October 2014): 319-325.

19. Vallor, Technology and the Virtues, 163.

20. Friedland and Cole, "From Homo-Economicus to Homo-Virtus."

21. K. Aquino and A. Reed II, "The Self-Importance of Moral Identity," Journal of Personality and Social Psychology 83, no. 6 (December 2002): 1423-1440.

22. S. Bowles, The Moral Economy: Why Good Incentives Are No Substitute for Good Citizens (New Haven, CT: Yale University Press, 2016).

23. K. Hwang and H. Kim, "Are Ethical Consumers Happy? Effects of Ethical Consumers' Motivations Based on Empathy Versus Self-Orientation on Their Happiness," Journal of Business Ethics 151, no. 2 (2018): 579-598.

24. J. Sadowski, S.G. Spierre, E. Selinger, et al., "Intergroup Cooperation in Common Pool Resource Dilemmas," Science and Engineering Ethics 21, no. 5 (October 2015): 1197-1215.

25. R.B. Cialdini, C.A. Kallgren, and R.R. Reno, "A Focus Theory of Normative Conduct: A Theoretical Refinement and Re-Evaluation of the Role of Norms in Human Behavior," Advances in Experimental Social Psychology 24 (December 1991): 201-234.

26. J. Sadowski, T.P. Seager, E. Selinger, et al., "An Experiential, Game-Theoretic Pedagogy for Sustainability Ethics," Science and Engineering Ethics 19, no. 3 (September 2013): 1323-1339.

27. A. Gopaldas, "Marketplace Sentiments," Journal of Consumer Research 41, no. 4 (Dec. 1, 2014): 995-1014.

28. Hwang and Kim, "Are Ethical Consumers Happy?"

29. L. Columbus, "2017 Roundup of Internet of Things Forecasts," Forbes, Dec. 10, 2017, www.forbes.com.

i. Kahneman, Thinking, Fast and Slow.

ii. M.R. Calo, "Against Notice Skepticism in Privacy (and Elsewhere)," Notre Dame Law Review 87, no. 3 (October 2013): 1027-1072.

Reprint 60401.

PDFs ■ Reprints ■ Permission to Copy ■ Back Issues