CUTTER CONSORTIUM
Access to the Experts

Cutter Business Technology JOURNAL

Vol. 31, No. 2, 2018

Management, Innovation, Transformation

Artificial Intelligence:
Fear It, Face It, or Embrace It

"Artificial intelligence is creating a lot of excitement and hype among professionals and across all kinds of business and industry, as well as among individuals. It is a new, innovative approach for solving some of the challenging problems that we encounter in practice, as well as an enabler of disruptive innovations."

— San Murugesan,
Guest Editor

Cutter Business Technology JOURNAL

As business models for creating value continue to shift, new business strategies are constantly emerging and digital innovation has become an ongoing imperative. The monthly *Cutter Business Technology Journal* delivers a comprehensive treatment of these strategies to help your organization address and capitalize on the opportunities of this digital age.

Cutter Business Technology Journal is unlike academic journals: Each monthly issue, led by an expert Guest Editor, includes five to seven substantial articles, case studies, research findings, and/or experience-based opinion pieces that provide innovative ideas and solutions to the challenges business technology professionals face right now — and prepares them for those they might face tomorrow. *Cutter Business Technology Journal* doesn't water down or delay its content with lengthy peer reviews. Written by internationally known thought leaders, academics, and practitioners — you can be certain you're getting the uncensored perspectives of global experts.

You'll benefit from strategic insight on how the latest movements in digital innovation and transformation, IoT, big data analytics and cloud, to name a few, are changing the business landscape for both new and established organizations and how cutting-edge approaches in technology leadership, enterprise agility, software engineering, and business architecture can help your organization optimize its performance and |transition to these new business models.

As a subscriber, you'll also receive the *Cutter Business Technology Advisor* — a weekly bulletin featuring industry updates delivered straight to your inbox. Armed with expert insight, data, and advice, you'll be able to leverage the latest business management thinking to achieve your organization's goals.

No other journal brings together so many thought leaders or lets them speak so bluntly — bringing you frank, honest accounts of what works, what doesn't, and why. Subscribers have even referred to *Cutter Business Technology Journal* as a consultancy in print and likened each month's issue to the impassioned discussions they participate in at the end of a day at a conference!

Get the best in thought leadership and keep pace with the technologies and business models that will give you a competitive edge — subscribe to *Cutter Business Technology Journal* today!

Cutter Business Technology Journal®

Founding Editor: Ed Yourdon
Publisher: Karen Fine Coburn
Group Publisher: Christine Generali
Managing Editor: Cindy Swain
Copy Editors: Jennifer Flaxman, Tara Meads
Production Editor: Linda Dias
Client Services: service@cutter.com

Cutter Business Technology Journal® is published 12 times a year by Cutter Information LLC, 37 Broadway, Suite 1, Arlington, MA 02474-5552, USA (Tel: +1 781 648 8700; Fax: +1 781 648 8707; Email: cbtjeditorial@cutter.com; Website: www.cutter.com; Twitter: @cuttertweets; Facebook: Cutter Consortium). ISSN: 2475-3718 (print); 2475-3742 (online).

Subscription rates are US $485 a year in North America, US $585 elsewhere, payable to Cutter Information LLC. Reprints, bulk purchases, past issues, and multiple subscription and site license rates are available on request.

☐ Start my print subscription to *Cutter Business Technology Journal* ($485/year; US $585 outside North America).

Name _____ Title _____

Company Address _____

City _____ State/Province _____ ZIP/Postal Code _____

Email (Be sure to include for weekly *Cutter Business Technology Advisor*) _____

Fax to +1 781 648 8707, call +1 781 648 8700, or send email to service@cutter.com.
Mail to Cutter Consortium, 37 Broadway, Suite 1, Arlington, MA 02474-5552, USA.

Request Online License Subscription Rates

For subscription rates for online licenses, email or call: **sales@cutter.com** or **+1 781 648 8700.**

CUTTER CONSORTIUM
Access to the Experts

Opening Statement

by San Murugesan

Old is new again — for good reasons. Artificial intelligence (AI) is now 62 years old; the term was coined in 1956 at a Dartmouth College workshop. Despite intense initial interest, AI developments never quite took off back then and failed to permeate into practice. However, artificial intelligence has now seen a resurgence and has yielded success in practice on several fronts. It has received a facelift, promoting a fresh new look. AI's renaissance is driven by recent complementary developments, including major advances in the AI arena, realistic expectations, and proven success in its application in various domains, such as finance, healthcare, manufacturing, and agriculture, particularly to address some complex, challenging problems.

Artificial intelligence is no longer just the theme of science fiction essays and movies; it is real and is transforming the way we live, work, and do business. AI encompasses several related technologies, including machine learning (ML), deep learning (DL), natural language processing and translation, and chatbots.

Along with developments in big data, cloud computing, analytics, and the Internet of Things (IoT), artificial intelligence is driving rapid change across all industries. Its impact on individuals and on business and industry will be profound. Some predict that its impacts will surpass those of electricity and industrialization. Indeed, large investments are being poured into research, development, and marketing of AI products, tools, and services. There is also greater awareness among stakeholders than ever before of the advances in AI and of the promises, limitations, and concerns around it. Technology professionals, executives, and businesses need to be aware of ongoing developments in the emerging AI landscape and should harness AI's potential — both known and hidden — while addressing its risks and concerns.

This scenario raises a few pertinent questions among business executives and technology professionals: Where is AI headed? What new applications and innovations will emerge? What are the real and perceived risks, limitations, and concerns, if any? How will AI transform businesses and industry sectors? What is an appropriate AI strategy? What will evolve as the "new normal"? What new opportunities will arise for the IT industry and technology professionals, and how should they prepare now for excelling in a new age of AI?

In this issue, we examine some of these questions along with the drivers of AI global trends and their implications — now and in the future. Our contributing authors provide insights on key opportunities, strategies, and approaches for realizing AI's potential and discuss emerging issues and concerns, including how AI may impact jobs and businesses.

The AI Renaissance

Artificial intelligence is creating a lot of excitement and hype — much more than any other new technology — among professionals and across all kinds of business and industry, as well as among individuals. It is a new, innovative approach for solving some of the challenging problems that we encounter in practice, as well as an enabler of disruptive innovations. In the past five years, there have been significant developments in AI, particularly in ML, DL, and chatbots. These developments facilitate autonomous decision making and operations and enable AI applications to learn to perform better from their interactions with us and their environments, to reason with purpose, and to interact more naturally with humans and other smart systems. Helpful applications of AI are emerging in areas such as research and discovery, decision support and advisory, customer engagement, customer experience management, healthcare, the IoT, and cybersecurity. We are also beginning to examine how AI-based systems and humans work together and complement each other's capabilities. And companies across industries are seriously examining how they can benefit from applying AI in their organizations.

Like any new technology, AI offers promises but also has limitations and presents some risks and concerns.

Concerns include its social implications, potential bias in AI systems, the ethics of some of its applications, and its lack of empathy. To benefit from AI's potential, AI applications must be seen as trustworthy and dependable to gain the acceptance of users.

Advances in artificial intelligence and judicious applications of the advances to address real-life problems have the potential to bring innovations that truly improve business operations and customer engagement, change the nature of work, and boost socio-economic development in the developed and developing worlds. They also have the potential for better human-machine fusion, complementing and exploiting each other's unique strengths and capabilities. These innovations will result in disruption so we must acknowledge the new challenges AI will present — not only technical and organizational challenges, but also social, legal, and regulatory hurdles.

IT professionals need to prepare for the new age of AI with education and (re)training. We also need to create new platforms and tools for development, testing, and deployment, and, perhaps more important, a new mindset among developers, users, and regulators. While AI seems to offer huge opportunities, researchers, developers, and businesses alike should focus on addressing meaningful and purpose-driven problems that confront business and society, rather than attempting to address hypothetical or toy problems and coming up with useless or fanciful applications.

Upcoming Topics

Disciplined Agile: Roadmap to Business Agility
Scott Ambler

Blockchain Technology
Phil O'Reilly

Transforming the Customer Experience
Jeanne Bliss

The Intelligent Edge of IoT
Ron Zahavi

While the new age of AI seems real, how it will transform what we currently do and how we address as-yet-unresolved problems and the extent of its impact and penetration — not only in business and industry but also in common use by people of different profiles and needs — will depend upon how we collectively harness its potential, understanding its limitations and addressing the challenges. Let's prepare ourselves — and the society we are expected to serve — for a better future shaped by AI and other supporting technologies.

This issue aims to reveal some of the emerging AI landscape and serve as inspiration. We present seven thoughtfully selected articles written by established practitioners, consultants, and researchers in AI.

In This Issue

Artificial intelligence can be viewed from many angles and through multiple lenses, and depending how you view it, you will come up with different perspectives on the technology. As a starting point for this issue, Cutter Fellow Steve Andriole presents a brief, multi-faceted overview of AI — "the good, the disruptive, and the scary" — and sets the backdrop for further exploration. He outlines some recent advances in and key limitations of AI and explores how AI could disrupt several domains, such as insurance, banking, law, real estate, and education. He then discusses the impact that the deployment of intelligent systems will have on jobs and the professional opportunities that will arise.

AI comprises different technologies and approaches. To exploit its full potential in a given context or application, it is important to connect and embrace those "dots" collectively. In our second article, Cutter Senior Consultant Borys Stokalski, Bogumił Kamiński and Przemysław Szufel emphasize the need to connect and collectively harness advances in different elements of AI and outline autonomous business entities as examples of convergence of AI. The authors discuss the application of AI, not only to improve business operations, but also for product adaptation and to enable and support business model innovations, thereby making the entire business "smart." They also explore "data labs" and "data factories," which facilitate business model innovation. Finally, the authors argue that while AI-driven, radical automation of businesses will replace human work in some areas, humans will remain relevant in others.

Next, we draw your attention to the design, development, deployment, and refinement of cognitive

computing systems (CCSs). While CCSs are deployed in a variety of fields yielding benefits exceeding expectations, there are also major failures. According to authors Kevin Desouza, Lena Waizenegger, and Gregory Dawson, lack of appreciation for the differences inherent in developing a CCS versus a traditional software system is key to these failures. To assist in developing successful CCSs and to derive benefits from them, they offer a set of 10 key recommendations based on their examination of over two dozen systems. They conclude that CCSs will be a dominant technology that will permeate all business operations for the foreseeable future.

The next three articles showcase how AI is being deployed in three major sectors: banking, government, and insurance. First up, Hemamalini Kumaran, Prema Sankaran, and Raj Gururajan discuss how AI is transforming the banking sector. They outline how Indian and US banks are using AI to gain significant benefits and offer an enhanced customer experience. The authors examine the key drivers that inspire banks to embrace AI, the challenges involved in implementing it, and what needs to be considered in applying AI to best serve customers.

Perhaps surprisingly, AI is gaining interest from the government sector, too. In their article, Vipin Jain and Seema Jain discuss the opportunities emerging from artificial intelligence and how cognitive technologies will fundamentally change the way government works. They outline how the US public sector is currently adopting and planning to embrace AI and ML in various applications. They also highlight priorities for federally funded research in the US. To help developers in conceiving and developing AI applications, the authors present an AI adoption framework and briefly discuss the categories of AI-branded services available from leading cloud service providers. They finish with a consideration of whether AI is a job creator or a job destroyer.

Artificial intelligence will compel the adoption of new business models. In his article, Raj Ramesh discusses business model transformation with a focus on the insurance sector. He covers the potential of AI in insurance and then expands his discussion to the ingredients necessary for AI to provide good value to the business in any sector.

Finally, we draw your attention to a key barrier to wider adoption of AI and ML: trust. Researchers Keng Siau and Weiyu Wang examine prevailing concepts of trust in general and in the context of AI applications and human-computer interaction in particular. They discuss the three types of characteristics that determine trust in this area: human, environment, and technology. They emphasize that trust building is a dynamic process and outline how to build trust in AI systems in two stages: initial trust formation and continuous trust development.

Imagining New Realities of AI

Artificial intelligence is driving rapid change across all industry sectors. We are entering new territory and hoping to benefit from the opportunities that AI and ML present. On one hand, ongoing developments and the resulting possibilities are impressive and likely to further transform the world. On the other hand, there are concerns regarding the ethical implications and potential harmful effects of AI applications on society. Moreover, to gain dominance in the AI landscape, there is fierce competition among technology giants as well as startups.

We hope the articles in this issue present perspectives and ideas on fulfilling the promise of artificial intelligence and that you'll find them interesting, insightful, and practical. We also hope this issue of *Cutter Business Technology Journal* will help you "imagine the new possible" and inspire and encourage you to harness advances in AI in your domain of interest, addressing any concerns and limitations, for good.

San Murugesan (BE [Hons], MTech, PhD; FACS) is a Senior Consultant with Cutter Consortium's Data Analytics & Digital Technologies *practice, Director of BRITE Professional Services, and an Adjunct Professor in the School of Computing and Mathematics at Western Sydney University, Australia. He is Editor-in-Chief Emeritus of the IEEE's* IT Professional*. Dr. Murugesan has four decades of experience in both industry and academia, and his expertise and interests include AI, the Internet of Everything, cloud computing, green computing, and IT applications. He offers certificate training programs on key emerging topics and keynotes.*

Dr. Murugesan is coeditor of a few books, including Encyclopedia of Cloud Computing *and* Harnessing Green IT: Principles and Practices*. He is a member of the COMPSAC Standing Committee, and a fellow of the Australian Computer Society. Dr. Murugesan held various senior positions at Southern Cross University, Australia; Western Sydney University; the Indian Space Research Organization, Bangalore, India; and also served as Senior Research Fellow of the US National Research Council at the NASA Ames Research Center. He can be reached at smurugesan@cutter.com.*

DOOMED FOR DISPLACEMENT DAMAGE

AI: The Good, the Disruptive, and the Scary

by Steve Andriole

Those who develop and sell "intelligent" applications understand the extraordinary implications — and opportunities — of artificial intelligence (AI) solutions. Amazon, IBM, SAP, IBM, Oracle, Google, Cisco, Microsoft, Alibaba, Comcast, HP, Intel, Facebook, Apple, eBay, Thermo Fisher Scientific, Samsung, Dell, Foxconn, Huawei, and Baidu — just about all the players — are racing across the value chain to sell what they have to sell: smart applications that can save money *and* make money. CEOs, COOs, CIOs, CISOs, CDOs, CFOs, and CTOs are *all* waiting impatiently to deploy applications that will save them time, effort, and money — especially money they now spend on expensive humans. They see AI as a cost manager and a profit center. Is this a revolution waiting to happen? Absolutely yes.

AI will continue to impact industries and professions vulnerable to deductive replacements and routine automation. But — for the first time — AI will also displace a whole lot of knowledge workers — *well-educated professionals* — especially in the banking, legal, education, real estate, and insurance industries, among a broad range of service industries. Its impact on the transportation and manufacturing industries will accelerate as well. In fact, there's no limit to the applied potential of AI, which will conceivably become as ubiquitous as databases, networks, and personal computing devices of all kinds.

The Good

The good news is that artificial intelligence is already many things to many people and represents a variety of alternative methods, tools, and techniques either already in the field or under aggressive development. An important distinction among intelligent solutions is strength.[1] "Weak" AI solutions are ones essentially preprogrammed to "solve" problems in well-bounded, usually deductive domains like routine task management (e.g., record-keeping, FAQs and answers, and customer service). "Strong" AI solutions are those capable of extrapolating and "reasoning" beyond well-bounded domains, such as those that can "diagnose" problems in real time.

Weak AI system "intelligence" is defined as the consistent, persistent performance of repetitive, repeatable tasks. Filing a tax return only requires weak AI, since tax rules are well-defined (well-bounded), deductive, and codified. Strong AI systems can "think" by generalizing intelligence across problem domains. While we're some years away from intelligent systems with "consciousness," strong AI systems will eventually mimic human capabilities, though there's heated debate about how "human-like" they can ever become.[2]

The good news is that weak AI is exploding. But don't confuse "weak" with ineffective. Some of the domains to which weak AI is applied include financial trading, market analysis, insurance underwriting, fraud protection, plagiarism checking, email management (including spam filtering), and multiple forms of tax preparation. These applications are inexpensive to develop, easy to deploy, and therefore extraordinarily cost-effective. In fact, their use has already become best practice.

Strong AI is on its way in several forms. Intelligent automation, especially robotic process automation,[3] is adaptive to changing tasks. Machine learning — especially deep learning[4] — super-charges intelligence with powerful knowledge representation techniques like neural network modeling.[5] Enabling these applications are tools like machine vision,[6] robotics, and natural language processing.

Billions of dollars are pouring into the development of both weak and strong AI. According to CB Insights (and other industry investment trackers),[7] AI remains one of the largest investment areas for private equity venture capitalists, corporate venture capitalists, and companies across multiple industries that the technology industry has ever seen. High levels of financial interest are predictors of technological progress.

The Disruptive

"Disruption" is a matter of perspective. It has been placed in quotes here because artificial intelligence is indeed disruptive, which many industries, companies, and executives find exciting as well as threatening (which may explain why most disruption comes from startups and not industry incumbents). There are some industries that AI will disrupt more (and faster) than others, such as insurance, real estate, banking, law, and education. Each of these industries will be disrupted not just by weak and strong AI, but also by regulatory changes that will clear the way for major disruption. Weak AI is already hard at work in these industries, but stronger AI is banging on the door. Let's look at some of the major processes in these industries and how AI will disrupt not only these processes and but whole business models.

Insurance

The insurance industry has already been attacked by digital agents, but the digital army is now poised for a takeover. Most millennials do not use home, auto, and life insurance agents:

> Instead of working with local agents to find the right coverage, 67% of millennials are purchasing directly from insurance companies, leaving agents out of the picture.[8]

Why in the world would anyone with a computer or smartphone make an appointment with a human being and physically travel to an office? If present trends continue, insurance agents will disappear completely about the same time we bury the last baby boomers. Agents and the entire process will be replaced by intelligent chatbots and virtual assistants, which will optimize insurance products according to the individual circumstances of each client. The same assistants will handle claims. The processes that are especially "vulnerable" include underwriting, claims processing, transaction management, fraud detection, risk management, and, perhaps most importantly, insurance planning. The shift from weak to strong AI will be disruptive, especially when intelligent virtual assistants replace human insurance agents, which will occur steadily over the next five to seven years.

Real Estate

The real estate industry is under attack from companies like Open Listings and the larger "for sale by owner" (FSBO) community. But the traditional players have some powerful friends that lobby endlessly to keep their hold on how real estate is bought and sold. There are so many hands in the typical transaction that it's impossible to easily identify all of the financial vested interests in real estate transactions — which makes the industry difficult, but not impossible, to disrupt. Consider executive coach and author Bruce Kasanoff's point of view:[9]

> 95% of a broker's role could be handled better by well-designed technology systems. Bidding, for example, could be handled by an automated system that includes legally-binding documents that would be instantly accessible to each party's attorney … the fact is that one thing keeps the broker's role alive today: the regulations that govern the real estate industry.

The implications for AI here are everywhere. The entire list/bid/buy/sell/regulate process can become victim to automation. Much of the automation only requires weak AI; strong AI will manage advanced bidding and negotiation. Customized listing processes will be developed, implemented, and managed by intelligent systems. Property search will be automated based on preferences and life circumstances that virtual assistants will calculate. The property search process will become proactive the moment a buyer expresses interest in moving or renting, when buyer and market databases will merge into personalized and customized options.

Banking

According to Moven founder Brett King (as reported by Eric Rosenbaum of CNBC):[10]

> The biggest banks in the world in 2025 will be technology companies, and banks that grew through branch acquisitions in the '80s and '90s, that grew by physical bank presence, will have a real problem.

Money is also disappearing. Way back in 2012, *International Business Times* correspondent Jacey Fortin reported that:[11]

In Sweden, monetary transactions made with physical cash are down to three percent of the national economy. In most Swedish cities, public buses don't accept cash; tickets are prepaid or purchased with a cell phone text message.

The US lags, of course, but it's only a matter of time and money — especially because of the control that cashless transactions provide governments and the financial gains banks accrue from closing physical branches and going cashless. What aspects of banking *cannot* be automated? Which processes would *not* benefit from weak and strong AI?

AI-with-blockchain will be the preferred banking architecture, but basic banking processes are not the only ones vulnerable to intelligent systems. Wealth and portfolio management will be empowered by advanced analytics capable of customized predictive *and prescriptive* planning, the holy grail of financial management.

> *Let's acknowledge the lack of intelligence around artificial intelligence.*

Law

Why are there so many lawyers? There's no need for so many physical, organic, living, breathing legal professionals in the digital era. Expertise defined around "rules" can be automated and distributed at the touch of a key, a verbal command, or a reasonably intelligent (even weak) assistant. Automated reasoning (strong AI) will replace many lawyers (and, for that matter, doctors, accountants, professors, and engineers), primarily because "the law" and other professions are well-bounded, codified domains — precisely what intelligent systems require to excel. Disruption has already arrived. According to data privacy lawyer Sterling Miller, three events are imminent:[12]

1. Some legal jobs will be eliminated (e.g., those [that] involve the sole task of searching documents or other databases for information and coding that information are most at risk).

2. Jobs will be created, including managing and developing AI (legal engineers), writing algorithms for AI, and reviewing AI-assisted work product (because lawyers can never concede the final say or the provision of legal advice to AI).

3. Most lawyers will be freed from the mundane task of data gathering for the value-added task of analyzing results, thinking, and advising their clients.

Consider also legal vendors like LegalZoom and Atrium that understand how the process from weak to strong AI will evolve. Weak AI applications — what we still call "expert systems" — power many of the self-service functions at LegalZoom and Atrium. Soon, legal advice dispensed by human lawyers will be replaced by intelligent assistants with de facto JDs informed, again, by personalized and customized databases.

Education

And why are there so many teachers? There are so many teachers, trainers, and professors because there's a tremendous need for their services. But there are problems. Many of the domains that they teach, such as science, technology, engineering, and mathematics (STEM), are dynamic. The currency of the teachers, trainers, and professors who teach domains that frequently change is a challenge. Pedagogy also tends to be static among veteran teachers, trainers, and professors.

Weak AI can automate the basics of education, such as enrollment, curriculum optimization, and grading. It can also customize and personalize learning experiences, correlating with personal and behavioral data that will improve the learning process. Intelligent analytics can measure effectiveness, which results in teacher/trainer/professor assessments, as well as assessments about the effectiveness of curriculum content. Said differently, total quality management can be automated by intelligent assistants. Intelligent tutors can assist where necessary.

The Scary

First, let's acknowledge the lack of intelligence around artificial intelligence. Members of the US Congress, for example, know little or nothing about the technology — which is worrisome on so many levels, especially when we consider the technology's inevitable impact on the US and global economies. Most CEOs — and even most CIOs and CTOs — also know relatively little about AI — though, when surveyed, they list AI as one of the most important technologies of the 21st century. The judicial system has its head in the sand. The general population only appears to understand AI in the way

Hollywood dramatizes it, like the way it was exhibited in 1992 in *Minority Report*, in 1999 in *The Matrix*, and, more recently, in the movie *Her*, HBO's *Westworld*, and several episodes of Netflix's *Black Mirror*.

Try this: go to a party and randomly ask the guests what they think about AI. I've done it several times and the word cloud illustrates robots, Alexa, Watson, and Westworld, but nothing about machine learning, knowledge representation, or neural networks. The gap here is scary.

Weak and strong AI will displace labor, but what happens when widespread job displacement occurs? Hardly any of the pundits describe specific displacement management strategies or plans. This is the scary part of the story (not AI hostages or AI-instigated Armageddon). How many industries and companies will know how — *or even want* — to manage displacement, especially if displacement management offsets savings from labor reduction?

Corporate HR departments will explode with complaints and lawsuits, and collapse under the weight of the exit packages they'll be forced to give, at least temporarily. Young *and* aging factory workers — along with lawyers, accountants, and educators — will forget their purpose. Politicians will stare into the technology headlights — again — frozen by their own confusion and vested self-interests. Executives and shareholders will squeal with profitable delight. Universities will adjust their curricula or rapidly lose customers. Pain will pervade the corridors (but not the boardrooms) of the hard *and* soft industrial worlds, though this time the corridors will be wider and prettier than they've been in past displacement revolutions because knowledge workers typically ply their trades in prettier places.

Transition — and displacement management — will be the challenge. This is the part of the AI story that deserves much more attention, though very few analysts talk about what happens after displacement occurs. What happens to the paralegals, lawyers, accountants, medical diagnosticians, manufacturers, supply chain managers, and customer service representatives when they're displaced? Where will they go? What will they do?

Consider the doom-and-gloom proposition from strategy and management consultant Xavier Mesnard at the *World Economic Forum:*[13]

> The risk we are facing in the near future is mass unemployment for some categories of workers, combined with lack of skills in other categories — and the political and social implications of such imbalances.

While Mesnard's predictions may be scary, they deserve serious attention. (What if he's right?)

The assumption is — as it always is — that we have time to think all this through (or ignore it for as long as we can); perhaps even as much time as it took cars to totally replace the horse-and-buggy, or about a quarter of a century. But displacement this time will be much faster and more brutal than it's been in the past. There won't be anywhere near as much time as in past industrial revolutions because this revolution is digital and therefore, by definition, pervasive and exponential. We're already digitally connected, so displacement has an infrastructure in place, which is an unprecedented revolutionary reality (note, for example, how the adoption of automobiles was constrained by the lack of a highway infrastructure).

One example boldly illustrates the point: self-driving vehicles will be widely deployed in five to seven years — not decades from now. The enabling digital technologies that locate, direct, and power autonomous vehicles are in place, which will accelerate adoption and displacement. What will happen to the operators of these vehicles and those who enable the entire vehicle and transportation supply chain?

The emerging crisis is not about inevitable displacement but about reaction and replacement policy. Since corporations will largely benefit from displacement (in the same way they benefit from cheaper global labor markets), governments will be expected to intervene at some level. But given how clueless government already is, and how slowly it moves even when it's informed and committed, the prospects for effective displacement management are low. It may be that we're focusing on the wrong problem — and the wrong problem-solvers. It may be that completely unpredictable events will define displacement management policy. So how *do* we manage displacement?

We have less than a decade to figure all this out, which is a problem given the current level of intelligence about AI. There's a lot of education and analysis necessary to minimize transitional and displacement damage and optimize another global transition. The backdrop is complicated by a variety of other competing priorities. The first step is awareness.

Now What?

We have seen whole industries disrupted by digital technology over the past two decades. But all that disruption combined only represents Disruption 1.0. The five industries discussed here represent just the beginning of Disruption 2.0. Like the iPhone, there are likely to be many incarnations over time. As machines get smaller, smarter, and cheaper, we'll see more and more industries disrupted by digital professionals and their digital tools. The implications of continuous disruption are extensive and unpredictable. The world as we think we know it will never be (just as it never was).

AI has the potential to quickly supplement and displace huge numbers of hard *and* soft skills. We need some *real* intelligence to avoid a rocky process that could have major social, economic, and political consequences. Are there members of government willing to launch these discussions at the national level? Are there editors of major media willing to raise red flags? Would it make sense to seek global agreement? (CNN has time to analyze the power of music and meals. Maybe it could do a series on AI.)

This is clearly a multidisciplinary problem. Colleges and universities should focus on the implications of displacement. Government and foundations should fund research in the area. Corporations should develop displacement management plans. The major social platforms — Facebook, Instagram, Twitter, and so on — should proactively encourage discussions about the displacements about to occur.

AI will disrupt knowledge- *and* production-based professions and the fields that prepare and maintain these professionals. Note that "trickle-down disruption" will be much more impactful and sustained than "first-order disruption." You can decide how much about AI is "good," "disruptive," and "scary." There's enough to go around, that's for sure.

Endnotes

[1]Regnard (anonymous). "Difference Between Weak and Strong AI." Computer Science Exchange, 21 August 2014 (https://cs.stackexchange.com/questions/29283/difference-between-weak-and-strong-ai).

[2]Nogrady, Bianca. "The Real Risks of Artificial Intelligence." BBC Future, 10 November 2016 (www.bbc.com/future/story/20161110-the-real-risks-of-artificial-intelligence).

[3]"What Is Robotic Process Automation?" Institute for Robotic Process Automation & Artificial Intelligence (https://irpaai.com/what-is-robotic-process-automation/).

[4]Hof, Robert D. "Deep Learning." *MIT Technology Review* (www.technologyreview.com/s/513696/deep-learning/).

[5]Hardesty, Larry. "Explained: Neural Networks." MIT News, 14 April 2017 (http://news.mit.edu/2017/explained-neural-networks-deep-learning-0414).

[6]"Machine vision." Wikipedia (https://en.wikipedia.org/wiki/Machine_vision).

[7]Norton, Steven. "Corporate VCs Invest $1.8 Billion in AI in First Half of 2017: CB Insights." *The Wall Street Journal*, 27 July 2017 (https://blogs.wsj.com/cio/2017/07/27/corporate-vcs-invest-1-8-billion-in-ai-in-first-half-of-2017-cb-insights/).

[8]Narcisco, Eric. "5 Reasons Millennials Aren't Buying Insurance from Local Agents." ThinkAdvisor, 16 September 2015 (www.propertycasualty360.com/2015/09/16/5-reasons-millennials-arent-buying-insurance-from?slreturn=1518548002).

[9]Kasanoff, Bruce. "Are Real Estate Brokers Obsolete?" *Forbes*, 4 April 2014 (www.forbes.com/sites/brucekasanoff/2014/04/04/real-estate-brokers-obsolete/#344caa997953).

[10]Rosenbaum, Eric. "Retail Bank Branch Is Doomed, and Banks Don't Know It." CNBC, 2 June 2015 (www.cnbc.com/2015/06/02/retail-bank-branch-is-doomed-and-banks-dont-know-it.html).

[11]Fortin, Jacey. "Sweden Going Cashless: Pros and Cons of Paper Money." *International Business Times*, 19 March 2012 (www.ibtimes.com/sweden-going-cashless-pros-cons-paper-money-214321).

[12]Miller, Sterling. "Artificial Intelligence and Its Impact on Legal Technology (Part II)." Above the Law, 7 August 2017 (https://abovethelaw.com/2017/08/artificial-intelligence-and-its-impact-on-legal-technology-part-ii/).

[13]Mesnard, Xavier. "What Happens When Robots Take Our Jobs?" World Economic Forum, 21 January 2016 (www.weforum.org/agenda/2016/01/what-happens-when-robots-take-our-jobs/).

Stephen J. Andriole is a Fellow with Cutter Consortium's Business Technology & Digital Transformation Strategies *and* Data Analytics & Digital Technologies *practices and the Thomas G. Labrecque Professor of Business Technology at Villanova University. Dr. Andriole was the Director of the Cybernetics Technology Office of the Defense Advanced Research Projects Agency (DARPA); the CTO and Senior VP of Safeguard Scientifics, Inc.; and the CTO and Senior VP for Technology Strategy at Cigna Corporation. His most recent books include* Ready Technology: Fast Tracking New Business Technologies *and* The Innovator's Imperative: Emerging Technology for Digital Transformation. *He has published articles in* MIT Sloan Management Review, Communications of the ACM, IEEE IT Professional, *and* European Business Review, *among others. He can be reached at sandriole@cutter.com.*

PRIME TIME FOR AI

Connecting the AI Dots:
Radical Scenarios for Business Automation

by Borys Stokalski, Bogumił Kamiński, and Przemysław Szufel

We are witnessing explosive growth in the practical applications of various flavors of artificial intelligence (AI), from chatbots powered by cognitive engines, through advanced image recognition, smart robots and vehicles, to advanced analytical solutions based on machine learning (ML) methods and tools. Moreover, the proponents of artificial general intelligence (AGI) platforms have made significant progress in their effort to create machines that exhibit creativity and have the potential to master complex games on a level unmatched by human players.

Furthermore, there is a growing portfolio of tools for advanced analytical processing that eliminates the need to master the advanced engineering skills of high-volume, parallel data processing. This allows data analysis experts, already proficient in the design and evaluation of complex analytical models, to work much more productively. Consequently, the entire discipline of data science will accelerate in terms of the number of people involved, progress in methods and practices, and value delivered.

While numerous specialized AI applications can clearly provide solid ROI, two key questions remain: How can the "dots" be connected? And what are the key "points of convergence" in AI evolution? This article focuses on one such point of convergence — autonomous business entities. This idea has already been hailed in such experiments as the AI-based hedge fund from Aidyia[1] and a "decentralized autonomous organization" (DAO),[2] a blockchain-based concept of a "company without a company."

The Evolution of Process Automation

Programmable machines invaded manufacturing processes in the early 1970s. The number of heavy-duty industrial robots operating today is close to 2 million, and that number is growing constantly. Information workflow automation exploded some two decades later, borrowing the paradigms from the manufacturing world and concepts developed by Frederick Winslow Taylor, the father of scientific management.[3] It has quickly become evident that such an approach delivers results mostly when applied to routine and predictable processes.

In today's dynamic business environment, however, information work has become ever-challenging with two opposing factors: (1) the flexibility required to allow for fast innovation and real-time responsiveness and (2) the efficiency achieved through Lean, optimized processes. Moreover, hierarchical process control must give way to emergent control. Agility, so much desired in modern business, requires the "frontline actors" to decide how to achieve goals within such constraints as time and resources available. Indeed, increasing competition and shorter product lifecycles require faster and cheaper "to market" activities and, at the same time, better alignment with customer expectations.

While numerous specialized AI applications can clearly provide solid ROI, two key questions remain: How can the "dots" be connected? And what are the key "points of convergence" in AI evolution?

This puzzle has given rise to service-oriented architectures and adaptive, goal-based process management. What does that mean? That processes must be injected with operational intelligence. By applying sophisticated models to analyze operational data coming from sensors, mobile devices, or embedded systems, the entire business cycle can be automated. This includes the cognitive aspects, which have long been human monopolies. Image recognition, auditory capabilities, and natural language processing are already handled "well enough" by machines, and the rate of improvement of machines' cognitive capabilities will accelerate due to the consumerization of technology. Machines are currently outsmarting humans in many areas, so it seems possible and perhaps even necessary to follow

a radically different paradigm: virtual business entities engaging customers, transacting business, and optimizing resources with minimum human intervention (or, should we say "consultation").

Predictive modeling is one piece of a three-element puzzle required to deliver best-in-class solutions. Other ingredients are optimization and simulation methods.

The Smart Supply Chain

Business is not just about bits; it is also about atoms. Things that are manufactured need to be handled in the supply chain all the way from basic components down to the finished product arriving at the customer's doorsteps or shop shelves. But as those "things" become smarter, the physical value chains and information processes converge.

Much of today's R&D and startup investment is going into smart logistics. Swarms of smart robots handle the material streams on manufacturing shop floors or at distribution centers. Robotic support for logistics ranges from entire systems to smart autonomous devices and components of store infrastructure (e.g., Amazon Robotics). Independent autonomous robots capable of transporting a variety of mobile containers while safely navigating through crowded plant spaces using just their sensors and onboard intelligence can be integrated through API and mission-planning applications with third-party warehouse management software. A good example is the Warsaw-based VersaBox.[4] This robotic company manufactures autonomous robots that roam the factories of car manufacturers and automotive component suppliers.

Smart logistics has its broader twin in the evolution of unmanned transport vehicles, such as driverless cars, drones, and trains. Legislators around the globe are currently working on laws that will enable driverless vehicles to operate on public roads. With initiatives like "Uber Elevate" (Uber working with NASA and air traffic experts on a proposed air transport system for the 2028 Los Angeles Summer Olympics), it is not unlikely that autonomous airborne logistics will become part of our daily experience in the coming decade.

The next logical step after autonomous transport services become part of the daily infrastructure is the concept of "smart cargo." Smart cargo can communicate with the sender, receiver, and transport services to optimize the way to its destination. As a result, order fulfillment processes will become distributed across many agents of the smart world — a tangle of digital services ecosystems.

Such multi-agent networks are currently being actively researched for their qualities and management strategies. Predictive modeling is one piece of a three-element puzzle required to deliver best-in-class solutions. Other ingredients are optimization and simulation methods, two workhorses of the operations research discipline. The true game-changer is the real-time combination of these methods' predictive models, fed with an unprecedented level of granularity of data that can be used to tune them online. Apart from giants like Uber or Amazon, who naturally have the scale to develop and maintain such solutions, it is interesting to observe that such technology can be readily applied by smaller players and even startups.

Take Instacart, "a same-day grocery delivery service."[5] Instacart customers can order products via an application and get the goods delivered by a personal shopper that same day. To squeeze out maximum performance of deliveries and maintain the ability to ensure high service levels, Instacart has built a complex analytical ecosystem. Building on the predictions of customer demand and shopper characteristics, the retailer derives optimal staffing levels and routing plans for deliveries. The whole process is wrapped in a stochastic simulation framework, which assesses the expected service level of each operational decision. Such a strong backing by analytical solutions allows Instacart to grow fast — and investors like it. The startup has had a fresh, big round of financing this year, putting it at a value of US $4.2 billion.

Automating Customer Engagement

Every business is only as sound as its capabilities to identify potential customers, present them with a value proposition at the right time and in the right context, and turn such contacts into a transaction, or, better yet, a profitable relationship.

These activities can in many cases be executed by machines without significant intervention of humans.

AI provides us with numerous ways to automate marketing, sales, and customer support processes. Examples include state-of-the-art chatbots based on AI/ML platforms and programmatic marketing services that can be used as effective, real-time segmentation and communication tools to crunch data generated by millions of Web users, which has been collected via global search, content, and social media platforms.

A well-known example of this AI technology are digital assistants, or conversational service hubs, such as Google Home or Amazon Alexa. These sophisticated Web services communicate with users through a range of interoperable devices either sold by the vendor or by third-party providers. Such cognitive tools can also be used as third-party components to enhance products. Car manufacturers such as BMW are already announcing the intentions to implement Alexa directly inside of cars. Another practical use case for AI technology is the automation of call centers where humans are only the last line of support. Predefined software components necessary for the delivery of such a solution are available today from cloud platform providers.

Adaptive Products

Customer engagement processes offer a broad spectrum of automation opportunities spanning all aspects of a customer's journey. The act of managing the product portfolio, however, remains in most cases a human responsibility. It is quite difficult to imagine fully automated product portfolio management — especially in the case of covering innovations related to customer value proposition. However, if we look at the product innovation process from the consumer's perspective, we might view it as a process of advanced product or service personalization, opening new possibilities for automation. Instead of designing a product, we build a model of a customer and design communication tools that help them configure an optimum customer experience. By combining cognitive technologies, advanced analytics, and flexible product delivery mechanisms (e.g., 3D printing, APIs to product functionalities, value-added partners and ecosystems), we can envision some form of "automated hyper-personalization," which may replace the need for internally driven product improvement/innovation as currently practiced.

Such an approach is unlikely to deliver breakthrough product innovations. It is also unrealistic to expect physical products such as cars, smartphones, or refrigerators to redesign or reshape themselves to better reflect changing customer preferences. However, as more product features are delivered through software, products become platforms, interconnected and open for integration with third-party products and services. As such, assets requiring extensive service and maintenance (e.g., cars) can be converted into services offerings. Replacing ownership with usage fuels the rise of a new category of smart products, including Google Home or Alexa Skills,[6] that can serve as aggregators of customer needs and delivery hubs for digital features provided by various parties. Therefore, the role of "smart" objects and places (e.g., smart cars, smart homes, smart offices) as mediators between an increasingly complex world of interconnected intelligent things and personalized human needs will become significant.

Customer engagement processes offer a broad spectrum of automation opportunities spanning all aspects of a customer's journey.

Adaptive Business Models

We have reviewed how business automation can take advantage of AI solutions in various aspects of business operations. We believe that connecting the "dots" provides a convincing argument that as the technology evolves, human involvement in operations will transform from execution to configuration, oversight, and refinement of autonomous operations so that the technology itself can accurately implement and execute the entire business model. Therefore, the next logical step in our quest for radical automation is to look for ways to use AI to automate activities beyond operations. Specifically, we want to look at making businesses adapt to changes in the environment by introducing changes in the various business cycles that implement the business model.

A digital online entertainment service business model might include typical operational business cycles covering customer engagement, onboarding, provisioning, monetization, and revenue assurance along with content management and even content design and production. We can look at how business cycles are implemented through the lens of an "observe, orient,

decide, and act" (OODA) loop[7] — a reference model of an adaptive event-response system. From this point of view, every business cycle involves coordination and execution of actions (sometimes complex ones) that are triggered by events recognized through observation (gathering data) and orientation (translating data into business-relevant events). In this view, adaptability involves two capabilities:

1. "Reprogramming" actions
2. Evaluating and improving models used for orientation

Current practices in process automation that best support adaptability involve separation of "execution logic" from "decision management logic." Execution is driven by a portfolio of simple, highly granular services representing atomic actions, such as manipulation of attributes of specific business objects, managing relationships between objects, and so on. These are then "brought to life" by decision logic comprised of rule sets activated by business events.

In such an approach, the question of adaptability of business models can be reduced to the adaptability of business cycles and further reduced to the evaluation of the efficiency of rule sets controlling the execution logic (ORIENT->DECIDE->ACT or ORIENT->ACT threads of the OODA loop) and improvement/modification of these rules.

Current practices in process automation that best support adaptability involve separation of "execution logic" from "decision management logic."

Lessons Learned from Automating Business Models

Companies that successfully incorporate smart analytics and achieve constant improvement of their OODA loops implement and orchestrate the following three elements to automate optimization of business models:

1. Alignment, usually supported by an appropriate incentive scheme, of the business and IT departments toward experimentation and innovation
2. Data collection and governance throughout the whole organization
3. Flexibility of IT architecture and supporting infrastructure

Alignment

Alignment of all stakeholders toward implementation of advanced analytics seems an obvious recommendation. However, there is a recurring pattern that leads to failure in this area: the people responsible for implementation of advanced analytics solutions are so strongly motivated to demonstrate their "hard value" for the organization that they never admit inevitable occasional failures. An organizational solution described by Thomas C. Redman and Bill Sweeney that aims to avoid this problem is a division of responsibilities between two units: the data lab and the data factory.[8]

The *data lab* is often centralized within the organization and has a strong focus on analytical skills. Its main objectives are fostering innovation across the organization and the creation of prototypes in fail-fast mode. The data lab team usually has long-term objectives to ensure that its creativity is not hindered by "end of quarter" syndrome. However, the prototypes that the data lab produces are low value unless someone ensures that they are implemented operationally.

This is where the *data factory* team comes into play. This team must possess strong business processes and technology expertise. Its task is to ensure the smooth transformation of the products of the data lab into working production solutions. The cultural backbones of the data factory team are quality and efficiency, so that the relevant innovations are delivered and bring value to the organization.

Data Collection and Governance

The second key challenge area is data collection and governance. We often hear that data is the "new oil," and storage is cheap. Less advertised is the fact that this oil needs serious analysis and refinement to be used effectively for analytical models driving actual decisions. Data is often poorly understood, as few organizations have current data models of their corporate data assets, and external data sources evolve fast.

Unfortunately, there is no silver bullet to solve this issue. Organizations need to get a grip on this problem by the pragmatic implementation of data governance processes, starting from the management board level. There are well-developed guidelines in this area (e.g., *DAMA DMBOK*®[9]). Implementation of those initiatives is often mandated by governments attempting to regulate the "flow of data." A good example of this is the EU implementation of the General Data Protection Regulation (GDPR).

The second aspect of the data collection process is more challenging and hard to codify. Even if we collect and understand the data, it still might not contain the information required to develop appropriate meaningful analytics solutions. Too often, managers ignore the fact that without experimenting and exploration, they can't build a fully valid propensity-to-buy model for a campaign or insist on selecting the most efficient sales or marketing actions without some form of A/B testing.

Flexibility

The third key challenge is technology-related. Again, there are two aspects to this issue. The first is related to technical debt. The interests paid for technical debt are high costs of change and integration — actions such as introducing new functionality or creating a "single version of truth" are expensive and time-consuming. This impairs the capability to implement the data factory concept.

Refactoring of legacy architecture combined with the migration to cloud platforms can allow an organization to achieve radical improvements in IT performance — economic and technical aspects alike. The development of complex analytical solutions is related to the highly varying demand for computing power. The data lab team usually "thinks and investigates," applying modest computing resources. However, when it comes to model development, the team instantly requires several orders of magnitude more data crunching power. The only reasonable way to provision such an infrastructure is by moving computations to the cloud. Fortunately, the competition between leading public cloud providers drives the costs down and quality up. Spinning up a farm of several thousands of servers in a matter of minutes is normal practice today. However, running computations in the cloud creates several challenges, including:

- **Data migration.** Unless the company has totally moved its IT infrastructure to the cloud, there must be a reliable way to move the data from on premises to the cloud.
- **Cloud security.** Cloud services are secure, but it is easy to misuse them without proper governance; we hear about such problems all too often (e.g., the recent FedEx customer data leak[10]).
- **Cost management.** Pricing models of cloud services are usually complex and require a different approach than on-premises architectures; financial departments need to adjust their monitoring policies to a pay-as-you-go cost model.

Common practices used in on-premises environments are cost-inefficient. They do not take advantage of auto-scaling offered by the cloud computing services or serverless approaches (where components are provisioned, managed, and scaled by the cloud provider). Hence, a properly designed cloud architecture and basic cost-optimization techniques can lead to cost reductions of up to 90% compared to simply replicating the assets of an on-premises software system in the cloud.

Refactoring of legacy architecture combined with the migration to cloud platforms can allow an organization to achieve radical improvements in IT performance — economic and technical aspects alike.

Automating the Data Factory

As we look for viable radical scenarios, let's examine to what extent the design, evaluation, and constant improvement of decision logic can be supported by technology and automated. In an ideal world, starting with some initial set of services forming the execution logic and an initial set of rules forming the decision logic, the process of business model evolution should run automatically — going back to humans only when something starts to fail (which should also be automatically detected).

One obvious issue that needs to be addressed is the interoperability of the diverse tools and technologies

being used today by the data science community — in particular, the broad and quickly evolving open source toolbox. Without proper tools, it is virtually impossible to keep track of different versions of R, Python, or Julia and the associated packages. To solve those problems, analytics solutions more often use containerization technologies such as Docker or Kubernetes or packaging and deployment toolkits like R Suite for solutions developed in R.

Autonomous Business Agents

Arguably, the most radical opportunities for business automation are offered by blockchain technology. Blockchain is best known as the base for several cryptocurrencies, but this is not our focus. What seems interesting is the transaction mechanism in blockchain that can be used to monitor, record, and share arbitrary business events in a verifiable and permanent way. Smart contracts are an even more interesting mechanism directly enabling an autonomous business. A smart contract can be executed within a blockchain whenever a certain set of rules is met. A cryptocurrency cost to smart contract execution (called "gas") is paid by the executing party. The right to execute a smart contract can be conditioned by the ownership of a token.

These mechanisms allow the creation of an autonomous virtual business that operates solely based on a set of rules defined within smart contracts without any external actions (see Figure 1). This can be further combined with customer engagement platforms, such as digital assistants, automated marketing, outsourcing logistics, drop shipping, and automated decision making. Hence, today we already have a set of technologies sufficient to create an entire retail marketing company without a single employee — a DAO. One might ask why customers would be willing to enter into contracts with such an organization. The answer is "trust through transparency." Blockchain provides full transparency, and any party wanting to enter into a transaction with a DAO has full information on its operational rules and knows that those rules are irrevocable since they are stored in a blockchain ledger. Hence, the transparency creates the trust that allows the virtual business to operate.

There are still humans involved in the DAO model. People can join a given virtual business as stakeholders and propose rules and contracts to be voted on by other participants. So DAO offers a radical way to organize (there is no other legal infrastructure required apart from smart contracts) and operate (smart contracts execute the logic triggered by business events recognized by the DAO platform). Moreover, once a

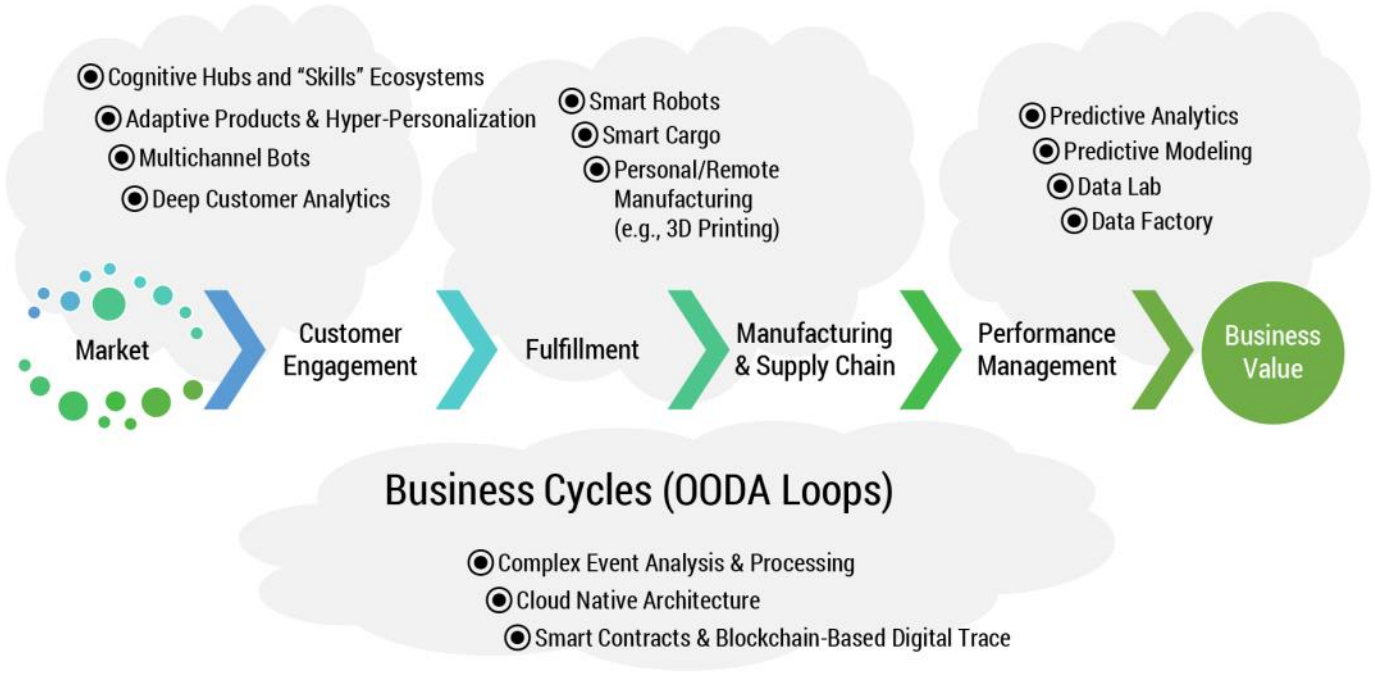

Figure 1 — Connecting the dots of autonomous business agent.

virtualized company is set up, it only incurs costs when a customer-initiated event occurs (e.g., an order is placed). Hence, every cost is strictly connected to a profit (e.g., a sale event). Therefore, such a company can operate with very little fixed costs and very predictable variable costs.

Since a DAO company is represented by its code, it will be easy for new business offerings to emerge, such as commercially available building blocks of various business functionalities to create blockchain companies by combining these functionalities. Combined with low fixed operational costs, this ability to easily create new companies can lead to situations where millions of blockchain autonomous companies will emerge. Instead of trying to manage product portfolios, companies will simply spin off swarms of "disposable," highly specialized virtual companies handling different product innovations. The absence of fixed costs also means that there is very low business risk related to a company having a very narrow specialization (e.g., offering only one product).

Challenges and Limitations of AI

Assuming that we know how to create autonomous operations that implement a given business model, now comes the big questions: can complex management tasks, which involve creativity, intuition, and insight, be automated? What are the boundaries and assumptions to be made to create a set of automated business cycles that result in some form of autonomous, adaptive, and fully automated business?

Big expectations are associated with deep reinforcement learning — a technology that demonstrates that machines can learn things that humans are not able to fully understand. There are several spectacular showcase examples of applying AGI in platforms developed by Google's DeepMind team around complex games. Recently, AlphaZero defeated the strongest "classical" chess-playing program "Stockfish" after learning to play chess for only eight hours. AlphaZero scored 28 wins and no losses in a 100-game match.[11]

Machines need data to learn. When learning to play chess, AlphaZero was able to generate the data itself (by simulating playing games). Sometimes such a simulation is also possible in a real business environment (e.g., Instacart). However, often this is hard to design as the algorithm must obtain feedback from the system about the possible outcomes of its actions.

Another challenge is more prosaic. As of right now, humans have more versatile senses than a computer. What we mean by this is best explained by a real-life example from a large retail chain. Even if you were to build a sophisticated tool for retail demand forecasting, that tool would occasionally make a bad mistake simply because it lacks essential information that humans can comprehend and operationalize in the blink of an eye. One instance of such a gross error occurred when a machine did not consider that a road leading to a store was undergoing renovation. Any store manager would have easily predicted a drop in sales due to the construction. The machine failed because no one thought of implementing such an explanatory variable in the data warehouse. And even if this had been done, such an event is probably rare enough that there would be a risk that the ML algorithm would not be able to efficiently incorporate the signal provided by such a variable in predictions.

When it comes to autonomous cars, they may have a sharper camera and radar than the human eye, but it can only capture the data it was designed to capture. On the contrary, the human brain trains all its life to process various kinds of often seemingly unrelated information. This human experience pays off when we capture and correlate new facts in new contexts, and cognitive flexibility remains a human advantage. Such problems can be solved by building a set of (usually complex) business rules on top of the AI model. In a recent failure of AI, the absence of such rules turned an experimental chatbot developed by Microsoft into a "virtual hate speech offender" after only one day of activity on Twitter.[12]

A final caveat: increasing adoption of AI technologies will affect future trust levels. A recent international survey of CEOs by PricewaterhouseCoopers reveals some potential problems involving trust in AI:[13]

- There is always a risk that the designer of an algorithm is malevolent or that some adversary breaks into the system and tweaks an algorithm.
- The complexity of algorithms means that people may not fully understand what the machines would do in every situation. Indeed, humans may have difficulty understanding *why* a machine has performed in a way it did. This is probably acceptable if we are talking about marketing campaigns, as the potential losses are limited, but is a major problem if we, for example, let machine policemen autonomously patrol the streets.

- Often, there is no clear guidance as to how the machine should perform in a given situation, especially when all courses of action are potentially harmful to humans; such a situation is a standard problem, for example, in the design of autonomous cars. Indeed, Uber and autonomous car manufacturers are now dealing with the recent tragedy of the first fatality by a self-driving car, when neither the self-driving system nor the person behind the wheel applied the brakes when a pedestrian suddenly appeared on the roadway at night.[14]

These are the challenges of the present and the future, and many companies are already establishing proper governance policies to face them.

Is the human mind the equivalent of a Turing machine? This may be the fundamental question of human relevance in the age of smart machines.

Possible Impacts and Consequences of Radical Automation

There are many promising ways to integrate artificial intelligence, advanced business intelligence, and modern "cloud native" service architectures to advance the way we approach business automation. Indeed, we can envision several areas where human work will likely be replaced, and machines will create value streams. Soon, many businesses may become software constructs, using cloud services to implement and improve operations and business models or using smart objects where atoms must be handled (i.e., physical goods or customer touchpoints). The integrity of the business can be maintained through reliable, trusted digital traces of events and automatically executed smart contracts. The factors that will drive such radical automation may be diverse — starting with reduced transaction costs (one of the major drivers of business automation over decades), through reducing time and costs of business cycles, and ending with better control by eliminating human errors, imperfections, and performance variations. It remains to be seen to what extent such expectations will be satisfied.

With every radical change, there will be consequences. Thus, we need to understand the role of human actors in the future economy of autonomous business agents. We believe the following three roles will continue to require human capacity:

1. **Filling in the gaps in machine capabilities.** As we have pointed out, machines make great progress in learning and using acquired knowledge to optimize operational performance, but there are still pockets of human superiority, mostly where synthetic, holistic thinking or a correlation of loosely related facts are concerned.

2. **Curating the autonomous organizations** by monitoring their decisions and courses of action, analyzing results and issues, trimming rule sets, and enforcing and improving policies. There is a very well-known example of a DAO (known as "The DAO") running within the Ethereum cryptocurrency with a failure in its implementation.[15] This made it possible for people exercising this vulnerability to transfer all money to their cryptocurrency accounts. The crisis has been addressed by Ethereum, and over 50% of the cryptocurrency miners agreed to revoke the transaction. But the story is an early caveat of some consequences of radical business automation.

3. **Analyzing and improving the rules and algorithms "invented" by autonomous businesses** to design new business models and better architectures for next-generation entities.

But is the human mind the equivalent of a Turing machine? This may be the fundamental question of human relevance in the age of smart machines. There are many serious arguments that this is not the case; hence, the limitations of what can be achieved by a program (even a sophisticated one) executed by a machine might not apply to the human mind. If this is true, then the evolution of computers requires humans — we might be outsmarted, but we will not be outthought. At least until we invent a computing architecture capable of implementing sentient thinking and mathematical intuition.

Endnotes

[1]Metz, Cade. "The Rise of the Artificially Intelligent Hedgefund." *Wired*, 25 January 2016 (https://www.wired.com/2016/01/the-rise-of-the-artificially-intelligent-hedge-fund).

[2]"Decentralized Autonomous Organization." Wikipedia (https://en.wikipedia.org/wiki/Decentralized_autonomous_organization).

[3]"Frederick Winslow Taylor." Wikipedia (https://en.wikipedia.org/wiki/Frederick_Winslow_Taylor).

[4]"VersaBox" (http://www.versabox.pl).

[5]"Instacart" (https://www.instacart.com).

[6]For an interesting overview of third-party Alexa Skills see: Martin, Taylor. "50 Most Useful Alexa Skills." CNET, 29 January 2018 (https://www.cnet.com/how-to/amazon-echo-most-useful-alexa-skills).

[7]For more about the OODA loop and its role in the smart business architecture, see: Stokalski, Borys, and Bogumił Kamiński. "Outsmarting the Competition: Information Superiority in the Digital Age." Cutter Consortium Business Technology & Digital Transformation Strategies *Executive Update*, Vol. 20, No. 6, 2017 (https://www.cutter.com/article/outsmarting-competition-information-superiority-digital-age-495151).

[8]Redman, Thomas C., and Bill Sweeney. "To Work with Data, You Need a Lab and a Factory." *Harvard Business Review*, 23 April 2013 (https://hbr.org/2013/04/two-departments-for-data-succe).

[9]*DAMA International Guide to Data Management Body of Knowledge (DAMA DMBOK®)*. DAMA International, 2015 (https://dama.org/content/body-knowledge).

[10]Forrest, Conner. "Leaked FedEx Customer Data Was Stored on Amazon S3 Server with No Password." TechRepublic, 15 February 2018 (https://www.techrepublic.com/article/leaked-fedex-customer-data-was-stored-on-amazon-s3-server-with-no-password).

[11]Klein, Mike. "Google's AlphaZero Destroys Stockfish in 100-Game Match." Chess.com, 6 December 2017 (https://www.chess.com/news/view/google-s-alphazero-destroys-stockfish-in-100-game-match).

[12]Vincent, James. "Twitter Taught Microsoft's AI Chatbot to be a Racist Asshole in Less than a Day." *The Verge*, 24 March 2016 (https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist).

[13]"20 Years Inside the Mind of the CEO ... What's Next?" PricewaterhouseCoopers, 2017 (https://www.pwc.com/gx/en/ceo-survey/2017/pwc-ceo-20th-survey-report-2017.pdf).

[14]Naughton, Keith, and Ryan Beene. "The Self-Driving Uber in Fatal Crash Didn't Have a Vision Problem." *Bloomberg Businessweek*, 21 March 2018 (https://www.bloomberg.com/news/articles/2018-03-21/for-self-driving-cars-seeing-everything-isn-t-always-enough).

[15]"The DAO (organization)." Wikipedia (https://en.wikipedia.org/wiki/The_DAO_(organization)).

Borys Stokalski is a Senior Consultant with Cutter Consortium's Business Technology & Digital Transformation Strategies *practice. Mr. Stokalski is also a partner in the digital business design studio re:Think, a seasoned advisor, manager, and investor with over 25 years' experience in building and managing one of the leading consulting and IT solution implementation companies on the Polish market, recently acquired by a European Top 10 software integrator. Today, Mr. Stokalski remains an active advisor and investor focusing on the growing digital business in Poland. He can be reached at bstokalski@cutter.com.*

Bogumił Kamiński is Head of the Decision Analysis and Support Unit at the Warsaw School of Economics, Poland, and Adjunct Professor at Data Science Laboratory, Ryerson University, Toronto. Dr. Kaminski has authored over 50 research articles on theory and business applications of prediction, simulation, and optimization methods. During the last 15 years, he has delivered over 100 data science projects for the largest Polish corporations. He can be reached at bkamins@sgh.waw.pl.

Przemysław Szufel is an Assistant Professor in the Decision Support and Analysis Unit at the Warsaw School of Economics, Poland. His current research focuses on distributed systems and methods for execution of large-scale simulations for numerical experiments and optimization. Dr. Szufel is currently working on asynchronous algorithms for the parallel execution of large-scale computations in the cloud and distributed computational environments. He is the author/coauthor of several open source tools for high-performance and numerical simulation (including SilverDecisions, which enjoys over 20,000 users). He can be reached at pszufe@gmail.com.

KEEP YOUR EYES WIDE OPEN

9 Recommendations for Designing, Developing, Deploying, and Refining Cognitive Computing Systems

by Kevin C. Desouza, Lena Waizenegger, and Gregory S. Dawson

Cognitive computing systems (CCSs) are debuting in a variety of fields and, in many of those deployments, the outcomes have exceeded expectations, causing interest in CCSs to spike. At the same time, there have also been spectacular failures. Most often, these failures are due to organizations underappreciating the differences inherent in developing a CCS versus a traditional software system. In this article, we draw on our examination of more than two dozen CCSs to put forth a series of recommendations that can guide organizations in recognizing the differences in developing a CCS versus developing a traditional system.

An Introduction to CCSs

Organizations are increasingly leveraging CCSs to address a wide range of problems — from optimizing solutions (e.g., processing data, analyzing risks, and offering recommendations) to solving complex social problems (e.g., mapping global poverty). CCSs are a suite of technologies that encompass machine learning (ML), natural language processing (NLP), data mining, large-scale data processing, pattern recognition, and visualization components. They come in many flavors — from autonomous vehicles to digital assistants and chatbots to advanced decision support systems. CCSs can do the following:[1]

- Learn in an iterative manner from both data and human interactions through the application of ML algorithms, continuously ingesting data and building new knowledge and learning models through synthesizing data.
- Personalize interactions with humans through NLP.
- Be context-sensitive, thereby recall history and tailor outcomes in a personalized manner.
- Provide confidence-weighted recommendations (outcomes) to humans and/or other computing systems autonomously.
- Be adaptive (i.e., adapt based on new data and information and tolerate ambiguity and incompleteness).
- Have memory (i.e., remember previous interactions and begin where something was left off).
- Be iterative (i.e., work through a series of steps to solve a problem).

Examples of CCSs abound. Hong Kong's Immigration Department uses CCSs to sort visa, passport, and residency applications.[2] Applications are classified as clearly approved, clearly denied, or in a grey area. An officer then reviews each application, considers the initial recommendations on similar cases by the CCS, and makes a final decision on the application. This saves time and effort, allowing the officer to "sanity check" the assignments after performing an initial determination.

Singapore is also a frontrunner in exploring and adopting artificial intelligence (AI) technologies. A virtual assistant ("Ask Jamie"), for instance, has significantly reduced the number of calls to government call centers.[3] Ask Jamie can pull up information from websites across multiple agencies, thereby enabling users to resolve queries more effectively and efficiently. Singapore residents and businesses can also utilize another chatbot ("Ask Jasmine") for support while completing tasks such as filing tax returns.[4]

Instances where CCSs have failed to deliver on their promise, or have had unintended consequences, also abound. For example:

- UK insurance company Admiral was blocked in October 2016 from analyzing Facebook posts to categorize the risk level of potential drivers based on Admiral's interpretation of what their posts indicated about their personality and lifestyle.[5]
- Two Swiss artists created a Random Darknet Shopper (RDS) chatbot.[6] The RDS used $100 Bitcoins

to purchase illegal and counterfeit items on the Darknet or Dark Web. The Kunst Halle art gallery in St. Gallen, Switzerland, displayed these items as a public display and called it an awareness-raising campaign. While it was a hit, the Swiss police nailed the bot when the gallery displayed a packet of drugs as the bot's latest purchase. This raises the question of accountability and liability when an autonomous CCS engages in illegal behavior and highlights the apparent unpredictability of CCS operations.

- Microsoft's chatbot "Tay" turned into a racist, misogynist, sexist, and Holocaust-denying bot in about 24 hours after going online.[7] Tay was trained to leverage ML and converse with people on Twitter, Kik Messenger, and GroupMe. Microsoft developers created Tay to learn from its interactions and respond in an entertaining manner. The bot initially started as a naive teenager on Twitter; however, within hours it learned and started repeating offensive statements and propagating conspiracy theories (e.g., 9/11 was an inside job). In short, Tay was unable to distinguish between good and bad learning and leveraged both equally.

Clearly, organizations must think differently about designing and deploying CCSs than they would with traditional information systems. CCSs require more care than traditional systems for the following three reasons:

1. They need to learn from data and human experts to perform.
2. They are continuously learning and evolving and hence are never completely "built" or "finished."
3. They trigger key organizational issues (e.g., workforce planning, business process reengineering) that arise from automation and augmentation.

Strategic oversight and governance are crucial during the four phases of the CCS implementation lifecycle: design, development, deployment, and refinement (see Figure 1). In the remainder of this article, we walk through the key considerations and recommendations within each phase.

Designing CCSs

Designing CCSs requires a strong case for the investment into those systems. Organizations must not only be able to justify the initial investment into developing a CCS, but also think through the investments that will be needed to ensure it can be refined and enhanced over time. Unlike traditional information systems that are deployed when fully functional and then are updated (patched) on a regular basis, CCSs are different. CCSs are not deployed when they are ready to work in an autonomous manner. Instead, they are initially deployed in a training mode, and then refined over time until they achieve a given level of performance. After that, they continue to learn and update and are never completely "finished."

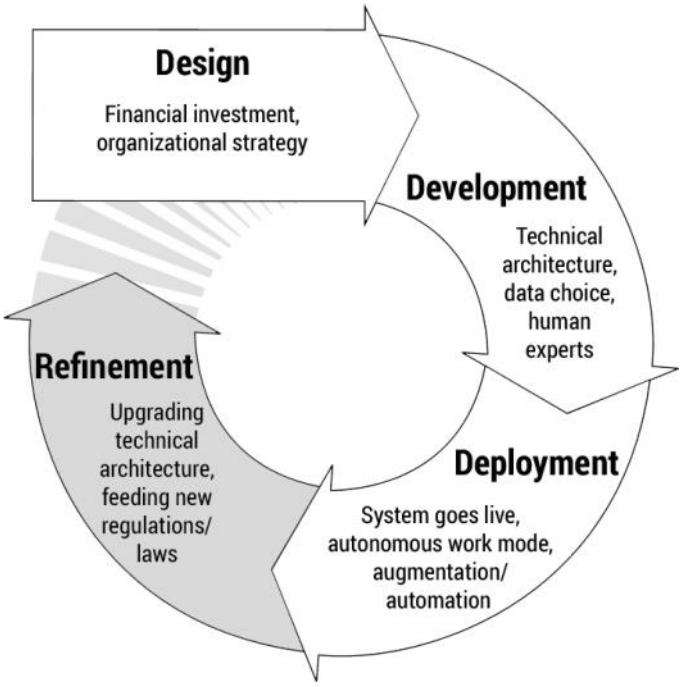

Figure 1 — The CCS lifecycle.

CCSs also call for organizational transformation investments to realize business value. A strategic framework must be developed to guide the organization as it takes a long-term view of how CCSs will influence the overall posture of the organization, including its value proposition, service delivery, stakeholder engagement, and future value.

Recommendation 1: Build the Strategic Case

Think broadly about the future of your organization in the age of CCSs. Engage key stakeholders across business lines. Develop a strategic framework that can guide your efforts in the near and medium term. Organizations that lack a strategic framework often waste technology investments in CCSs due to poor appreciation of how to embed them within the existing organizational fabric and, more important, how to derive value from them.

CCS efforts should be linked to clear and measurable business outcomes. As such, it is vital for an organization to look at its existing product and service lines, the organization's current efforts in these spaces, the external trends that might impact the future trajectories of those product and service lines, and then situate the role of CCSs from a value proposition perspective.

The strategic framework should also account for current major business processes and personnel management issues. Any CCS implementation will require some level of business process reengineering and will impact how existing personnel conduct their work. Thinking through the process and personnel issues that will need to be addressed as an outcome of digital transformation will enable the organization to ensure that it is preparing the current environment for the CCS as it is being conceived and developed, rather than waiting for it to be ready and then acting on these complicated issues.

A viable strategy for organizations getting into the CCS environment is to choose to develop "edge" cases for CCSs.

Recommendation 2: Build a Capacity for Experimentation

Organizations should have the capacity to experiment with CCSs. This involves making CCSs part of the organization's overall innovation and investment strategies. Employees should be empowered to suggest, design, and develop CCS prototypes. Mechanisms to test prototypes and run experiments to gauge their efficacy should be in place. Employees working in product and service lines should be able to collaborate with IT personnel on CCS ideation.

Toward this end, organizations must also undertake a detailed examination of their IT innovation capacity. To build CCSs in-house, an organization must have the requisite expertise available, including talent in ML, data science, and so on. If not, then the organization should investigate whether its current IT vendors have the expertise and capability necessary to leverage as part of existing and/or new projects.

Recommendation 3: Develop Edge Cases with Live Systems Running in Parallel

A viable strategy for organizations getting into the CCS environment is to choose to develop "edge" cases for CCSs (i.e., a system that pushes the edge on what's possible with technology). These edge cases can be deployed alongside existing live systems, but they are very challenging to get right. For these edge cases, the risk is simply too high for the CCS to be implemented without an appropriate safeguard of an existing live system. This allows the organization to avoid costly mistakes associated with deploying these live without enough testing. In addition, it provides the organization with an opportunity to compare the performance and outcomes of the CCSs with current system performance.

Developing CCSs

CCSs begin their development in a controlled setting, where developers build the ontologies and algorithms the system is running on. The initial system is fed with a training set and generates adaptive learning mechanisms that correlate inputs with the correct outputs. The system also learns from human experts and developers, who fix incorrect outputs and provide feedback to improve the system's performance. This phase is where developers deal with the data issues to ensure that the system uses the right kind of data: data that is free from biases (e.g., no gender or racial bias) so that ML systems can go to work on it. Bias-free data will always be a challenge, and the developers and human experts will need to periodically audit data to ensure that bias has not crept in (see Recommendation 9).

In this phase, experts are engaged to develop and train the system. While the system is being built, a simultaneous effort is taking place to ensure that the necessary changes are being made in the organization to accommodate the CCS (e.g., changes in job assignments, rethinking workflows). This positions the organization for a successful go-live.

Recommendation 4: Invest in Data Governance

CCSs live and die based on the data they ingest to build their learning models. As such, it is vital for organizations to take great care when choosing which projects might be best suited for CCSs given their data peculiarities. Two common strategies are (1) to begin with

projects where the data is already structured, curated, and stored effectively, thereby limiting data governance challenges; or (2) to choose small, well-defined, and bounded problems that, due to their limited scope, can be tackled even if the data set is small or there might be a need for significant data-cleansing efforts. At the very minimum, organizations need to ensure that the data being fed into a CCS during its development phase is accurate and representative of the problem and solution space.

Recommendation 5: Engage Human Experts

It is important that CCSs are not only fed with a diverse data set but also trained by a variety of human experts with different viewpoints and knowledge backgrounds to avoid any "trainer bias." Human experts need to spend a significant amount of time to assess the accuracy of the system's output regarding known problems and solutions pairs and to correct the system's answer if it is wrong. The system saves that feedback and uses it in future iterations. Assessing and revising the output of the CCS requires time but is crucial for the performance and accuracy of the system. The organization will need to assume the up-front costs of pulling knowledge experts from tasks that might generate revenues in the near term to allow them to participate in developing the system. Organizations that diligently involve experts in the development of the system are more likely to witness better outcomes with CCSs.

Recommendation 6: Get Human Expert Buy-In on Decisions

Closely related to our previous recommendation, the human experts must not only be involved in designing the system but also must, eventually, be willing to accept the recommendation the CCS makes unless there is evidence to suspect that the recommendation is unacceptable. Organizations will need to alleviate the human experts' fears of being replaced by machines (or losing value due to those machines). Once the system is operational, human experts must become comfortable with the decisions the system recommends, although those human experts may occasionally disagree with those decisions. CCS protocols need to ensure that a human expert can override the system's decisions when warranted. One might consider this akin to flying an airplane: under most conditions, commercial flights can operate on autopilot, but when situations warrant due to changes in the internal or external environments, the pilot can take over and command the aircraft.

Deploying CCSs

Initially, CCSs are deployed in a controlled setting, where they work on tasks alongside humans. The outputs of the CCSs are compared with those of the humans to measure performance (e.g., correctness or validity of recommendation, accuracy, speed, time). In cases where a CCS gives an erroneous response, human experts provide feedback to the CCS so that it can learn from and incorporate the correct solutions in future interactions. Over time, it is expected that the performance of the CCSs will improve and meet the required criteria to be deployed autonomously.

It is important that CCSs are not only fed with a diverse data set but also trained by a variety of human experts with different viewpoints and knowledge backgrounds to avoid any "trainer bias."

Recommendation 7: Proactively Monitor for Unexpected Outcomes

CCSs are continuously evolving entities. They are constantly learning based on feedback from expert users but also when they ingest and analyze new data and refine their learning functions. It would be fatal to leave these systems to their own devices. It is important to implement monitoring and governance structures to avoid unintended negative outcomes. For example, as described above, Microsoft's AI chatbot Tay turned into a racist, violence-glorifying, and sexist tweeter.[8] While the chatbot started out very nicely with tweets like "@mazanak_jee can I just say that I'm super stoked to meet u? humans are super cool," it became a Nazi in less than 24 hours and tweeted statements like "@brightonus33 Hitler was right I hate the jews." The chatbot was originally deployed to improve Microsoft's understanding of conversational language, especially language used by teenagers. However, AI-based technologies can be hijacked to perpetuate social biases and undesirable attitudes. Thus, companies must think about developing mechanisms to constantly monitor

and assess these technologies for biases as they interact with humans.

Refining CCSs

On a regular basis, CCSs need to be refined and upgraded, in a similar way to how human employees need training and development. It is important to remember that CCSs are living systems, so unless they are kept current, they will become obsolete (much like human workers), resulting in poor performance. Furthermore, in cases of regulatory or political changes to laws and rules, it may be necessary to rework underlying learning algorithms or to feed the system a new set of parameters.

CCSs are becoming ubiquitous. However, despite our heavy reliance on these tools, very few people know how these algorithms actually work or the logic behind them.

Recommendation 8: Bolster Your Cybersecurity Capacity

Cybersecurity has never been more important. The rise of CCSs will only intensify the pressure on companies to have robust cybersecurity mechanisms. Active monitoring of systems and proactive actions to thwart attacks will become even more important. CCSs that operate in autonomous mode can be hijacked with devastating outcomes. Given the fact that these systems will be operating with a network of other systems, a compromise in one system can easily trigger cascading failures. For example, in 2013, the Associated Press's Twitter account was hacked and misinformation regarding an attack on the White House and President Obama being injured was tweeted.[9] In the three minutes it took for people to realize that the tweet was fake, US $136 billion in value was wiped out from the US equity market, due to actions by humans on the trading floor who believed the information was valid. Imagine the chaos that would result from a compromise to a system of systems set up with information flows moving between autonomous systems. As CCSs are refined and improved over time, it is vital that organizations ensure that security protocols are bolstered on a regular basis.

Recommendation 9: Conduct Regular Audits

CCSs are becoming ubiquitous in our society. However, despite our heavy reliance on these tools, very few people know how these algorithms actually work or the logic behind them. A critical concern with CCSs are the algorithm biases that might arise due to the data they ingest over time, which impacts their learning functions. In addition, changes in the operating environment may render the existing algorithm outdated. Thus, CCSs must be audited on a regular basis to avoid embarrassing outcomes.

For example, a study ran an algorithm on Google News articles to test for biases in word embedding.[10] The algorithm associated "man is to computer programmer as woman is to homemaker." Another study found that Google displayed coaching ads for high-paying executive jobs far less often to women than it did to men.[11] Algorithms pick up these biases and stereotypes from their creators and the data they ingest. In a third example, Google's object detection and recognition algorithm mistakenly tagged black people as gorillas.[12] In all three cases, the CCS learned a bias and acted upon it. Periodic auditing could have caught these biases earlier. Particular attention should be paid to the auditing of CCSs when they are being scaled or being deployed in new environments.

Conclusion

With all nine recommendations for designing, developing, deploying, and refining CCSs, we note that, in addition to these CCS-centric recommendations, good systems development practices still need to exist.

There is little doubt that CCSs will be the dominant technology for the foreseeable future. It is incumbent on business professionals to understand the strengths and the weaknesses of CCSs and to implement them with their eyes wide open about both. At present, successfully developing and implementing a CCS likely represents a significant competitive advantage for firms that can do it and do it well, but at some point CCSs will become a necessary part of all business operations.

Acknowledgments

This paper draws heavily from Kevin Desouza's talks at the 2016 *Workshop on Artificial Intelligence, Cognitive Computing, and Autonomous Applications* held in Dublin, Ireland, and the 2016 *BIG Ideas Conference* in Dublin, Ohio, USA. We are indebted to the comments received during these presentations. Partial support for this project is provided by the IBM Center for the Business of Government.

Endnotes

[1]Desouza, Kevin C. "Delivering Artificial Intelligence in Government: Challenges and Opportunities." IBM Center for the Business of Government, 2018 (http://www.businessofgovernment.org/sites/default/files/Delivering%20Artificial%20Intelligence%20in%20Government.pdf).

[2]Chun, Andy Hon Wai. "Using AI for e-Government Automatic Assessment of Immigration Application Forms." *Proceedings of the 19th National Conference on Innovative Applications of Artificial Intelligence*, Volume 2, AAAI Press, 2007 (https://www.aaai.org/Papers/AAAI/2007/AAAI07-273.pdf).

[3]"Whole of Government 'Ask Jamie' Virtual Assistant." GovTech Singapore (https://www.tech.gov.sg/Programmes-Partnerships/Programmes-Partnerships/Initiatives/Virtual-Assistant).

[4]Tan, Lorna. "Filing Tax Returns Less Taxing Now." *The Straits Times*, 27 March 2016 (www.straitstimes.com/business/invest/filing-tax-returns-less-taxing-now).

[5]"Facebook Blocks UK Insurance Company from Using the Social Media Network's Data." Beat, 20 December 2017 (www.beat102103.com/tech/facebook-blocks-uk-insurance-company-using-social-media-networks-data).

[6]Grant, Katie. "Random Darknet Shopper: Exhibition Featuring Automated Dark Web Purchases Opens in London." *The Independent*, 11 December 2015 (www.independent.co.uk/life-style/gadgets-and-tech/news/random-darknet-shopper-exhibition-featuring-automated-dark-web-purchases-opens-in-london-a6770316.html).

[7]Vincent, James. "Twitter Taught Microsoft's AI Chatbot to be a Racist Asshole in Less than a Day." *The Verge*, 24 March 2016 (www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist).

[8]Vincent (see 7).

[9]Jackson, David. "AP Twitter Feed Hacked; No Attack at White House." *USA Today*, 23 April 2013 (https://www.usatoday.com/story/theoval/2013/04/23/obama-carney-associated-press-hack-white-house/2106757).

[10]Bolukbasi, Tolga, et al. "Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings." *Advances in Neural Information Processing Systems*, 2016 (https://arxiv.org/abs/1607.06520).

[11]Gibbs, Samuel. "Women Less Likely to be Shown Ads for High-Paid Jobs on Google, Study Shows." *The Guardian*, 8 July 2015 (www.theguardian.com/technology/2015/jul/08/women-less-likely-ads-high-paid-jobs-google-study).

[12]Mulshine, Molly. "A Major Flaw in Google's Algorithm Allegedly Tagged Two Black People's Faces with the Word 'Gorillas.' " *Business Insider*, 2 July 2015 (https://www.businessinsider.com.au/google-tags-black-people-as-gorillas-2015-7).

Kevin C. Desouza is a Foundation Professor in the School of Public Affairs at Arizona State University, where he served as the inaugural Associate Dean for Research for the College of Public Service & Community Solutions. Dr. Desouza is also a Nonresident Senior Fellow at the Brookings Institution. He has authored, coauthored, and/or edited nine books and has published more than 130 articles in journals across a range of disciplines, including software engineering, information science, public administration, political science, technology management, and urban affairs. Dr. Desouza's work has been featured in Sloan Management Review, Harvard Business Review, Stanford Social Innovation Research, Bloomberg Businessweek, The Wall Street Journal, Government Executive, *and* Computerworld, *among others. He has advised, briefed, and/or consulted for major international corporations, non-governmental organizations, and public agencies on strategic management issues ranging from management of information systems, to knowledge management, competitive intelligence, government intelligence operations, and crisis management. Dr. Desouza has received more than US $1.7 million in research funding from both private and government organizations. He can be reached at kevin.c.desouza@gmail.com.*

Lena Waizenegger is a Lecturer in business information systems at Auckland University of Technology, New Zealand. Her main research areas are ubiquitous connectivity, artificial intelligence, and human-robot interaction, and she has been published in various peer-reviewed outlets. Previously, Dr. Waizenegger was a project manager of an international research project at the University of Auckland, New Zealand. She was awarded the renowned Marietta Blau Scholarship from the Austrian Federal Ministry of Science, Research and Economy and has participated in two renowned doctoral consortia. Dr. Waizenegger received her PhD in information systems from the University of Innsbruck, Austria. She can be reached at lena.waizenegger@aut.ac.nz.

Gregory S. Dawson is a Clinical Associate Professor at the W.P. Carey School of Business at Arizona State University (ASU) and a Research Fellow at the Center for Organization Research and Design within ASU's College of Public Service & Community Solutions. Prior to working in academia, he was a Partner in the Government Consulting Practice at PricewaterhouseCoopers, focusing on public sector outsourcing and information systems implementation, and a Director at Gartner, working in the state and local government area. Dr. Dawson's research has been published in numerous academic and practitioner outlets, and he has won several teaching awards at the graduate and undergraduate teaching levels. He received his PhD in information systems from the Terry College of Business at the University of Georgia. He can be reached at GregorySDawson@gmail.com.

Transforming Banks Through AI

by Hemamalini Kumaran, Prema Sankaran, and Raj Gururajan

Artificial intelligence (AI) is a blend of machine learning (ML), natural language processing (NLP), and cognitive computing that enables machines to simulate the intelligence of humans. It is already being used in the telecom, automotive, financial services, retail, entertainment, education, healthcare, and travel industries. AI impacts our everyday lives from Siri on our phone, to Cortana or Google Assistant on our computer, to Netflix on our smart TV for personalized recommendations.

The banking and retail industries reportedly spend the most on cognitive and AI systems.[1] Artificial intelligence has been widely used with chatbots and in such areas as anti–money laundering (AML) and fraud detection. Voice technology, through voice services such as Amazon Alexa, is becoming the future of digital interactions to, for example, check an account balance, hear payment due date, get account transaction history, and make payments.[2] The banking and financial services industry is one of the top users of AI to increase productivity and customer satisfaction. Figure 1 shows industries with high, medium, and low adoption of AI.[3]

AI in the Banking Sector

Artificial intelligence has become a significant enabler in the banking and financial services due to the importance of making decisions faster and quicker.

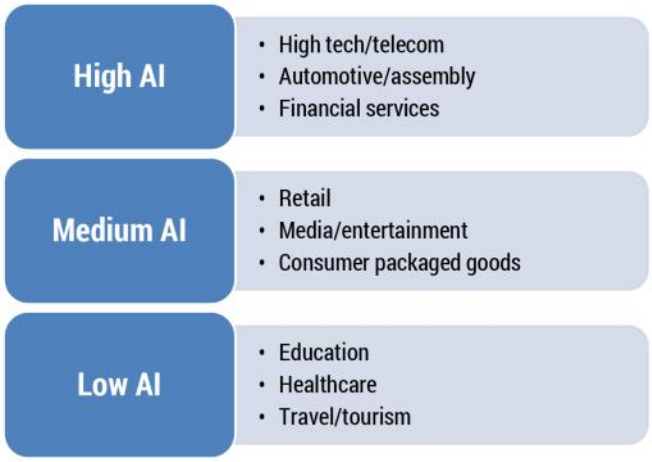

Figure 1 — Digital adopters of AI.
(Source: McKinsey Global Institute.)

AI technologies are being applied in the areas of knowledge management, market analysis, customer relationships, fraud reduction, and risk control. With faster computational power and big data analytics, computers using efficient and powerful AI algorithms can efficiently apply recommendation engines in the financial sector. The industry can now work with huge amounts of historical data for decision-making processes. Leveraging AI in the banking industry will result in a significant shift from a heavy reliance on human interactions to digitally managed customer service.

Technology-driven banks compete by launching AI solutions to provide faster services. Indeed, the digital transformation budget is spent on software, internal IT, cybersecurity, innovation on IT systems, and training. With an increasing percentage of the budget devoted to AI innovation, banks' AI capabilities are growing significantly across all business functions and processes. Table 1 shows some of the AI applications banks are using today in different functional areas.

Certainly, AI will revolutionize banking operations and the way banks interact with customers. According to customer experience provider Servion, artificial intelligence will power 95% of customer interactions by 2025.[4] By allowing customers to use voice-banking technology on an app, AI dramatically improves customer satisfaction. Moreover, AI applications monitor the online banking behavior of customers through deep learning algorithms for potential fraud detection. In addition, AI-powered chatbots serve as personal banking assistants, monitoring customers' financial habits and providing personalized recommendations. These growing advantages have increased AI's potential in the banking sector to provide reliable data for effective decision making.

The growing interest and increasing initiatives to introduce AI in banks reflect the need in today's business environment to stay competitive and enhance customer satisfaction. In this article, we describe the ways that banks have deployed artificial intelligence and the benefits of the AI applications that have been introduced in the banking sector. We also highlight a

Area/Application	Purpose	Used by
Chatbot	Answer customer queries	HDFC Bank
Robotic assistant	Guide customers through robotic help desk	HDFC Bank
Real-time feedback	Capture feedback based on facial expressions	SBI
Intelligent assistant	Provide instant answers to queries	SBI
Software robotics	Perform repetitive task	ICICI Bank
Natural language processing	Answer FAQs	Axis Bank
Robotic process automation	Manage processes	Axis Bank
Contract intelligence	Document analysis	JPMorgan Chase & Co.
Chatbot through Facebook Messenger	Answer customer queries through social media	Wells Fargo
Intelligent virtual assistant	Provide predictive analytics and cognitive messaging	Bank of America
Personal finance app	Save, invest, and track finances	Citibank

Table 1 — AI applications in banks.

few AI applications that Indian and US banks have implemented to enhance customer experience. Finally, we examine the key drivers behind the adoption of AI at banks and the challenges involved in implementing artificial intelligence.

AI Banking Initiatives in India

According to AI market researcher TechEmergence, the impact of AI on Indian banks is in the areas of real-time feedback solutions, chat assistants, robotic process automation (RPA), and NLP apps.[5] The following sections highlight some of the AI applications that Indian banks use.

HDFC Bank's Chatbot "Eva" to Address Customer Queries

Indian banking firm Housing Development Finance Corporation (HDFC) has launched AI-based chatbot "Eva" to handle customer queries.[6] This chatbot has addressed more than 2.7 million customer queries and has interacted with more than 530,000 unique users since March 2017. The electronic chatbot can assimilate knowledge from thousands of sources and provide answers within seconds. With the help of this chatbot, customers can receive product and services information quickly. In addition, HDFC Bank is experimenting with an Intelligent Robotic Assistant (IRA). The bank hopes to use the IRA to create a robotic help desk to guide customers to relevant counters.[7] The main advantage of using robotic assistance is a reduction in repetitive work for employees.

SBI Intelligent Assistant for Real-Time Feedback Solutions

The State Bank of India (SBI) uses AI-based real-time feedback solutions on customer satisfaction to gather feedback from branch-installed cameras that capture customer facial expressions.[8] This solution provides immediate feedback on whether the customer is satisfied. This helps gauge the effectiveness of the service provided based on the feedback captured. For ethical consideration, SBI has sent a disclaimer to customers to explain that this feedback is captured for internal training and quality purpose only. In addition,

SBI has launched the SBI Intelligent Assistant (SIA), an AI-powered chat assistant that provides instant answers to customer queries.[9] This chatbot has responded to millions of queries and can respond to 10,000 inquiries per second.

ICICI Bank's Software Robotics and Chatbot "iPal"

Software robotics facilitate human-robot interactions to automate and perform repetitive, high-volume, and time-consuming business tasks at the Industrial Credit and Investment Corporation of India (ICICI) Bank. These robots can execute data entry and validation, automated formatting, text mining, workflow acceleration, reconciliation, and currency exchange rate processing. In addition, ICICI has launched an AI-based chatbot ("iPal"), which has interacted with 3.1 million customers and has answered 6 million queries with a 90% accuracy rate.[10] The chatbot helps distinguish between customer (i.e., those that purchase a product for individual consumption) email and distributor (i.e., those that purchase a business pack to sell to other independent distributors or customers) email, thereby reducing the time it takes to process the status of transactions.

Axis Bank's NLP-Enabled App for FAQs

Axis Bank has launched an AI and NLP-enabled app to answer frequently asked questions. The bank has implemented AI across 125-plus processes and cognitive automation for 90 processes, significantly reducing response time to customer queries.[11] The bank uses RPA for such processes as account maintenance and servicing, loan disbursements, bulk transactions, and ATM support to reduce turnaround time. RPA reduces cost, minimizes human intervention in the execution of tasks, and increases efficiency.

AI Banking Initiatives in the US

In research on banks in the US, TechEmergence reports that chatbots and conversational interfaces are trending.[12] The impact of artificial intelligence on US banks is primarily in process automation, in the areas of document analysis, chatbots, virtual assistants, and customer service apps. Next, we discuss some of the applications in use by US banks.

JPMorgan Chase's COIN for Document Analysis

The COIN (contract intelligence) platform introduced by JPMorgan Chase is used to analyze legal documents and extract important data points and clauses within seconds.[13] In analyzing legal documents, AI can search through relevant data and analyze contracts for errors, missing information, and inconsistent language. Before COIN, a manual review required 360,000 hours; the same review now takes only seconds. The COIN platform automates the reading and interpreting of commercial loan agreements, which, done manually, is a time-consuming task. With the automation of review, research, and quality control, there is more time savings, a reduction in loan servicing mistakes from human error, and an increased focus on building customer relationships.

Wells Fargo's Chatbot for Facebook Messenger

Wells Fargo launched an AI-driven chat messenger through Facebook to support its customers. As most of its customers are familiar with Facebook and already users of messenger services, Wells Fargo's initiative facilitates huge time savings for busy customers.[14] The bank's effort to support customers through desktop, smartphone, or mobile devices will enable the bank to reach customers through social media. AI-based innovation helps the bank deliver an outstanding customer experience, build trust, and provide information to make financial decisions.

Bank of America's Intelligent Virtual Assistant "Erica"

Bank of America recently launched an intelligent virtual assistant ("Erica") to provide predictive analytics and cognitive messaging.[15] Through this assistant, customers can perform day-to-day transactions 24/7. The chat assistant also provides recommendations to help customers reach their financial goals. Moreover, Bank of America has leveraged mobile banking by making its intelligent virtual assistant easily accessible to clients.

Citibank's Clarity Money and Feedzai

CitiBank, through Citi Ventures, has made investments in AI, including Clarity Money.[16] Clarity Money has analyzed more than $10 billion in customers transactions, resulting in a savings of $300 per customer.

Through Citi Ventures, Citibank has also made strategic investments in Feedzai technology for detection of malicious threats.[17] This technology uses ML and AI to detect threats and help customers make data-backed decisions.

Benefits of Redefining Banking Through AI

AI has become all-pervasive as banks begin to realize the potential of the technology in streamlining operations, improving offerings, and enhancing the customer experience. From some of the applications we examined, we can observe that AI applications in banking and financial services are primarily centered on robotic advice, fraud detection, document analysis, customer recommendations, and chatbots, among others (see Figure 2). AI applications provide benefits in banking functions in the following ways:

- **Understanding and interpreting customer behavioral patterns.**[18] AI can be used to understand customer behavior and know customer preferences at a personal level to engage in deeper interactions, which in turn leads to accurate predictions of future behavior based on past preferences. AI can offer personalized recommendations to customers based on customer activity and buying patterns.
- **Approving credit applications (either for new or existing customers).**[19] AI provides lenders the benefit of examining and scoring a customer's creditworthiness. This technology can examine the interactions among multiple variables and make accurate predictions for credit risk and approve credit applications to the appropriate people. Banking and financial services lenders can reduce their lending risk by making accurate decisions using statistical and ML techniques. The use of these techniques not only enables the correct lending decision but also facilitates fast and efficient decisions. The main disadvantage of using AI in credit approval is that the analytical process will be hidden from the consumers and does not provide room for transparency.
- **Tracking customer purchases for any fraudulent activity**.[20] Fraud detection is one of the fields that has received a massive boost in providing accurate and superior results with the intervention of artificial intelligence. In any transaction, customer purchases are regularly tracked by software to detect any fraudulent transaction pattern. For example, if a customer has accessed their account from more than seven different IP addresses within a span of a week, this will raise an alarm. However, this may be a false alarm if, for example, the customer is a frequent traveler and has accessed his account from different

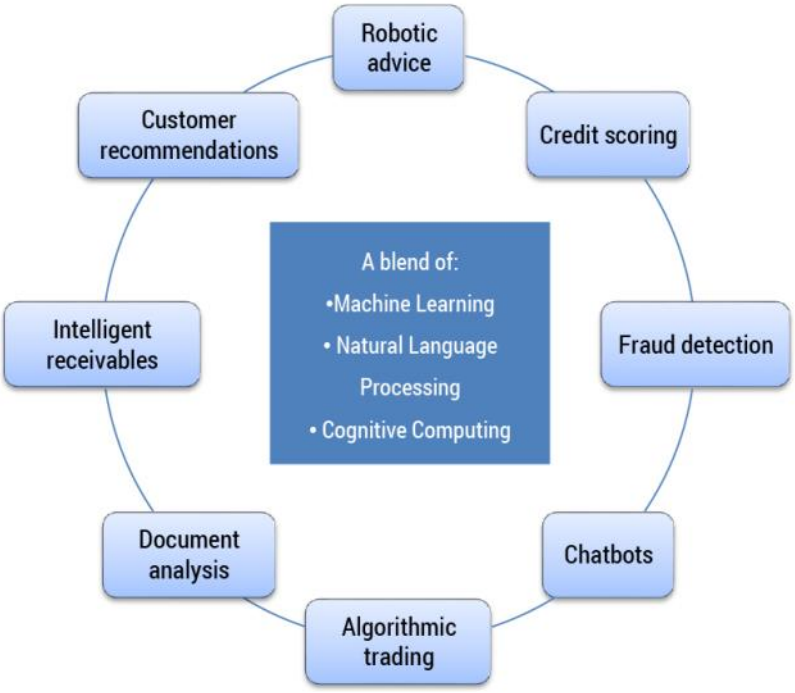

Figure 2 – Key areas of AI application in banking and financial services.

IPs. To overcome this problem, banks have launched AI applications to analyze spending patterns closely.

- **Searching customer transactions for any money-laundering activity**.[21] Banks are currently shifting to AI to identify money-laundering patterns. The use of artificial intelligence to detect money laundering goes beyond the abilities offered by rules-based systems. AI and ML impact the Know Your Customer (KYC) compliance process, which is typically used by banks to verify the identity of clients and regulate anti–money laundering activities. Machine learning, in particular, makes use of real-time, transaction-based KYC to identify peculiar transactions and to generate a suspicious activity report (SAR). In addition, AI facilitates more powerful customer segmentation. For example, AI takes customers' banking transactions in aggregate and generates "archetypes" of customer behavior. Each customer is a mixture of these archetypes and by using clustering techniques based on these archetypes, the technology forms customer clusters. Customers that do not fall within any of these formed clusters are considered suspicious. Thus, AI can be leveraged to redefine fraud management through real-time transaction data analysis, deeper understanding of the parties involved, and automatic generation of SARs.

Most banks in the US have not yet adopted AI for reasons such as operations, regulations, and budget or resource constraints.

- **Forecasting customer collections.**[22] AI-assisted forecasting helps identify when a customer will default on payment so that proactive steps can be applied to recover the amount already lent to the customer and to take action to suppress any future disbursement. Past customer behavior provides reliable information to predict future behavior.

- **Processing receivables intelligently.** AI capabilities also extend to the area of corporate processing of accounts receivable. With its innovative Intelligent Receivables, Bank of America brings together AI, ML, and optical character recognition to overcome reconciliation problems caused by incomplete remittance information.[23] Intelligent Receivables is helpful in achieving greater efficiency and cost savings in managing accounts receivable.

These are only a few of the AI applications currently used in banks. AI technologies are not limited to customer service; they have been extended to sales and marketing, fraud management, and financial advisory to provide wealth management decisions. From the applications discussed, it is clear that AI technologies can be successfully incorporated into several banking functions to great advantage. In fact, developments in AI provide the opportunity to completely reengineer banking processes in the customer service, sales, risk management, financial assistance, and reconciliation areas.

AI Key Drivers and Implementation Challenges

In a survey by AI tech company Narrative Science, 32% of the firms surveyed confirmed using AI technologies.[24] However, according to PricewaterhouseCoopers, most banks in the US have not yet adopted AI for reasons such as operations (e.g., IT operations, managing complex data), regulations, and budget or resource constraints.[25] Adoption of AI in the banking sector is influenced by four business and technological drivers, which face the following challenges (see Figure 3):

1. **Transparency.** Regulatory measures demand a greater awareness of risk that banks manage, and the effectiveness of AI controls is important to mitigate risk. It is a challenge to manage the entire financial system as a single optimization problem because the system is complex. While AI excels at managing known risk, it can also augment the risks that culminate in unknown risk, such as a financial crisis. For instance, letting robotic software rather than a human decide who can obtain a loan could cause financial risk if incorrect decisions are made and the steps involved in making the decisions are not made transparent. The banking and financial services sector operates in a highly regulated environment. AI software uses hidden analytical processes to make decisions. This may not be acceptable to regulators. To overcome this hurdle, a level of transparency is encouraged to allow users to audit the decisions made by the intelligent systems. Transparency can be achieved through AI technologies such as natural language generation that transforms structured data into explanatory language. In addition, deciding what technology

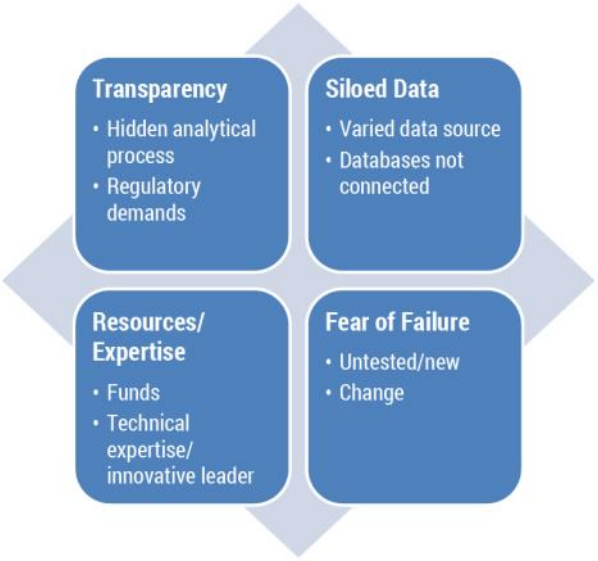

Figure 3 — Four key drivers and associated challenges in AI implementation.

to use is a challenge in AI implementation, as AI involves multidisciplinary capabilities related to ML, deep learning, cognitive computing, NLP, and robotics.

2. **Siloed data.** The difficulty in handling data variety in multiple data silos is a challenge faced by some of the organizations implementing AI. The data repositories are siloed, and there is a growing need to integrate data across the organization to make data consistent and easily available. Siloed data exists when departments within an organization are not sharing information with other departments. This becomes a major issue in AI adoption, since AI relies primarily on data for analytics. Data silos are criticized for lack of data integrity because the content may differ from department to department in an organization. To resolve this, banks need to create more consistent data in terms of format, data types, and sources. Many organizations are moving to cloud-based backup and archiving solutions to ensure data integration across all departments. This approach unifies the data within the organization, while also making information a strategic asset and a key differentiator among competitors.

3. **Resources/expertise.** AI implementation in banks is challenged by lack of technical expertise, no clear internal ownership of testing emerging technologies, and low availability of funds for innovation. However, funds are not a major constraint for larger banks. Nevertheless, banks need to assess the level of AI requirement needed for their operations and evaluate whether they are technology-enabled to build it or buy it. For example, if a bank has several repetitive processes, there is an opportunity to automate through AI solutions. The next major decision would be to identify whether the bank wants to build internally or procure the AI solution. If the bank decides to build in-house, it is vital to consider internal skills, technology readiness, and the deadline to achieve the goal. On the other hand, if the bank relies on external partners to provide AI solutions, it should evaluate the breadth and depth of offerings, technical expertise, and the delivery model. The challenge of no clear internal ownership of testing emerging technologies is another aspect of resource or expertise constraint. The Narrative Science survey reported that only 6% of those surveyed had a dedicated innovative leader to test new ideas or processes.[26]

4. **Fear of failure.** This challenge may be due to the uncertainty involved in implementing new solutions as banks may not be sure about the security issues involved in handling data. There is also a fear of losing jobs because of lack of digital skills and fear of change. Employees' fear of change is a major concern in the adoption of artificial intelligence (in any industry). AI is perceived to replace human resources because most processes are automated. On the contrary, AI can be effectively used to perform monotonous tasks while banking personnel focus on more complex tasks.

Conclusion

Artificial intelligence is rapidly changing how banks and financial institutions are operated and regulated. Some functions of banks and financial institutions are naturally well-suited for AI, such as risk management and day-to-day financial supervision, where AI provides cost savings and increased operational efficiency. Indeed, increased use of artificial intelligence in banks may result in exceptional management of day-to-day risk, via more logical rules and automatic compliance, yielding much lower costs. The usefulness of AI comes from its ability to find data representations that are naturally nontransparent to the human mind.

Banks that want to stay competitive need to be on the lookout for new applications, technologies, and innovations to deliver the best customer experience.

The application of AI has grown in recent years as banks strive to stay competitive. Presently, most AI initiatives undertaken by banks are primarily centered on customer insight, customer support, risk management, and operations. Accelerated interest is the result of increased growth in data, technologies, competition, regulation, and customer expectations. Overall, artificial intelligence is playing a greater role in digital transformation and will accelerate retail banking functions to the next level in terms of customer satisfaction and efficiency. AI impacts many industries and is expected to grow exponentially in banking.

Although there is increased interest in AI in the banking sector, most banks are still explorers with AI-related initiatives. Banks must update their technological infrastructure while at the same time embark on a cultural shift to use data in their day-to-day business. Thus, banks need to develop strategic approaches to take full advantage of this new opportunity. Banks should begin to integrate and automate all processes across major areas, including mortgage, customer service, wealth management, settlements, risk management, dispute management, and anti–money laundering without neglecting regulatory compliances. As banking is one of the world's most data-intensive industries, the capability to analyze all data will be of growing importance. If AI automation in banks can achieve increased compliance, reduced complaints, and improved customer service, it will bring major benefits to banking operations.

Banks that want to stay competitive need to be on the lookout for new applications, technologies, and innovations to deliver the best customer experience. Those that embrace AI technologies will be winners over the next 10 years, as this technology has the remarkable potential to augment human efforts in handling banking operations. With AI, the banking industry will reinvent several banking processes to great advantage and gain productive results in the front, middle, and back office.

Endnotes

[1]"IDC Spending Guide Forecasts Worldwide Spending on Cognitive and Artificial Intelligence Systems to Reach $57.6 Billion in 2021." IDC, 25 September 2017 (https://www.idc.com/getdoc.jsp?containerId=prUS43095417).

[2] "Amazon's Alexa Speaks to US Bank Customers." FinTech Futures, 7 September 2017 (http://www.bankingtech.com/2017/09/amazons-alexa-speaks-to-us-bank-customers).

[3]Bughin, Jacques, et al. "Artificial Intelligence: The Next Digital Frontier?" McKinsey Global Institute, June 2017 (https://goo.gl/R1BEmP).

[4]"AI Will Power 95% of Customer Interactions by 2025." Servion, 14 March 2017 (http://servion.com/blog/ai-will-power-95-customer-interactions-2025).

[5]Baruah, Ayushman. "AI Applications in the Top 4 Indian Banks." TechEmergence, 28 December 2017 (https://www.techemergence.com/ai-applications-in-the-top-4-indian-banks).

[6]"HDFC Bank OnChat." HDFC Bank (https://www.hdfcbank.com/htdocs/onChat/index.html).

[7]Jha, Sneha. "HDFC Bank Bets Big on AI with Robotic Assistant IRA." ETCIO.com, 10 January 2017 (https://cio.economictimes.indiatimes.com/news/strategy-and-management/hdfc-bank-bets-big-on-ai-with-robotic-assistant-ira/56421444).

[8]Baruah (see 5).

[9]Shetty, Mayur. "SBI Deploys AI-Based Financial Solution SIA." *The Times of India*, 26 September 2017 (https://timesofindia.indiatimes.com/business/india-business/sbi-deploys-ai-based-financial-solution-sia/articleshow/60835330.cms).

[10]Sawant, Nimish. "ICICI Bank's AI-Powered Chatbot iPal Completes 6 Million Responses, Interacts with 3.1 Million Customers in Eight Months." Firstpost, 17 October 2017 (http://www.firstpost.com/tech/news-analysis/icici-banks-ai-powered-chatbot-ipal-completes-6-million-responses-interacts-with-3-1-million-customers-in-eight-months-4151269.html).

[11]Baruah (see 5).

[12]Sennaar, Kumba. "AI in Banking — An Analysis of America's 7 Top Banks." TechEmergence, 24 October 2017 (https://www.techemergence.com/ai-in-banking-analysis).

[13]Galeon, Dom, and Kristin Houser. "An AI Completed 360,000 Hours of Finance Work in Just Seconds." Futurism, 8 March 2017 (https://futurism.com/an-ai-completed-360000-hours-of-finance-work-in-just-seconds).

[14]"Wells Fargo Testing Bot for Messenger Featuring New Customer Service Experience." Press release, Wells Fargo, 18 April 2017 (https://newsroom.wf.com/press-release/community-banking-and-small-business/wells-fargo-testing-bot-messenger-featuring-new).

[15]"Bank of America Previews the Future of Banking and Payments Technology at Money20/20." Press release, Bank of America, 2016 (http://newsroom.bankofamerica.com/files/press_kit/additional/2016_Money2020_BAC_Announcement.pdf).

[16]"Citi Ventures Leads $11 Million Round in AI-Based Personal Finance Advocate Clarity Money." Finextra Research, 23 March 2017 (https://www.finextra.com/newsarticle/30328/citi-ventures-leads-11-million-round-in-ai-based-personal-finance-advocate-clarity-money).

[17]"Citi Ventures Makes Strategic Investment in Feedzai." Press release, Feedzai, 19 October 2016 (https://feedzai.com/press-releases/citi-ventures-makes-strategic-investment-feedzai).

[18]Mohammed, Umair. "How Artificial Intelligence Is Changing Customer Engagement and Experience." Entrepreneur India, 22 June 2016 (https://www.entrepreneur.com/article/277926).

[19]Crosman, Penny. "Is AI Making Credit Scores Better, or More Confusing?" American Banker, 14 February 2017 (https://www.americanbanker.com/news/is-ai-making-credit-scores-better-or-more-confusing).

[20]Sekar, Prasanna Kumar. "Potential of AI in Fraud Detection." *FINspeak*, Infosys, 18 September 2017 (http://www.infosysblogs.com/finspeak/2017/09/potential_of_ai_in_fraud_detec.html).

[21]Horan, T.J. "AI Meets AML: How Smart Analytics Fight Money Laundering." *FICO Blog*, 13 February 2017 (http://www.fico.com/en/blogs/analytics-optimization/ai-meets-aml-how-smart-analytics-fight-money-laundering).

[22]Hall, Owen P., Jr. "Artificial Intelligence Techniques Enhance Business Forecasts." *Graziadio Business Review*, Vol. 15, No. 2, 2002 (https://gbr.pepperdine.edu/2010/08/artificial-intelligence-techniques-enhance-business-forecasts).

[23]"Bank of America Applies AI to Streamline Client Payment Processing." Finextra Research, 23 August 2017 (https://www.finextra.com/newsarticle/30993/bank-of-america-applies-ai-to-streamline-client-payment-processing).

[24]"The Rise of AI in Financial Services." *Research Brief*, Narrative Science (https://narrativescience.com/Offers/The-Rise-of-AI-in-Financial-Services).

[25]Zhou, Adelyn. "Banks Eager for Artificial Intelligence, But Slow to Adopt." *Forbes*, 30 June 2017 (https://www.forbes.com/sites/adelynzhou/2017/05/29/financial-services-industry-banks-artificial-intelligence-slow-adoption/#2843c1cf597e).

[26]Narrative Science (see 24).

Hemamalini Kumaran is a Marketing and Communications Manager at A1PlusSoft Private Limited, India. She has more than 10 years' experience in technical IT communications. Dr. Kumaran has worked with clients in Switzerland documenting API guides and at Fidelity in Boston, Massachusetts, USA, for its HR management system. Her contributions to research are in the areas of consumer attitudes toward online shopping, market research, e-commerce, and product research. Dr. Kumaran received a PhD in management from Bharathiar University, India. She can be reached at hema8448@gmail.com.

Prema Sankaran is a Professor of Management at the Firebird Institute of Research, India. With a demonstrated history of working in the education management industry, she is highly skilled in health informatics, portfolio management, business development, administration, and stock market guidance. Dr. Sankaran received a PhD in management from Bharathiar University, India, focused on finance and financial services. She can be reached at premsank@gmail.com.

Raj Gururajan is a Professor of Information Systems at the School of Management and Enterprise, University of Southern Queensland, Australia. He has spent more than three decades in teaching and research, including 15 years of PhD supervisory experience. Dr. Gururajan currently teaches courses on global IT, big data visualization, and digital innovation. His research interests are in the areas of health information systems, nutrition and obesity studies, and qualitative research methodologies. Dr. Gururajan received a PhD in information systems from Edith Cowan University, Australia. He can be reached at Raj.Gururaja@usq.edu.au.

THE GENIE IS OUT

AI Enables Efficient and Effective Digital Government

by Vipin Jain and Seema Jain

> A computer would deserve to be called intelligent if it could deceive a human into believing that it was human.
>
> — Alan Turing

Artificial intelligence (AI) is quickly becoming more science fact than science fiction. And despite concerns about thinking machines taking human jobs — or one day outsmarting their makers — the US federal government's investment in AI technology is growing and could be applied to a wide range of functions, from healthcare to public safety and defense/intelligence applications to employee performance. Can cognitive technologies do government employees' thinking for them? Not yet. But they can augment employees' capabilities and free up billions of labor hours for more critical tasks — while delivering faster, better, and cheaper services.

The application of AI technologies can reduce backlogs, redesign work, cut costs, overcome resource constraints, enable more accurate predictions, and do myriad tasks not previously practical for humans to do on their own. Because of these potential benefits, the US government has invested in AI research for many years, and various committees and taskforces continue to investigate the technologies. Indeed, the opportunities that emerge from AI will help to transform the relationship between people, process, and technology (see Figure 1).

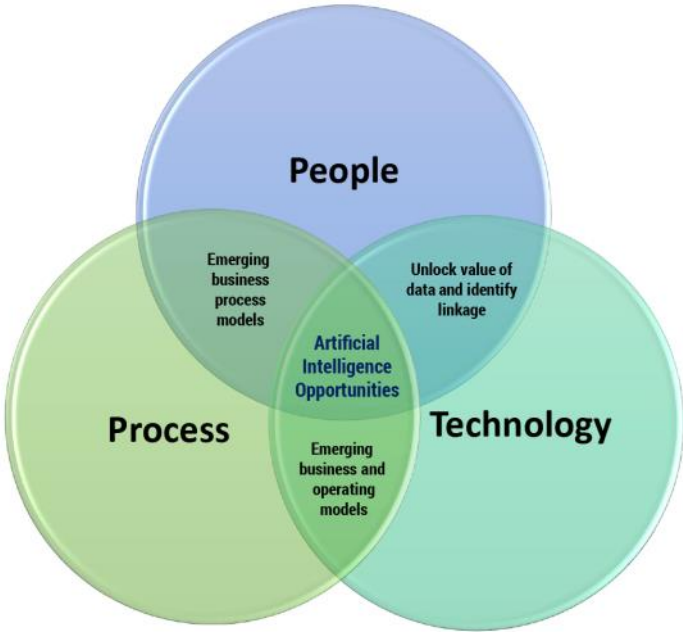

Figure 1 — AI opportunities.

AI in the US Government: Strategic Directions

The US National Artificial Intelligence Research and Development Strategic Plan[1] establishes a set of objectives for federally funded AI research — research occurring within the government as well as federally funded research occurring outside of government, such as in academia. The goal of this research is to produce new AI knowledge and technologies that provide a range of positive benefits to society, while minimizing the negative impacts. To achieve this goal, the Strategic Plan identifies seven priorities and makes the following recommendations for federally funded AI research:

1. **Make long-term investments in AI research.** Prioritize investments in the next generation of AI that will drive discovery and insight and enable the US to remain a world leader in AI.

2. **Develop effective methods for human-AI collaboration.** Rather than replace humans, most AI systems will collaborate with humans to achieve optimal performance. Research is needed to create effective interactions between humans and AI systems.

3. **Understand and address the ethical, legal, and societal implications of AI.** We expect AI technologies to behave according to the formal and informal norms to which we hold our fellow humans. Research is needed to understand the ethical, legal, and social implications of AI and to develop methods for designing AI systems that align with ethical, legal, and societal goals.

4. **Ensure the safety and security of AI systems.** Before AI systems are in widespread use, assurance is needed that the systems will operate safely and

securely, in a controlled, well-defined, and well-understood manner. Further progress is needed to address this challenge of creating AI systems that are reliable, dependable, and trustworthy.

5. **Develop shared public data sets and environments for AI training and testing.** The depth, quality, and accuracy of training data sets and resources significantly affect AI performance. Researchers need to develop high-quality data sets and environments and enable responsible access to high-quality data sets as well as to testing and training resources.

6. **Measure and evaluate AI technologies through standards and benchmarks.** Essential to advancements in AI are standards, benchmarks, testbeds, and community engagement that guide and evaluate progress in AI. Additional research is needed to develop a broad spectrum of evaluative techniques.

7. **Better understand the national AI R&D workforce needs.** Advances in AI will require a strong community of AI researchers. An improved understanding of current and future R&D workforce demands in AI is needed to help ensure that sufficient AI experts are available to address the strategic R&D areas outlined in this plan.

As a contribution toward preparing the US for a future in which AI plays a major role, the National Science and Technology Council's subcommittee on Machine Learning and Artificial Intelligence was chartered in May 2016. It is responsible for fostering interagency coordination, providing technical and policy advice on topics related to AI, and monitoring the development of AI technologies across industry and the research community. The US government's report, "Preparing for the Future of Artificial Intelligence,"[2] includes a comprehensive list of recommendations for further specific actions by federal agencies and other supporting organizations.

AI in the US Government: Focus Areas

> Strategic, widespread use of artificial intelligence could save government up to 1.2 billion work hours and US $41.1 billion annually.
>
> — IBM Center for the Business of Government[3]

AI is ready to digitally transform the business of government to an extent that we have not seen since the last industrial revolution. The key to successful implementation of AI lies in its being done with a defined scope and skilled workers. AI is not a single technology but rather a collection of multiple advanced technologies allowing machines to sense, understand, perform, and learn from experience continuously. If used strategically, AI offers tremendous opportunity to help government agencies reduce their costs, complement the current workforce's capabilities to focus on critical tasks while AI does some of their routine work faster, and deliver improved and timely services to citizens. Many federal, state, and local agencies have been experimenting with the use of AI, and that use is expected to further accelerate as the agencies become more comfortable with AI and realize its benefits. Following are some areas in government where AI is either currently used or being considered.

If used strategically, AI offers tremendous opportunity to help government agencies reduce their costs, complement the current workforce's capabilities, and deliver improved and timely services to citizens.

Healthcare

As one of the fastest-growing areas and with much of that growth fueled by analytics, medical imaging, and diagnostic capabilities, healthcare AI could and should be leveraged and considered for the following:

- Identifying possible spread of diseases early on and narrowing down the people and locations at risk
- Supporting and expediting FDA processes for drug testing and approval for consumption by citizens
- Identifying people at risk of post-traumatic stress disorder and other diseases, and providing them with timely treatments
- Analyzing healthcare data to identify possible violations of health codes at all levels

As an example of AI's use in healthcare, the US Department of Veteran Affairs is using AI to better predict medical complications and improve treatment of severe combat wounds at the Walter Reed National Military Medical Center, leading to better patient outcomes, faster healing, and lower costs.[4]

Social Media and Sentiments

From Facebook posts to sensor readings, we generate far too much data for humans to make sense of it in a timely manner. Cognitive technologies can help sift through data much faster and understand sentiments concerning national security, crime prevention, or national disaster, allowing authorities to make faster decisions.

Fraud Detection

There will likely always be a small group of people who will cheat the system for their personal gain. AI can be leveraged to identify those people and protect the rest of society from fraud. For example, AI could:

- Identify fraudulent Medicaid and Medicare claims, help catch culprits, and support reduced operating costs
- Identify and stop insider trading to avoid any loss of confidence in our financial system

Cognitive technologies can help sift through data much faster and understand sentiments concerning national security, crime prevention, or national disaster, allowing authorities to make faster decisions.

Electronic Document Discovery

In our current system, we generate enormous amounts of paper, especially within legal and healthcare segments. Without AI, it is very difficult, time-consuming, and resource-intensive to find relevant documents when needed. The proper use of AI and other relevant technology-based solutions will facilitate the following:

- Identifying relevant documents in the discovery phase of legal cases, reducing cost and time
- Expediting the FDA approval process for medicines

Performance Measurement

AI can help in monitoring and measuring whether new or upgraded policies or systems are producing the expected results. For example:

- Did a new childcare policy help reduce the potential harm done to children in foster care?
- Did a new capability reduce the time required to process a citizen request?

Citizen Experience and Engagement

AI has great potential for improving customer experience and engagement by:

- Understanding and sensing citizens' needs from data collected from social media, blogs, or feedback, and using it to help develop or update policies, regulations, and business solutions and addressing citizens' needs in a timely manner
- Developing AI-enabled service desks, including suggesting additional relevant services and resources to citizens who are applying for something such as Medicaid benefits, driving licenses, or business permits
- Having chatbots handle common service requests like password resetting, freeing workers to handle more complex requests and giving citizens quick answers to important questions — improving service while reducing costs and backlog

Public Safety

The public infrastructure is dated and requires ongoing maintenance. AI can help with the following:

- Understanding weather conditions and traffic patterns; predicting traffic congestion and accidents and identifying alternate routes
- Anticipating road, bridge, port, or other infrastructure maintenance and replacement needs
- Monitoring social media for quick notification of emergency situations (e.g., flood, crime, agitation)

Analytics and Prediction

Machine learning (ML) and natural language processing can reveal patterns, enabling better predictive capabilities. When your email program flags a message as spam, or your credit card company warns you of a potentially fraudulent use of your card, ML is probably involved. Some examples where technology is already being used in these areas include:

- The US Army is developing wearable monitors that use an ML algorithm to determine wound seriousness, helping medics prioritize treatment.[5]
- The US Department of Energy's self-learning weather and renewable forecasting technology uses ML, sensor information, and cloud motion physics derived from sky cameras and satellite observations to improve solar forecasting accuracy by 30%.[6]

Training

The US Defense Advanced Research Projects Agency (or DARPA), intending to reduce from years to months the time required for new US Navy recruits to become experts in technical skills, now sponsors the development of a digital tutor that uses AI to model the interaction between an expert and a novice.[7]

Robotic Process Automation

Robotic process automation will allow agencies to significantly reduce administrative tasks and have their employees maximize their time for mission-focused work. Some examples include:

- A Colorado survey found child-welfare caseworkers spending 37.5% of their time on documentation and administration, versus just 9% on actual contact with children and their families.[8]
- At the federal level, research indicates that simply documenting and recording information consumes a *half-billion* staff hours each year. "Bots" can automate such activities, ranging from invoice processing to filling in forms, from data entry to writing budget-reporting documents.[9]

By freeing up time, we can create a more effective government, empowering employees to do the work that really matters: serving citizens in need.

As these examples illustrate, cognitive technologies eventually will fundamentally change how government works, and the changes will likely come much sooner than many think. As cognitive technologies advance in power, government agencies will need to bring more creativity to workforce planning and work design. The most forward-leaning jurisdictions will see cognitive technologies as an opportunity to reimagine the nature of government work itself — making the most of complementary human and machine skills.

An AI Adoption Framework

AI presents government agencies with new opportunities to innovate that previously may have been impossible. Cognitive computing leverages several AI components, such as ML, to understand, reason, and learn much like the human brain does. Cognitive complements AI, enabling computers both to think like humans and to understand how humans think. Current AI systems and subsystems are able to learn, make decisions, and solve complex problems. Examples include speech recognition, self-driving cars, online assistants, and image recognition.

As cognitive technologies advance in power, government agencies will need to bring more creativity to workforce planning and work design.

As government agencies are realizing the need to start using AI to develop and deliver new business capabilities to meet their mandates, our recommendation is that they take an incremental approach to adopting AI. Irrespective of the maturity and experience of the AI team, the members should have a good understanding of the agency's overall digital strategy and, more specifically, its AI strategy if already in place. AI could be extremely helpful in improving the speed and efficiency of many tasks, but it is not suited for every task. The AI team should have a good understanding of the role of AI and the outcomes to be expected from its use in the workplace.

AI is not plug and play, and agencies cannot simply "buy intelligence" and apply it to address their requirements. Many AI services are available in the market, but the hard work of managing the interplay of planning, data, processes, and technologies happens and gets managed in-house. The conceptual framework for putting AI to work is quite simple. This framework can and should be customized based on an agency's environment and requirements.

The framework, as shown in Figure 2, has five steps to be executed from left to right, but this linear progression is not a requirement. For an organization new to AI, it will be best to go through all the steps, but a

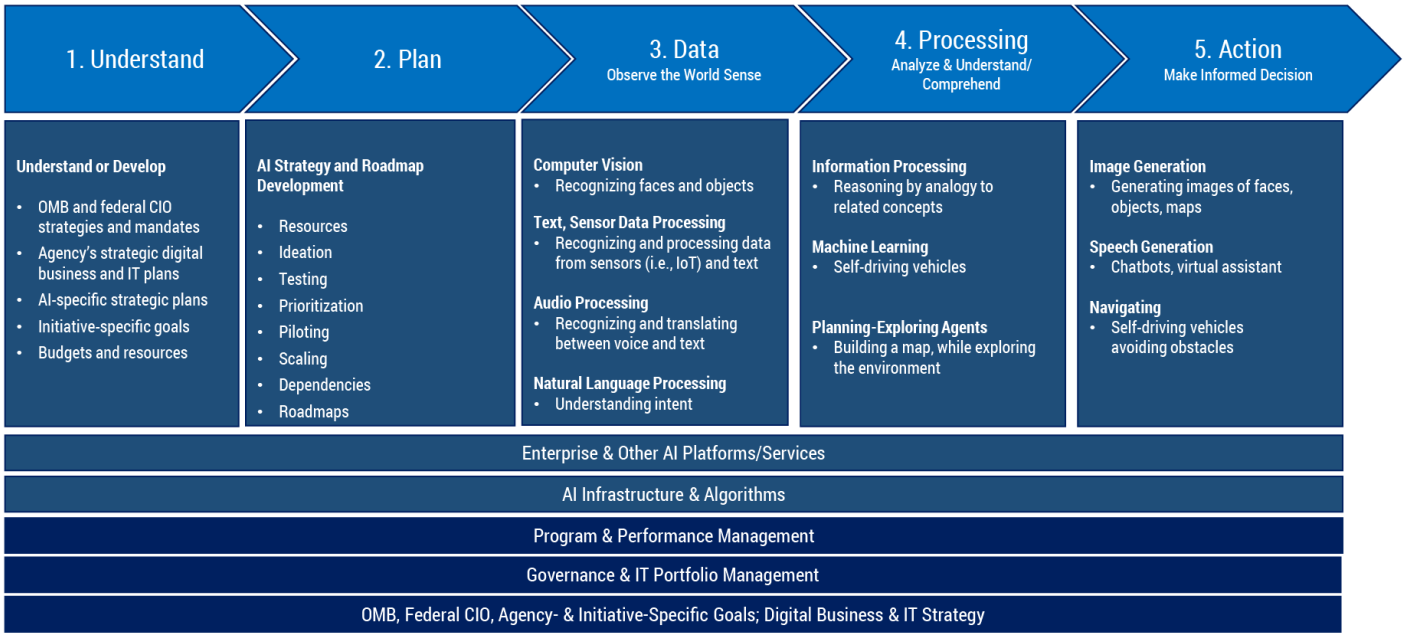

Figure 2 — A conceptual AI adoption framework.

mature organization may want to focus on specific areas based on its needs. AI algorithms consume data, process it, and then help with actions. Processing of the data depends on proper planning and integration of technologies.

AI Service Providers

Leading cloud service providers (CSPs), including Amazon Web Services, Azure, Google, and IBM, offer a wide array of AI and ML services that are quickly maturing. These services are expected to dominate enterprise AI in the near term and may also affect smaller startups. Most AI-branded services from the CSPs fall into one of four categories:

1. **Computer vision services.** These provide APIs that allow digital images or digital video to be passed to a pretrained ML algorithm for analysis. Each of these purpose-built services focuses on a different analysis use case.

2. **Language-processing services.** These provide the cloud services and APIs used for analyzing human language — either in text or in audio. In most cases, these services provide support for several languages. As with computer vision services, language-processing services and their pretrained ML models are segmented into distinct domains, with each service solving one or more aspects of the language-processing problem. For example, some services analyze and extract speech from digital recordings; others enable development of conversational interfaces, translation to and from a wide array of languages, and robust text-to-speech processing.

3. **ML services.** This category comprises general-purpose ML services. Unlike the previous two categories, which include task-specific, pretrained algorithms, ML services allow developers, data practitioners, and architects to generate their own models. ML services also allow users to train and evaluate those models using their own data.

   These services can perform the most common types of ML analysis, including classification, clustering, and regression. Among ML services, there will be a range of options, even from a single provider. Some of these ML services are built to be extremely simple to use. However, as is often the case, simplifying the solution might mean less flexibility or functional limitations.

4. **Miscellaneous services.** This category includes those services that do not necessarily fall into any of the other three categories.

AI Demand: Job Creator or Destroyer?

As government agencies and commercial organizations are starting to deploy AI within their environments, they are faced with a shortage of AI skills, resulting in

delays, increased cost, and, at times, failed initiatives. Demand for AI talent is far outpacing the availability of skilled resources worldwide. Industry research suggests that there are about 300,000 skilled AI workers available today, and while more are retooling to acquire the needed skills, current demand for AI workers runs into the millions, leaving a large gap.[10] Agencies requiring employees to have clearances could face an even wider gap.

Large commercial firms like Amazon, Baidu, and Microsoft are paying higher premiums to attract AI talent. This competition to hire the limited AI talent available will have an impact on the budget to hire and retain AI experts and should be considered as part of the strategic plan by governmental agencies and the service providers supporting them.

An important concern associated with AI is its impact on current jobs. Will it create or destroy jobs? How will job profiles evolve? And what types of skills will be in demand? The answers to these questions are critical to business leaders and policy makers as they seek to take full advantage of the opportunities arising from AI by ensuring that an appropriately skilled workforce is in place to capture them.

Looking at some past industrial revolutions and their impacts on jobs may provide some insight into what we might expect in the near future. Past revolutions have transformed and dramatically improved industrial productivity: steam power in the 19th century, electricity in the early 20th century, automation in the 1970s, and the Internet in the late 1990s. These waves of technological advancement did not reduce overall employment, however. Although the number of manufacturing jobs decreased, new jobs emerged and the demand for new skills grew.

Today, AI — along with a few other emerging technologies, such as blockchain, sensors, robotics, 3D printing, and others collectively known as part of Industry 4.0 — represents ongoing workforce transformation. These technologies are disrupting almost all industries and all kinds of jobs, be they in government, manufacturing, service, or white-collar sectors. We believe that, overall, the short-term impact of AI should be a positive one if we manage it well, even though jobs may require a better-educated and AI-aware workforce. In the long term, as AI becomes much stronger, we should be aware of the possible threats AI poses and begin to plan for them now.

Recent progress in AI is being driven by multiple complementary developments, including greater availability of data, cheaper and faster cloud-based computing power enabled by faster and more powerful graphics processing units, and the development of new algorithms. Venture capitalists are also making many investments in AI startups, and tech giants are opening up resources to enable others to develop better AI.

AI is accelerating the automation of white-collar jobs, and a growing wave of expert automation and augmentation software platforms are expected to usher in a new era of AI-enhanced productivity. Startups are targeting multiple industries, such as the legal field, journalism, wealth management, and healthcare, among others, which until now have been relatively safe from the threats of automation. AI-supported capabilities will allow for faster and better diagnosis or advice with fewer resources.

The short-term impact of AI should be a positive one if we manage it well, even though jobs may require a better-educated and AI-aware workforce.

Conclusion

> Artificial intelligence is the future, not only for Russia, but for all humankind. It comes with colossal opportunities, but also threats that are difficult to predict. Whoever becomes the leader in this sphere will become the ruler of the world.
>
> — Vladimir Putin

> I am in the camp that is concerned about artificial intelligence. First the machines will do a lot of jobs for us and not be super intelligent. That should be positive if we manage it well. A few decades after that, though, the intelligence is strong enough to be a concern. I agree with Elon Musk and some others on this and don't understand why some people are not concerned. Humans should be worried about the threat posed by artificial intelligence.
>
> — Bill Gates

The AI genie has already been released from the lamp, and any efforts to try to put it back or ignore it will be

futile and possibly harmful. AI is a disruptive technology with implications for humankind at all levels, be it as an individual citizen or independent business, or at the level of an industry, an individual country, or the whole world. The importance of AI-based startups, the pace of progress achieved in the past decade, and the adoption of AI by various organizations are ensuring that the capabilities and capacity of AI-based solutions will continue to accelerate.

> *AI has the potential to help citizens navigate government services, allow agencies to respond to threats more thoroughly, and improve agencies' overall effectiveness.*

AI has tremendous potential to transform our government and help us address many of our concerns and enhance our safety, education, health, and how various services are delivered. AI could and should be used to change the way public servants do their jobs. The AI field is advancing rapidly at a time when agencies are looking for new ways to achieve their missions. It can improve agencies' effectiveness, make data more understandable and easier to use, and help citizens navigate government services. And it could save government up to 1.2 billion work hours and $41.1 billion annually.[11]

According to Machine Intelligence Research Institute's Katja Grace and colleagues, researchers predict AI will outperform humans in many activities in the next 10 years, such as translating languages (by 2024), writing high school essays (by 2026), driving a truck (by 2027), working in retail (by 2031), writing a bestselling book (by 2049), and working as a surgeon (by 2053).[12] According to the same study, researchers also believe there is a 50% chance of AI outperforming humans in *all* tasks in 45 years and all human jobs will be automated in 120 years.

Even though we will not be here to witness the reality of all these predictions, developments such as Sophia, the first robot citizen of Saudi Arabia; Tesla's self-driving vehicles; robotic surgeries; or realistic conversational chatbots lead us to believe that AI will continue to be part of our daily lives and will make life easier.

Like many other disruptive technologies, in addition to all its promises, AI will have some downsides as well, including the risk of job elimination for those jobs that can be better performed by an AI solution. We are optimistic that even though some jobs may be lost due to AI, moving forward, AI will result in the creation of more, better-paying jobs and new job categories.

AI is here to stay, and so are the disruptions it will cause worldwide. Government agencies and businesses are just getting started in leveraging emerging technologies such as AI, blockchain, robotics, computer vision, automation, and so on, to provide enhanced citizen services, create new business models, and optimize business operations. Savvy leaders are not taking any chances and are facing the challenges head on. They are anticipating upcoming AI disruptions and working to convert them into strategic opportunities.

Although there will be challenges and detractors, AI has the potential to help citizens navigate government services, allow agencies to respond to threats and crises more thoroughly, and improve agencies' overall effectiveness. For some, AI promises to solve significant and long-standing problems, while for others AI presents a sense of peril. But all agree that AI will have a lasting impact on our lives.

Endnotes

[1]"The National Artificial Intelligence Research and Development Strategic Plan." US National Science and Technology Council, Networking and Information Technology Research and Development Subcommittee, October 2016 (www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf).

[2]"Preparing for the Future of Artificial Intelligence." Executive Office of the President, US National Science and Technology Council Committee on Technology, October 2016 (https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf).

[3]"The Future Has Begun: Using Artificial Intelligence to Transform Government." Partnership for Public Service, IBM Center for the Business of Government, 17 January 2018 (www.businessofgovernment.org/sites/default/files/Using%20Artificial%20Intelligence%20to%20Transform%20Government.pdf).

[4]Piechowski, Darcie. "Artificial Intelligence as a Driver for Innovation." IBM Center for the Business of Government, 2 February 2017 (www.businessofgovernment.org/blog/business-government/artificial-intelligence-driver-innovation).

[5]Ferdinando, Lisa. "Work: Exciting Technology Under Development." US Department of Defense, 4 March 2016 (www.defense.gov/News/Article/Article/685979/work-exciting-technology-under-development).

[6]"The SunShot Initiative." US Department of Energy, Office of Energy Efficiency & Renewable Energy (www.energy.gov/eere/solar/sunshot-initiative).

[7]Fletcher, J.D., and John E. Morrison. "Accelerating Development of Expertise: A Digital Tutor for Navy Technical Training." Institute for Defense Analyses, November 2014 (www.dtic.mil/get-tr-doc/pdf?AD=AD1002362).

[8]"Colorado Child Welfare County Workload Study." Colorado General Assembly, Department of Human Services, August 2014 (https://leg.colorado.gov/audits/colorado-child-welfare-workload-study).

[9]Eggers, William D., David Schatsky, and Peter Viechnicki. "How Artificial Intelligence Could Transform Government." Deloitte Insights, 26 April 2017 (www2.deloitte.com/insights/us/en/focus/cognitive-technologies/artificial-intelligence-government-summary.html).

[10]Vincent, James. "Tencent Says There Are Only 300,000 AI Engineers Worldwide, But Millions Are Needed." *The Verge*, 5 December 2017 (https://www.theverge.com/2017/12/5/16737224/global-ai-talent-shortfall-tencent-report).

[11]Eggers et al. (see 9).

[12]Grace, Katja, John Salvatier, Allan Dafoe, Baobao Zhang, and Owain Evans. "When Will AI Exceed Human Performance? Evidence from AI Experts." Cornell University Library, 30 May 2017 (https://arxiv.org/pdf/1705.08807.pdf).

Recommended Reading

"AI 100: The Artificial Intelligence Startups Redefining Industries." CB Insights, 12 December 2017 (www.cbinsights.com/research/artificial-intelligence-top-startups).

"Comparing Machine Learning as a Service: Amazon, Microsoft Azure, Google Cloud AI." AltexSoft, 14 December 2017 (www.altexsoft.com/blog/datascience/comparing-machine-learning-as-a-service-amazon-microsoft-azure-google-cloud-ai).

Davenport, Thomas H., and Rajeev Ronanki. "Artificial Intelligence for the Real World." *Harvard Business Review*, January–February 2018 (https://hbr.org/2018/01/artificial-intelligence-for-the-real-world).

"The Federal Big Data Research and Development Strategic Plan." US National Science and Technology Council, Big Data Senior Steering Group, Subcommittee on Networking and Information Technology Research and Development, May 2016 (www.nitrd.gov/PUBS/bigdatardstrategicplan.pdf).

"Federal Cybersecurity Research and Development Strategic Plan: Ensuring Prosperity and National Security." US National Science and Technology Council, Networking and Information Technology Research and Development Program, February 2016 (www.nitrd.gov/cybersecurity/publications/2016_Federal_Cybersecurity_Research_and_Development_Strategic_Plan.pdf).

Jha, Vishakha. "Amazon AWS Brings an Insight into Artificial Intelligence for Developers." TechLeer, 24 July 2017 (www.techleer.com/articles/213-amazon-aws-brings-an-insight-into-artificial-intelligence-for-developers).

Launchbury, John. "A DARPA Perspective on Artificial Intelligence." US Defense Advanced Research Projects Agency (www.darpa.mil/attachments/AIFull.pdf).

"Transforming Government Through Technology." IBM Center for the Business of Government, March 2017 (www.businessofgovernment.org/sites/default/files/Transforming%20Government%20Through%20Technology.pdf).

Zilis, Shivon, and James Cham. "The Current State of Machine Intelligence 3.0." ShivonZilis, 7 November 2016 (www.shivonzilis.com).

Vipin Jain is CTO for Optimal Technology Solutions, Inc. (OTSI), based in Chantilly, Virginia, USA. He leads OTSI's Digital Transformation Strategy and Architecture practice. Mr. Jain has over 30 years of experience in all aspects of technology management and solution delivery, including digital and IT strategy, enterprise and solution architecture, IT portfolio optimization and transformation, and deployment of global business solutions. Prior to joining OTSI, he held various executive-level positions with HPE, Microsoft, and Accenture, where he advised industry executives from Fortune *500 companies and US federal agencies delivering innovative business solutions. Mr. Jain also held executive-level positions within AIG, Merrill Lynch, Citibank, First Chicago, Empire Blue Cross Blue Shield, and other firms, managing technology across several enterprise-wide initiatives and performing due diligence of potential acquisitions. His breadth of experience includes work with large and small commercial companies, US federal and state government agencies, and management consulting firms. Mr. Jain has worked in multiple business domains, including banking, brokerage, insurance, healthcare, pharmaceuticals, energy, consulting, banking, and government. He is a member of industry groups focused on the public sector, digital and IT transformation, big data, and cloud. He can be reached at vipin@optimaltsi.com.*

Seema Jain is President and CEO for Optimal Technology Solutions, Inc. (OTSI), based in Chantilly, Virginia, USA. Ms. Jain has over 20 years of experience in all aspects of technology management and solution delivery, including digital and IT strategy, enterprise and solution architecture, IT portfolio optimization and transformation to cloud, and deployment of global business solutions. Prior to joining OTSI, she held various executive-level positions with IBM, Prudential, and EDS, where she advised industry executives from Fortune *500 companies and US federal agencies delivering innovative business solutions. Ms. Jain managed technology across several enterprise-wide initiatives, leading architecture review boards and due diligence of potential technology acquisitions. Ms. Jain has worked in multiple business domains, including financials, healthcare, consulting, and government. She is a member of industry groups like ACT-IAC, focused on the public sector, AI, digital and IT transformation, data analytics, and cloud. She can be reached at seema@optimaltsi.com.*

WILL YOU STAY IN THE GAME?

AI Is Driving New Business Models: How Do We Adapt?

by Raj Ramesh

> Business models for creating and capturing value are shifting, giving rise to the new "hub economy," in which networks and data are the organizing principles. This, in turn, is creating a winner-take-all world, where a small number of hub firms, such as Amazon, Alibaba, Facebook, Google, Apple, and Microsoft, dominate.
>
> — Harvard Business School Professor and Cutter Fellow Karim Lakhani

Amazon is crushing the retail business, with venerable companies like J.C. Penney, Macy's, and Sears coming under pressure to shut down stores because their business models are outdated.[1] These old business models stand no chance against the nimbler entrants. Amazon's business model, enabled by the Internet, makes retail efficient. Technology has enabled Amazon to do things differently, resulting in lower prices, higher convenience, and more product choices. Often, companies can differentiate themselves from their competitors by performing completely different activities, or by performing the same activities differently to achieve similar outcomes. Amazon's technology-enabled business model has allowed it to poach customers from revered established companies and propelled it to a leading position in retail, while Uber and Airbnb are enabling new business models for the cab and hotel businesses, respectively. Leveraging smartphone technology, these latter two companies have circumvented traditional channels to reach their customers.

As the Internet and smart mobile devices have led to the creation of new business models, artificial intelligence (AI) is poised to do the same. This change will drive some established businesses to bankruptcy unless they prepare themselves against such inevitability. Newer companies, or "digital natives," are introducing new business models that incumbents must catch up to.[2] Consider Amazon and Macy's, both of which have used the Internet but in completely different ways. While Macy's has used it to offer another channel for shopping, Amazon has leveraged it as part of its foundation. Likewise, many companies will adopt AI into their operations, but a few will use it to create superior business models — and *that* is what incumbents must worry about.

New Business Models

First, let's agree on what we mean by a business model. Businesses make money by selling products and services. The business model defines how they do that. Every business has a business model. For example, Facebook and YouTube make money by charging advertisers while enabling their customers to produce and share content freely. LinkedIn offers basic content and features to its customers for free but charges for premium services.

Technology also helps to create new business models. Amazon uses the Internet to sell products without having to establish retail locations. Uber uses the smartphone to communicate with its customers who can be almost anywhere as they hail a car. Walmart uses computers to optimize its distribution operations so they are efficient, passing on those savings to customers in the form of low prices. The Internet provided a new distribution channel, the smartphone provided a direct connection to individual customers, and computers provided a tool to make operations more efficient.

As new technology appears on the horizon and matures, it forces the creation of new business models. These new business models often compete with the existing ones, and since the newer models are in some way distinct — offering better products, better prices, or better customer experiences — existing businesses must adapt and account for this new competition.[3] To adapt, they must change their business model to fend off what Clay Christensen termed "disruptive innovators."[4]

New AI Models

AI is one of those technologies that could be the basis of new business models, much the way Amazon leveraged the Internet. The term is commonly used as an umbrella for many subdisciplines, including computer vision, machine learning (ML), deep learning, and natural language processing (NLP), among others. Most of these consume large volumes of data, build models, and use those models to make inferences. The models

essentially represent the patterns of the data that was used to build the models in the first place.

Some business processes generate large amounts of high-quality information; these are good candidates for automation by AI. Other processes that do not generate much data are best facilitated by humans. Consider an insurance company that has collected risk information over many years. This vast amount of data can be used to build ML models to make inferences about the level of risk new customers pose. This business process is thus a good candidate for AI-based automation. On the other hand, business strategy also considers past information, but current information — and informed guesses about the future — are most important. These decisions still must be made by humans because human creativity is vital, and AI systems are currently unable to formulate strategy.

Thus, there are situations in business where a collaborative solution between machines and humans is optimal. Consider credit card fraud. When a customer's credit card number is stolen and used for purchases, it is impossible for humans to spot variations in patterns across billions of transactions. It is also not possible to write specific business rules to spot fraud due to the large number of variations. ML algorithms, on the other hand, could be used to create alerts for uncommon charge patterns by customers. One customer may be charging large amounts on a credit card daily, while another might rarely charge large amounts. So the spending patterns are different across different customers. With ML, the system could alert human operators to look at a smaller subset of possible fraudulent transactions to identify fraud.

As one of the most powerful technologies humans have invented, AI will spawn new business models that will be vast and deep, the likes of which we have not yet seen. The use of AI introduces cognitive capabilities into the business. There are many subdisciplines of AI that enable machines to perform cognitive tasks. For example, computer vision helps machines recognize objects, NLP helps them respond to natural speech, and ML helps machines identify patterns in large data sets, to name a few.[5]

With machines taking over many tasks that humans previously performed, companies will change the makeup of their organizations, moving affected employees to areas that require human skills such as imagination and creativity. Humans will focus on things that AI cannot do well today. This shift, with its fundamental organizational changes, will create new business models.

The Evolution of Insurance

One industry being disrupted by AI is auto insurance. Think about how the distribution of auto insurance has evolved. Before the Internet, the best way to sell auto insurance was through agents located in physical offices. These agents interacted directly with customers, and agents held the knowledge — the rules and constraints — of how insurance could be sold. When an insurer placed agents in more locations than its competitors, it gained a strategic advantage. More agents meant more reach, better distribution, and ultimately a higher number of policies sold.

As one of the most powerful technologies humans have invented, AI will spawn new business models that will be vast and deep, the likes of which we have not yet seen.

Then came the Internet. Some companies collected and provided detailed information about car models to their customers. Other companies collected and provided detailed information about people (who would insure cars) to their customers. A new insurance company could get all the information it needed about a person, and all the information it required about the car model that person owned, to compute a premium based on its proprietary risk model. Auto insurance was commoditized.

So the Internet created a new distribution channel, and a Web portal could be implemented at a fraction of the cost of hiring and training agents. Practically anyone with access to the Internet could buy auto insurance without meeting with an agent. A new Web-only business model can leapfrog the limitations of older business models, allowing a company to reach a wider audience more cheaply.

AI Impacts to Insurance

The Internet offered wide access, and the smartphone offered direct and personalized access to the insurance customer. What can AI offer?

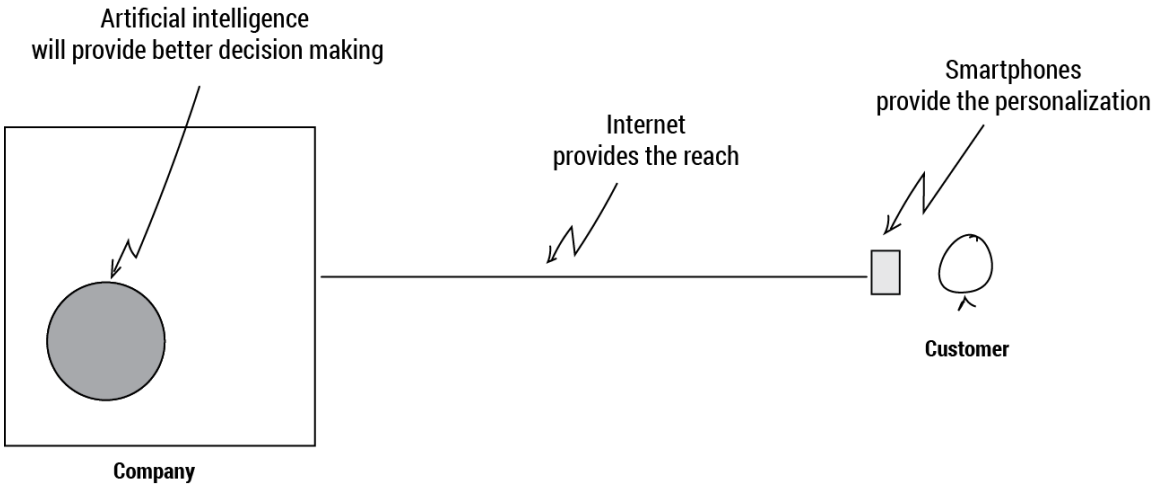

Figure 1 — AI enhances organizational capabilities.

AI is a cognitive engine that can process massive amounts of information and make quick decisions (see Figure 1). Today, there are many sources for information.

Consider the customer who uses a mapping app for directions and has opted to share such data with her insurance company. As she drives, the app can gather information about her destination, her speed at different parts of the journey, her braking patterns, and potentially other information to determine how safe a driver she is. This can only be done by AI, since there is a vast volume of real-time data being collected and processed. The company can then use the AI system's conclusion to price the customer's insurance premium, effectively giving lower rates for safe drivers and higher rates for reckless ones.

AI can handle other scenarios as well. For example, an AI system can be trained to look for patterns in customer questions and automatically answer the most common and easiest ones, freeing up time for customer representatives to handle those inquiries that require more complex answers. AI systems can also be trained to recommend the appropriate coverage to customers based on customer information and needs.

If AI takes over many of the tasks currently performed by humans, incumbent insurance companies may have to reassign people to other tasks. While we use insurance as an example, this is true for other industries. Today, many financial companies, for example, use robo-advisors to guide their clients' choice of investment options, offloading a task that a human financial advisor has traditionally performed.[6] As AI systems mature, they can perform many activities that humans have been doing. Reshuffling the organization to identify and/or create new roles for employees displaced by AI will become a major challenge and will become one of the strategic questions companies have to wrestle with in this new world of AI.

Moreover, new companies based on fully digital AI models are appearing in the insurance space. For example, Lemonade uses AI engines to compute premiums automatically by determining risk from data.[7] Previously, companies used a sizeable team of statisticians to build risk models from data. Now that AI can consume large amounts of data, models are not only more accurate, but can be kept up to date with a near-real-time data feed.

In the future, we may see customers migrate to such fully digital insurers that don't have the burdens of high employee costs and retail agent locations. Digital insurers could be the beginning of cheaper and faster insurance.

As another example of the level of disruption that's possible in insurance, imagine a new model of insurance from the customer's perspective. Your smartphone is the point of information collection, asking you a set of simple questions, and perhaps using the camera phone to get pictures of objects you want to insure. Using this information and the collaborative power of the Internet, the insurance company can pull in additional information about you and the objects you want to insure from other partner sites, through well-defined APIs. Using all the gathered information, the company can use AI to compute the best quote for you, to recommend other coverage you might need, and to ultra-personalize the delivery to your smartphone. If you accept the insurer's quote, the premium can be automatically debited from your account, and you are covered

immediately. Such quick response is possible only because of AI.

Where to Focus?

AI is a vast technology that includes many subdisciplines. For AI to provide good value to the business, it requires four major ingredients:

1. Acceptance by employees that AI will become ubiquitous, which in turn will impact their own roles
2. A reasonably large volume of good data for AI algorithms to consume
3. Appropriate algorithms to build the right models
4. Powerful machines

Companies whose primary focus is not AI are unlikely to have the resources and the employee expertise to develop AI. It is too much work for any company to develop, let alone a company whose main area of focus is not AI. Most non-AI companies should focus on and align their strategies with the first two ingredients listed above.

Building AI systems is a huge undertaking. Therefore, most companies should focus on helping employees adjust to the new world of AI, curating the right data and leaving the mechanics of building AI systems to vendors. The former is critically important.

Employee roles will certainly change with the use of AI systems. In a company with many employees, it is inevitable that some may lose their jobs to AI, just as some employees lost their jobs to automation during the computing revolution and some lost their jobs to mechanical systems during the Industrial Revolution. At Amazon Go, for example, there are no cashiers, and at Sam's Clubs, the Scan & Go app on the smartphone allows shoppers to skip the cashier.

AI will drive fundamental changes in the business model, such as being able to do more with fewer people. For example, if AI systems can evaluate risk and compute an insurance premium for an automobile, answer common customer questions, and even automatically process claims, then they are essentially taking over most of what an insurance agent currently does. This means that a company can potentially change its model from an agent-based distribution model to an Internet-based distribution model.

Data is the fuel for AI. Since data is so critical, companies will have to develop and execute on a data strategy, which should be part of the business model. The questions that such a strategy should answer include: What are our business objectives? Are we collecting the right data? Do we have to partner with others to get complementary data? What is our data quality? How do we make data an enterprise asset? Who owns the data? Do we have the right data expertise? And so on.

Data is the fuel for AI. Since data is so critical, companies will have to develop and execute on a data strategy, which should be part of the business model.

The company must collect, clean, organize, and govern its own unique data. In many cases, this data must be combined with external data to feed into the AI algorithms. As an example, if an insurance company wants to get a holistic understanding of an auto accident claim, it might get the car speed from Internet of Things devices that are managed by the auto manufacturer, driver information from its own data, pricing information from auto suppliers, and local weather information from weather information providers. The unique ways in which the company can combine multiple sources of data to build models and make decisions, using AI, give it a unique competitive advantage.

As mentioned earlier, integrating AI into the organization is a lot of work and must be done strategically. That strategy involves first focusing on the change mindset and data. For other aspects of the transformation, companies can seek the help of service providers. Companies such as Amazon, Microsoft, and Google are offering AI as a service (AIaaS), since they've realized that most companies will not be able to afford or would not want to build AI algorithms. If these service providers can build the AI algorithms and provide the computing power, then the consumers of their services can focus on the aspects that they need to get right, which are change management and data.

In summary, the prudent strategy for most companies is to own their data and to formulate change management, while leveraging AIaaS.

Conclusion

While AI drives the current changes in the business that include evaluating new business models, this is not the last time companies will have to adapt. The business environment is changing at a much faster rate than ever before, and it is imperative for companies to adapt continuously.

In the past, companies could afford to establish a business model and strategy and leave them in place for a few years before needing to consider the next change. Companies like Kodak did that successfully a couple of times in changing from wet to dry film processing, and again from black-and-white to color processing, but could not accept the need for change as people sought the convenience of digital cameras. Other companies, such as IBM, have adapted often and continue to be successful.

AI will surely compel the adoption of new business models, just as the Internet and smart mobile devices did. Companies need to understand the implications of AI and make some strategic choices. Those choices include whether it is sufficient to integrate AI into their business, or whether they should fundamentally change their business model to stay relevant. Either is acceptable, but such decisions must be made in the context of reality and with a deep understanding of the market forces that drive change.

Endnotes

[1]McGrath, Rita Gunther. "When Your Business Model Is in Trouble." Interview by Sarah Cliffe. *Harvard Business Review*, January–February 2011 (https://hbr.org/2011/01/when-your-business-model-is-in-trouble).

[2]Bollard, Albert, Ewan Duncan, Petko Rangelov, and Marta Rohr. "Accelerating the Shift to a Next-Generation Operating Model." *McKinsey Quarterly Review*, December 2017 (www.mckinsey.com/business-functions/digital-mckinsey/our-insights/accelerating-the-shift-to-a-next-generation-operating-model).

[3]Girotra, Karan, and Serguei Netessine. "Four Paths to Business Model Innovation." *Harvard Business Review*, July–August 2014 (https://hbr.org/2014/07/four-paths-to-business-model-innovation).

[4]Christiansen, Clay. "The Explainer: Disruptive Innovation." HBR Video, October 2013 (https://hbr.org/video/2688242135001/the-explainer-disruptive-innovation).

[5]For an overview of AI, see: Ramesh, Raj. "What Is Artificial Intelligence? In 5 Minutes." YouTube, August 2017 (www.youtube.com/watch?v=2ePf9rue1Ao).

[6]Davenport, Thomas H., Barry Libert, and Megan Beck. "Robo-Advisers Are Coming to Consulting and Corporate Strategy." *Harvard Business Review*, 12 January 2018 (https://hbr.org/2018/01/robo-advisers-are-coming-to-consulting-and-corporate-strategy).

[7]Lemonade (www.lemonade.com/).

Raj Ramesh has more than two decades of experience in business and IT across many domains. Dr. Ramesh has worked with senior-level business and IT leaders in Fortune *500 companies as an advisor and architect to foster deeper business-IT collaboration through consulting, advising, and building customized frameworks for common understanding. These frameworks are the tools to help the business and IT leaders achieve their specific objectives. He has written extensively and spoken at many national and international conferences. Dr. Ramesh loves to teach, and his hobby is to create short three- to five-minute videos on various business and IT topics. His three teenage children keep him busy the rest of the time. He can be reached at dr.raj.ramesh@topsigma.com.*

TRUST IS HARD TO COME BY

Building Trust in Artificial Intelligence, Machine Learning, and Robotics

by Keng Siau and Weiyu Wang

In 2016, Google DeepMind's AlphaGo beat 18-time world champion Lee Sedol in the abstract strategy board game *Go*.[1] This win was a triumphant moment for artificial intelligence (AI) that thrusted AI more prominently into the public view. Both AlphaGo Zero and AlphaZero have demonstrated the enormous potential of artificial intelligence and deep learning. Furthermore, self-driving cars, drones, and home robots are proliferating and advancing rapidly. AI is now part of our everyday life, and its encroachment is expected to intensify.[2, 3] Trust, however, is a potential stumbling block. Indeed, trust is key in ensuring the acceptance and continuing progress and development of artificial intelligence.

In this article, we look at trust in artificial intelligence, machine learning (ML), and robotics. We first review the concept of trust in AI and examine how trust in AI may be different from trust in other technologies. We then discuss the differences between interpersonal trust and trust in technology and suggest factors that are crucial in building initial trust and developing continuous trust in artificial intelligence.

What Is Trust?

The level of trust a person has in someone or something can determine that person's behavior. Trust is a primary reason for acceptance.[4] Trust is crucial in all kinds of relationships, such as human-social interactions,[5, 6] seller-buyer relationships[7, 8] and relationships among members of a virtual team.[9] Trust can also define the way people interact with technology.[10, 11]

Trust is viewed as: (1) a set of specific beliefs dealing with benevolence, competence, integrity, and predictability (trusting beliefs); (2) the willingness of one party to depend on another in a risky situation (trusting intention); or (3) the combination of these elements.

Table 1 lists some concepts and antecedents of trust in interpersonal relationships and trust people have toward specific types of technology, such as mobile technology and information systems. The conceptualization of trust in human-machine interaction is, however, slightly different (see Table 2).

Compared to trust in an interpersonal relationship, in which the trustor and trustee are both humans, the trustee in a human-technology/human-machine relationship could be either the technology per se and/or the technology provider. Further, trust in technology and trust in the provider will influence each other (see Figure 1).

> *Both initial trust formation and continuous trust development deserve special attention.*

Trust is dynamic. Empirical evidence has shown that trust is typically built up in a gradual manner, requiring ongoing two-way interactions.[12, 13] However, sometimes a trustor can decide whether to trust the trustee, such as an object or a relationship, before getting firsthand knowledge of the trustee — or having any kind of experience with the trustee. For example, when two persons meet for the first time, the first impression will affect the trust between these two persons. In such situations, trust will be built based on an individual's disposition or institutional cues.[14] This kind of trust is called *initial trust*, which is essential for promoting the adoption of a new technology.[15] Both initial trust formation and continuous trust development deserve special attention.[16] In the context of trust in AI, both initial trust formation and continuous trust development should be considered.

Trust Concepts and Antecedents (Factors That Lead to Trust)	Trust Context
Trust is the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party.[1] **Antecedents:** • *Ability/competence* — skills that enable a party to have influence within some specific domain (e.g., interpersonal communication) • *Benevolence* — the extent to which a trustee is believed to want to do good to the trustor; for instance, trustee has specific attachment to trustor (e.g., relationship between protégé and mentor) • *Integrity* — trustee adheres to a set of principles acceptable to trustor	Interpersonal trust in organization
Trust means that one believes in and is willing to depend on another party.[2] **Antecedents:** • *Disposition to trust* — consistent tendency of willingness to depend on others • *Institution-based trust* — sense of security one feels about a situation because of impersonal structures (e.g., regulations and guarantees) • *Cognitive process* — rapid and cognitive cues, including categorization process (e.g., reputation categorization) and illusions of control process	Initial trust formation in organization
Trust is the willingness of a party to be vulnerable to the actions of another party, irrespective of the ability to monitor or control the other party. **Antecedents:** • *Trustee attributes* — perceived ability, benevolence, and integrity • *Trustor attributes* — propensity to trust	Trust in virtual teams[3]
Trust has three characteristics: 1. Trustor and trustee 2. Uncertainty and risk 3. Honesty, benevolence, and lack of betrayal Two stages of trust building: 1. Initial trust formation 2. Continuous trust development	Trust in mobile commerce[4]
Trust is the belief that another is benevolent, competent, honest, or predictable in a given situation and includes the willingness to depend on another. **Antecedents:** • *Disposition to trust* — consistent tendency to be willing to depend on others • *Institution-based trust* — sense of security one feels about a situation due to impersonal structures (e.g., regulations and guarantees) • *Knowledge-based trust* — trust based on first-hand knowledge (i.e., familiarity) • *Calculative-based trust* — includes the calculation of pros and cons/costs and benefits	Trust in e-commerce[5, 6]
Trust is the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party. **Antecedents:** • *Personality base* — faith in humanity and trusting stance • *Cognitive base* — cognitive cues, including categorization process • *Calculative base* — calculation of pros and cons/costs and benefits • *Institutional base* — sense of security one feels about a situation due to institutional structures (e.g., regulations and guarantees)	Trust in information systems[7]

[1]Mayer, Roger C., James H., Davis, and F. David Schoorman. "An Integrative Model of Organizational Trust." *Academy of Management Review*, Vol. 20, No. 3, July 1995.

[2]McKnight, D. Harrison, Larry L. Cummings, and Norman L. Chervany. "Initial Trust Formation in New Organizational Relationships." *Academy of Management Review*, Vol. 23, No. 3, July 1998.

[3]Jarvenpaa, Sirkka L., Kathleen Knoll, and Dorothy E. Leidner. "Is Anybody Out There? Antecedents of Trust in Global Virtual Teams." *Journal of Management Information Systems*, Vol. 14, No. 4, Spring 1998.

[4]Siau, Keng, and Zixing Shen. "Building Customer Trust in Mobile Commerce." *Communications of the ACM*, Vol. 46, No. 4, April 2003.

[5]Gefen, David, Elena Karahanna, and Detmar W. Straub. "Trust and TAM in Online Shopping: An Integrated Model." *MIS Quarterly*, Vol. 27, No. 1, March 2003.

[6]McKnight, D. Harrison, Vivek Choudhury, and Charles Kacmar. "Trust in E-Commerce Vendors: A Two-Stage Model." *Proceedings of the 21st International Conference on Information Systems*. Association for Information Systems, 2000.

[7]Li, Xin, Traci J. Hess, and Joseph S. Valacich. "Why Do We Trust New Technology? A Study of Initial Trust Formation with Organizational Information Systems." *Journal of Strategic Information Systems*, Vol. 17, No. 1, 2008.

Table 1 — Conceptualization of trust and its antecedents.

Trust Concepts and Antecedents (Factors That Lead to Trust)	Trust Context
Trust is beliefs, attitudes, intentions, behavior, and goal-oriented. **Antecedents:**[1,2] • Human-related factors (e.g., expertise, personality traits) • Environment-related factors (e.g., culture) • Technology-related factors • Performance (i.e., reliability, predictability, and ability) • Process (i.e., algorithms of the automation) • Purpose (i.e., within the realm of designers' intent)	Trust in human-automation interaction
Five promoting factors:[3] 1. Reliability 2. Validity 3. Utility 4. Robustness 5. False-alarm rate	Trust in automation
Trust is the willingness of people to accept robot-produced information and follow robot's suggestions,[4,5] share tasks, contribute information, and provide support to robot.[6] **Antecedents:** • Human factors (i.e., abilities, characteristics/personalities) • Robot factors (i.e., performance and attributes) • Environment factors (i.e., team collaboration and nature of the tasks)	Trust in human-robot interaction
Three trust dimensions: 1. Performance (i.e., operational safety, data security, privacy protection) 2. Process (i.e., cognitive compatibility, usability, trialability [reducing concerns by inviting users to test the technology]) 3. Purpose	Trust in applied artificial intelligence:[7] • Technology • Innovating firm and communication

[1]Lee, John D., and Katrina A. See. "Trust in Automation: Designing for Appropriate Reliance." *Human Factors*, Vol. 46, No. 1, Spring 2004.

[2]Schaefer, Kristin E., et al. "A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Understanding Autonomy in Future Systems." *Human Factors*, Vol. 58, No. 3, March 2016.

[3]Hoffman, Robert R., Matthew Johnson, Jeffrey M. Bradshaw, and Al Underbrink. "Trust in Automation." *IEEE Intelligent Systems*, Vol. 28, No 1, January–February 2013.

[4]Hancock, Peter A., et al. "A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction." *Human Factors*, Vol. 53, No. 5, September 2011.

[5]Salem, Maha, Gabriella Lakatos, Farshid Amirabdollahian, and Kerstin Dautenhahn. "Would You Trust a (Faulty) Robot? Effects of Error, Task Type, and Personality on Human-Robot Cooperation and Trust." *Proceedings of the 10th Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2015.

[6]Oleson, Kristin E., et al. "Antecedents of Trust in Human-Robot Collaborations." *Proceedings of 1st International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support*. IEEE, 2011.

[7]Hengstler, Monika, Ellen Enkel, and Selina Duelli. "Applied Artificial Intelligence and Trust — The Case of Autonomous Vehicles and Medical Assistance Devices." *Technological Forecasting and Social Change*, Vol. 105, April 2016.

Table 2 — Trust conceptualization in human-machine interaction.

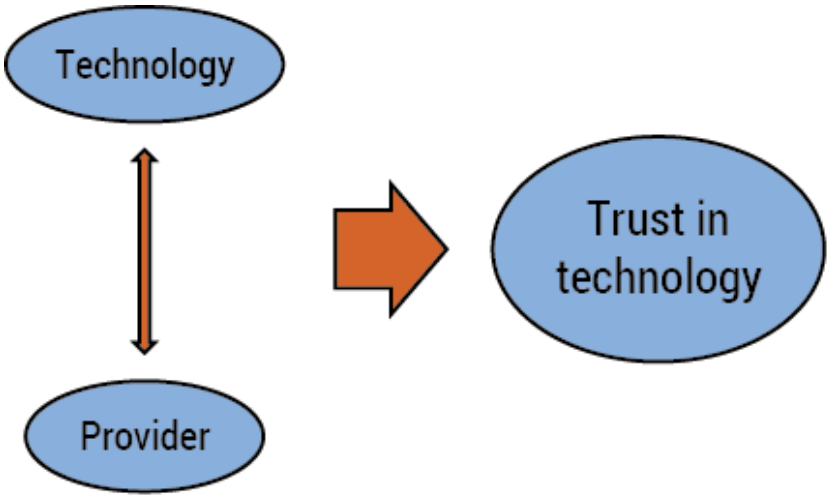

Figure 1 — Trust in technology interacts with trust in the provider of the technology.

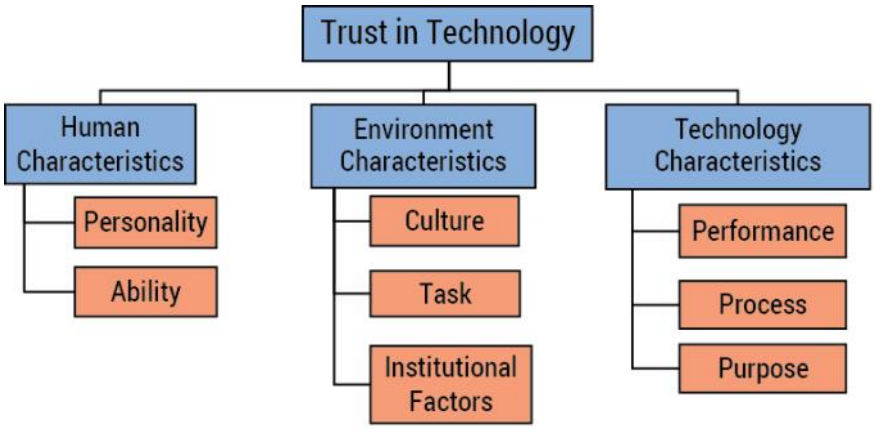

Figure 2 — Factors and dimensions of trust in technology.

What Is the Difference Between Trust in AI and Trust in Other Technologies?

Trust in technology is determined by human characteristics,[17] environment characteristics,[18] and technology characteristics.[19] Figure 2 shows the factors and dimensions of trust in technology.

Human characteristics basically consider the human's personality, the trustor's disposition to trust, and the trustee's ability to deal with risks. The trustor's personality or disposition to trust could be thought of as the general willingness to trust others, and it depends on different experiences, personality types, and cultural backgrounds. Ability usually refers to a trustee's competence/group of skills to complete tasks in a specific domain. For instance, if an employee is very competent in negotiation, the manager may trust the employee when he or she takes charge of negotiating contracts.

Environment characteristics focus on elements such as the nature of the tasks, cultural background, and institutional factors. Tasks can be of different natures. For example, a task can be very important or a task can be trivial. Cultural background can be based on ethnicity, race, religion, and socioeconomic status. Cultural background can also be associated with a country or a particular region. For instance, Americans tend to trust strangers who share the same group memberships, and Japanese tend to trust those who share direct or indirect relationship links.[20] Institutional factors indicate the impersonal structures that enable one to act in anticipation of a successful future endeavor. Institutional factors, according to literature, include two main aspects: the situational normality and structural assurances. Situational normality means the situation is normal, and everything is in proper order. Structural assurances refer to the contextual conditions such as promises, contracts, guarantees, and regulations.

No matter who or what the trustee is, whether it is a human, a form of AI, or an object such as an organization or a virtual team, the impact of human characteristics and environment characteristics will be roughly similar. For instance, a person with a high-trusting stance would be more likely to accept and depend on others, such as a new technology or a new team member. Similarly, it will be easier for a technology provided by an institution/organization with a high reputation to gain trust from users than it would be for a similar technology from an institution/organization without such a reputation.

Technology characteristics can be analyzed from three perspectives: (1) the performance of the technology, (2) its process/attributes, and (3) its purpose. Although human and environment characteristics are fairly similar irrespective of trustee, the technology characteristics impacting trust will be different for AI, ML, and robotics than they are for other objects or humans. Since artificial intelligence has many new features compared to other technologies, its performance, process, and purpose need to be defined and considered. Using a two-stage model of trust building,[21] Table 3 shows the technology features related to AI's performance, process, and purpose, and their impact on trust.

Building Initial Trust in AI

Several factors are at play during trust building. These factors include the following:

Initial Trust Formation	Continuous Trust Development
Performance: • Representation • Image/perception • Reviews from other users Process: • Transparency and ability to explain • Trialability	Performance: • Usability and reliability • Collaboration and communication • Sociability and bonding • Security and privacy protection • Interpretability Purpose: • Job replacement • Goal congruence

Table 3 — Technology features of AI that affect trust building.

- **Representation.** Representation plays an important role in initial trust building, and that is why humanoid robots are so popular. The more a robot looks like a human, the easier people can establish an emotional connection with it. A robot dog is another example of an AI representation that humans find easier to trust. Dogs are human's best friends and represent loyalty and diligence.

- **Image/perception.** Sci-fi books and movies have given AI a bad image — when the intelligence we create gets out of control. Artificial general intelligence (AGI) or "strong AI"[22] can be a serious threat. This image and perception will affect people's initial trust in AI.

- **Reviews from other users.** Reading online reviews is common these days. Compared with a negative review, a positive review leads to greater initial trust.[23] Reviews from other users will affect the initial trust level.

- **Transparency and "explainability."** To trust AI applications, we need to understand how they are programmed and what function will be performed in certain conditions. This transparency is important, and AI should be able to explain/justify its behaviors and decisions. One of the challenges in machine learning and deep learning is the black box in the ML and decision-making processes. If the explainability of the AI application is poor or missing, trust is affected.

- **Trialability.** Trialability means the opportunity for people to have access to the AI application and to try it before accepting or adopting it. Trialability enables enhancement of understanding. In an article in *Technological Forecasting and Social Change*, Monika Hengstler et al. state that "when you survey the perception of new technologies across generations, you typically see resistance appear from people who are not users of technology."[24] Thus, providing chances for potential users to try the new technology will promote higher initial trust.

Developing Continuous Trust in AI

Once trust is established, it must be nurtured and maintained. This happens through the following:

- **Usability and reliability.** Performance includes the competence of AI in completing tasks and finishing those tasks in a consistent and reliable manner. The AI application should be designed to operate easily and intuitively. There should be no unexpected downtime or crashes. Usability and reliability contribute to continuous trust.

- **Collaboration and communication.** Although most AI applications are developed to perform tasks independently, the most likely scenario in the short term is that people will work in partnership with intelligent machines. Whether collaboration and communication can be carried out smoothly and easily will affect continuous trust.

- **Sociability and bonding.** Humans are social animals. Continuous trust can be enhanced with social activities. A robot dog that can recognize its owner and show affection may be treated like a pet dog, establishing emotional connection and trust.

- **Security and privacy protection.** Operational safety and data security are two eminent factors that influence trust in technology.[25] People are unlikely to trust anything that is too risky to operate. Data security, for instance, is important because machine learning relies on large data sets, making privacy a concern.

- **Interpretability.** With a black box, most ML models are inscrutable. To address this problem, it is necessary to design interpretable models and provide the ability for the machine to explain its conclusions or actions. This could help users understand the rationale for the outcomes and the process of deriving the results. Transparency and explainability, as discussed in initial trust building, are important for continuous trust as well.

- **Job replacement.** Artificial intelligence can surpass human performance in many jobs and replace human workers. AI will continue to enhance its capability and infiltrate more domains. Concern about AI taking jobs and replacing human employees will impede people's trust in artificial intelligence. For example, those whose jobs may be replaced by AI may not want to trust it. Some predict that more than 800 million workers, about a fifth of the global labor force, might lose their jobs soon.[26] Lower-skilled, repetitive, and dangerous jobs are among those likely to be taken over by machines. Providing retraining and education to affected employees will help mitigate this effect on continuous trust.

Trust building is a dynamic process, involving movement from initial trust to continuous trust development. Continuous trust will depend on the performance and purpose of the artificial intelligence.

- **Goal congruence.** Since artificial intelligence has the potential to demonstrate and even surpass human intelligence, it is understandable that people treat it as a threat. And AI *should* be perceived as a potential threat, especially AGI! Making sure that AI's goals are congruent with human goals is a precursor in maintaining continuous trust. Ethics and governance of artificial intelligence are areas that need more attention.

Practical Implications and Conclusions

Artificial intelligence is here, and AI applications will become more and more prevalent. Trust is crucial in the development and acceptance of AI. In addition to the human and environment characteristics, which affect trust in other humans, objects, and AI, trust in AI, ML, and robotics is affected by the unique technology features of artificial intelligence.

To enhance trust, practitioners can try to maximize the technological features in AI systems based on the factors listed in Table 3. The representation of an AI as a humanoid or a loyal pet (e.g., dog) will facilitate initial trust formation. The image and perception of AI as a terminator (like in the *Terminator* movies) will hinder initial trust. In this Internet age, reviews are critical, as well as the ability of artificial intelligence to be transparent and able to explain its behavior/decisions. These are important for initial trust formation. The ability to try out AI applications will also have an impact on initial trust.

Trust building is a dynamic process, involving movement from initial trust to continuous trust development. Continuous trust will depend on the performance and purpose of the artificial intelligence. AI applications that are easy to use and reliable — and can collaborate and interface well with humans, have social ability, facilitate bonding with humans, provide good security and privacy protection, and explain the rationale behind conclusions or actions — will facilitate continuous trust development. A lack of clarity over job replacement and displacement by AI along with AI's potential threat to the existence of humanity breed distrust and hamper continuous trust development.

Trust is the cornerstone of humanity's relationship with artificial intelligence. Like any type of trust, trust in AI takes time to build, seconds to break, and forever to repair once it is broken!

Endnotes

[1]"AlphaGo Versus Lee Sedol." Wikipedia (https://en.wikipedia.org/wiki/AlphaGo_versus_Lee_Sedol).

[2]Siau, Keng. "Impact of Artificial Intelligence, Robotics, and Automation on Higher Education." Proceedings of the *23rd Americas Conference on Information Systems 2017 (AMCIS 2017)*, Boston, Massachusetts, USA, 10-12 August 2017 (http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1579&context=amcis2017).

[3]Siau, Keng, and Ying Yang. "Impact of Artificial Intelligence, Robotics, and Machine Learning on Sales and Marketing." Proceedings of *12th Annual Midwest Association for Information Systems Conference (MWAIS 2017)*, Springfield, Illinois, USA, 18-19 May 2017 (https://aisel.aisnet.org/mwais2017/48).

[4]Gefen, David, Elena Karahanna, and Detmar W. Straub. "Trust and TAM in Online Shopping: An Integrated Model." *MIS Quarterly*, Vol. 27, No. 1, March 2003 (https://pdfs.semanticscholar.org/e2e8/6748e6abeb7c1077ee7a93029fa4d7843d70.pdf).

[5]Hengstler, Monika, Ellen Enkel, and Selina Duelli. "Applied Artificial Intelligence and Trust — The Case of Autonomous Vehicles and Medical Assistance Devices." *Technological Forecasting and Social Change*, Vol. 105, April 2016 (https://www.sciencedirect.com/science/article/pii/S0040162515004187).

[6]McKnight, D. Harrison, Larry L. Cummings, and Norman L. Chervany. "Initial Trust Formation in New Organizational Relationships."*Academy of Management Review*, Vol. 23, No. 3, 1998 (https://www.jstor.org/stable/259290?seq=1#page_scan_tab_contents).

[7]Gefen, David. "E-Commerce: The Role of Familiarity and Trust." *Omega*, Vol. 28, No. 6, 2000 (https://www.sciencedirect.com/science/article/pii/S0305048300000219).

[8]Siau, Keng, and Zixing Shen. "Building Customer Trust in Mobile Commerce." *Communications of the ACM*, Vol. 46, No. 4, April 2003 (https://cacm.acm.org/magazines/2003/4/6855-building-customer-trust-in-mobile-commerce/abstract).

[9]Coppola, Nancy W., N. Rotter, and Starr Roxanne Hiltz. "Building Trust in Virtual Teams." *IEEE Transactions on Professional Communication*, Vol. 47, No. 2, 2004 (https://www.researchgate.net/profile/N_Rotter/publication/3230326_Building_Trust_in_Virtual_Teams/links/00b495177d91ab09c8000000/Building-Trust-in-Virtual-Teams.pdf).

[10]Li, Xin, Traci J. Hess, and Joseph S. Valacich. "Why Do We Trust New Technology? A Study of Initial Trust Formation with Organizational Information Systems." *Journal of Strategic Information Systems*, Vol. 17, No. 1, 2008 (https://www.sciencedirect.com/science/article/abs/pii/S0963868708000036).

[11]Siau, Keng, et al. "A Qualitative Investigation on Consumer Trust in Mobile Commerce." *International Journal of Electronic Business*, Vol. 2, No. 3, 2004 (http://cba.unl.edu/research/articles/645/).

[12]Gefen (see 7).

[13]McKnight et al. (see 6).

[14]McKnight et al. (see 6).

[15]Li et al. (see 10).

[16]Siau and Shen (see 8).

[17]Hengstler et al. (see 5).

[18]Oleson, Kristin E., et al. "Antecedents of Trust in Human-Robot Collaborations." *Proceedings of 1st International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support.* IEEE, 2011 (http://ieeexplore.ieee.org/document/5753439/).

[19]Schaefer, Kristin E., et al. "A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Understanding Autonomy in Future Systems." *Human Factors*, Vol. 58, No. 3, 2016 (https://www.ncbi.nlm.nih.gov/pubmed/27005902).

[20]Yuki, Masaki, et al. "Cross-Cultural Differences in Relationship- and Group-Based Trust." *Personality and Social Psychology Bulletin*, Vol. 31, No. 1, 2005 (http://journals.sagepub.com/doi/abs/10.1177/0146167204271305).

[21]Siau and Shen (see 8).

[22]AGI or "strong AI" refers to the AI technology that can perform almost (if not) all the intellectual tasks that a human being can.

[23]Kusumasondjaja, Sony, Tekle Shanka, and Christopher Marchegiani. "Credibility of Online Reviews and Initial Trust: The Roles of Reviewer's Identity and Review Valence." *Journal of Vacation Marketing*, Vol. 18, No. 3, 2012 (http://journals.sagepub.com/doi/abs/10.1177/1356766712449365).

[24]Hengstler et al. (see 5).

[25]Hengstler et al. (see 5).

[26]Gray, Alex. "These Are the Jobs Most Likely to Be Taken by Robots." World Economic Forum, 15 December 2017 (https://www.weforum.org/agenda/2017/12/robots-coming-for-800-million-jobs?utm_content=buffer8f3ff&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer).

Keng Siau is Chair of the Department of Business and Information Technology at the Missouri University of Science and Technology. Previously, he was the Edwin J. Faulkner Chair Professor and Full Professor of Management at the University of Nebraska-Lincoln (UNL), where he was Director of the UNL-IBM Global Innovation Hub. Dr. Siau has written more than 250 academic publications and is consistently ranked as one of the top information systems researchers in the world based on h-index and productivity rate. His research has been funded by the US National Science Foundation, IBM, and other IT organizations. Dr. Siau has received numerous teaching, research, service, and leadership awards, including the International Federation for Information Processing Outstanding Service Award, the IBM Faculty Award, and the IBM Faculty Innovation Award. He received his PhD in business administration from the University of British Columbia. He can be reached at siauk@mst.edu.

Weiyu Wang is currently pursuing a master of science degree in information science and technology at Missouri University of Science and Technology. Ms. Wang received an MBA from the Missouri University of Science and Technology. She can be reached at wwpmc@mst.edu.

Cutter Membership
The Ultimate Access to the Experts

Cutter Consortium Membership opens up multiple avenues to interact with Cutter's experts to brainstorm and gain guidance to transform your organization and boost success.

Like everything business technology, one size does not fit all.

That's why we encourage you to choose the Membership that's right for your organization. Whether you choose Digital Transformation & Innovation, Enterprise-wide, Practice-specific, or CIO Membership you'll see a strong return.

Contact us at +1 781 648 8700 or sales@cutter.com to arrange a sample inquiry call with a Cutter expert and see for yourself how quickly your return on Membership can be realized.

What Do You Get from Cutter Membership?

- Get guidance in leveraging new strategies, emerging technologies, and business management practices to enable digital transformation and boost competitive advantage.
- Learn how to mine data to create new products and services and improve customer experience.
- Get input on how to reduce expenses through more cost-effective strategies.
- Gain insights and get ideas on achieving sustainable innovation, successful change management, and prudent risk management.
- Get coaching and insights on leadership and team-building practices that boost productivity.
- Discover vendor-agnostic advice to ensure unbiased purchasing decisions.

Test Drive Cutter Today

Have a question? A challenge you're trying to overcome? Have you reached a bottleneck and need some expert advice? Find out for yourself how valuable Cutter Membership can be — request a sample inquiry with a Cutter expert now.

Call +1 781 648 8700 or email sales@cutter.com.

CUTTER CONSORTIUM
Access to the Experts

37 Broadway, Suite 1
Arlington, MA 02474 USA
Tel: +1 781 648 8700
Fax: +1 781 648 8707
Web: www.cutter.com
Email: sales@cutter.com

"I have personally been able to leverage Cutter's services since 1999. Among the attributes that differentiate Cutter from other firms, two remain at the top of my list: **Thought Leadership and Real Value**. Thought Leadership is driven by Cutter's experts. Real Value, driven by 'Access to the experts,' pushes us beyond just understanding the options. We can develop relationships with the experts, and tailor the options they present so that they can be quickly and practically executed within our organization, enabling our Business Technology team to continually improve, engage, and contribute to business growth."

— Doug Mikaelian,
VP Business Technology,
Dairy Farmers of America

"The Information Technology revolution continues to change how businesses work and how people work in and with them. Today's demands on software to provide useful, usable, secure, and scalable services stress all aspects of systems development. These include architecture, technology selection, product development and lifecycle management. Cutter provides comprehensive coverage for these crucial topics — through research reports and customized virtual training delivered globally — helping us to understand needs and solutions that span traditional silos."

— Eric Schoen
Director of Engineering, i2k Connect LLC;
Former Chief Software Architect,
Schlumberger

"Thanks to Cutter´s experts, in a short time we have been able to:

- Improve our capabilities, using the contract scorecard for outsourcing.
- Gain key insights and practical strategies for responding to challenges we face in building trust and partnerships within the organization and outside our institutional borders.
- Improve our methods for working with people and anticipating the challenges of cultural transformation within our organization.
- Have immediate access to key 'experts' at critical times, who spend quality time with our teams, provide state-of-the-art orientation, and help us implement projects.
- Implement agile, effective, and articulated innovation around a common vision for achieving quality education in Mexico."

— Arturo Cervantes
Director General of Information Systems for Evaluation Results, National Institute for Educational Evaluation, Mexico

Cutter Business Technology JOURNAL

About Cutter Consortium

Cutter Consortium is a unique, global business technology advisory firm dedicated to helping organizations leverage emerging technologies and the latest business management thinking to achieve competitive advantage and mission success. Through its research, training, executive education, and consulting, Cutter enables digital transformation.

Cutter helps clients address the spectrum of challenges technology change brings — from disruption of business models and the sustainable innovation, change management, and leadership a new order demands, to the creation, implementation, and optimization of software and systems that power newly holistic enterprise and business unit strategies.

Cutter pushes the thinking in the field by fostering debate and collaboration among its global community of thought leaders. Coupled with its famously objective "no ties to vendors" policy, Cutter's *Access to the Experts* approach delivers cutting-edge, objective information and innovative solutions to its clients worldwide.

For more information, visit www.cutter.com or call us at +1 781 648 8700.

CUTTER CONSORTIUM
Access to the Experts