McKinsey&Company

# NOTES FROM THE AI FRONTIER
# INSIGHTS FROM HUNDREDS OF USE CASES

**DISCUSSION PAPER**
APRIL 2018

Michael Chui | San Francisco
James Manyika | San Francisco
Mehdi Miremadi | Chicago
Nicolaus Henke | London
Rita Chung | Silicon Valley
Pieter Nel | New York
Sankalp Malhotra | New York

首席AI观旗下有3大类40余小类3000多AI社群，现进一步邀请行业人士加入。

| | | | | |
|---|---|---|---|---|
| **AI技术群** | 云计算与大数据　机器学习　自然语言处理（NLP）　语音识别<br>计算机视觉（CV）　物联网技术　VR/AR　AI芯片 | | | |
| **AI应用** | 自动驾驶　安防　机器人　智能家居　智能音箱　物流<br>新零售　金融　教育　影视游戏　医疗健康　通信服务<br>智慧城市　智慧建筑 | | | |
| **城市群** | 北京　深圳　上海　武汉　广州　杭州　成都　南京　苏州　西安 | | | |

入群方法：

1、添加小助手微信（zhidxcom001）

2、回复"入群"，根据提示操作即可。

# MCKINSEY GLOBAL INSTITUTE

Since its founding in 1990, the McKinsey Global Institute (MGI) has sought to develop a deeper understanding of the evolving global economy. As the business and economics research arm of McKinsey & Company, MGI aims to provide leaders in the commercial, public, and social sectors with the facts and insights on which to base management and policy decisions.

MGI research combines the disciplines of economics and management, employing the analytical tools of economics with the insights of business leaders. Our "micro-to-macro" methodology examines microeconomic industry trends to better understand the broad macroeconomic forces affecting business strategy and public policy. MGI's in-depth reports have covered more than 20 countries and 30 industries. Current research focuses on six themes: productivity and growth, natural resources, labor markets, the evolution of global financial markets, the economic impact of technology and innovation, and urbanization. Recent reports have assessed the digital economy, the impact of AI and automation on employment, income inequality, the productivity puzzle, the economic benefits of tackling gender inequality, a new era of global competition, Chinese innovation, and digital and financial globalization.

MGI is led by three McKinsey & Company senior partners: Jacques Bughin, Jonathan Woetzel, and James Manyika, who also serves as the chairman of MGI. Michael Chui, Susan Lund, Anu Madgavkar, Jan Mischke, Sree Ramaswamy, and Jaana Remes are MGI partners, and Mekala Krishnan and Jeongmin Seong are MGI senior fellows.

Project teams are led by the MGI partners and a group of senior fellows, and include consultants from McKinsey offices around the world. These teams draw on McKinsey's global network of partners and industry and management experts. Advice and input to MGI research are provided by the MGI Council, members of which are also involved in MGI's research. MGI Council members are drawn from around the world and from various sectors and include Andrés Cadena, Sandrine Devillard, Richard Dobbs, Tarek Elmasry, Katy George, Rajat Gupta, Eric Hazan, Eric Labaye, Acha Leke, Scott Nyquist, Gary Pinkus, Sven Smit, Oliver Tonby, and Eckart Windhagen. In addition, leading economists, including Nobel laureates, act as research advisers to MGI research.

The partners of McKinsey fund MGI's research; it is not commissioned by any business, government, or other institution. For further information about MGI and to download reports, please visit **www.mckinsey.com/mgi**.

---

## MCKINSEY ANALYTICS

McKinsey Analytics helps clients achieve better performance through data. We work together with clients to build analytics-driven organizations, helping them develop the strategies, operations, and capabilities to derive rapid and sustained impact from analytics. Over the past five years, we have worked with more than 2,000 clients across every industry and business function. McKinsey Analytics is led globally by Nicolaus Henke and Noshir Kaka, together with an executive committee comprised of 40 McKinsey senior partners representing all regions and practices. Today, McKinsey Analytics brings together more than 1,900 advanced analytics and AI experts and spans more than 125 domains (industry- and function-specific teams with people, data, and tools focused on unique applications of analytics). McKinsey Analytics includes several acquired companies such as QuantumBlack, a leading advanced analytics firm that McKinsey acquired in 2015.

Learn more at **www.mckinsey.com/business-functions/mckinsey-analytics/our-insights.**

# NOTES FROM THE AI FRONTIER: INSIGHTS FROM HUNDREDS OF USE CASES

For this discussion paper, part of our ongoing research into evolving technologies and their effect on business, economies, and society, we mapped traditional analytics and newer "deep learning" techniques and the problems they can solve to more than 400 specific use cases in companies and organizations. Drawing on MGI research and the applied experience with artificial intelligence (AI) of McKinsey Analytics, we assess both the practical applications and the economic potential of advanced AI techniques across industries and business functions. We continue to study these AI techniques and additional use cases. For now, here are our key findings:

- AI, which for the purposes of this paper we characterize as "deep learning" techniques using artificial neural networks, can be used to solve a variety of problems. Techniques that address classification, estimation, and clustering problems are currently the most widely applicable in the use cases we have identified, reflecting the problems whose solutions drive value across the range of sectors.

- The greatest potential for AI we have found is to create value in use cases in which more established analytical techniques such as regression and classification techniques can already be used, but where neural network techniques could provide higher performance or generate additional insights and applications. This is true for 69 percent of the AI use cases identified in our study. In only 16 percent of use cases did we find a "greenfield" AI solution that was applicable where other analytics methods would not be effective.

- Because of the wide applicability of AI across the economy, the types of use cases with the greatest value potential vary by sector. These variations primarily result from the relative importance of different drivers of value within each sector. They are also affected by the availability of data, its suitability for available techniques, and the applicability of various techniques and algorithmic solutions. In consumer-facing industries such as retail, for example, marketing and sales is the area with the most value. In industries such as advanced manufacturing, in which operational performance drives corporate performance, the greatest potential is in supply chain, logistics, and manufacturing.

- The deep learning techniques on which we focused — feed forward neural networks, recurrent neural networks, and convolutional neural networks—account for about 40 percent of the annual value potentially created by all analytics techniques. These three techniques together can potentially enable the creation of between $3.5 trillion and $5.8 trillion in value annually. Within industries, that is the equivalent of 1 to 9 percent of 2016 revenue.

- Capturing the potential impact of these techniques requires solving multiple problems. Technical limitations include the need for a large volume and variety of often labeled training data, although continued advances are already helping address these. Tougher perhaps may be the readiness and capability challenges for some organizations. Societal concern and regulation, for example about privacy and use of personal data, can also constrain AI use in banking, insurance, health care, and pharmaceutical and medical products, as well as in the public and social sectors, if these issues are not properly addressed.

- The scale of the potential economic and societal impact creates an imperative for all the participants—AI innovators, AI-using companies and policy-makers—to ensure a vibrant AI environment that can effectively and safely capture the economic and societal benefits.

# INTRODUCTION

Artificial intelligence (AI) stands out as a transformational technology of our digital age. Questions about what it is, what it can already do—and what it has the potential to become—cut across technology, psychology, politics, economics, science fiction, law, and ethics. AI is the subject of countless discussions and articles, from treatises about technical advances to tabloid headlines about its effects. Even as the debate continues, the technologies underpinning AI continue to move forward, enabling applications from facial recognition in smartphones to consumer apps that use AI algorithms to detect diabetes and hypertension with increasing accuracy.[1] Indeed, while much of the public discussion of AI focuses on science fiction-like AI realization such as robots, the number of less-noticed practical applications for AI throughout the economy is growing apace and permeating our lives.

This discussion paper seeks to contribute to the body of knowledge about AI by mapping AI techniques to the types of problems they can help solve and then mapping these problem types to more than 400 practical use cases and applications in businesses across 19 industries, from aerospace and defense to travel and the public sector, and nine business functions ranging from marketing and sales and supply-chain management to product development and human resources.[2] Drawing on a wide variety of public and proprietary data sources, including the experiences of our McKinsey & Company colleagues, we also assess the potential economic value of the latest generations of AI technologies. The AI techniques we focus on are deep learning techniques based on artificial neural networks, which we see as generating as much as 40 percent of the total potential value that all analytics techniques could provide.

Our findings highlight the substantial potential of applying deep learning techniques to use cases across the economy; these techniques can provide an incremental lift beyond that from more traditional analytics techniques. We identify the industries and business functions in which there is value to be captured, and we estimate how large that value could be globally. For all the potential, much work needs to be done to overcome a range of limitations and obstacles to AI application. We conclude with a brief discussion of these obstacles and of future opportunities as the technologies continue their advance. Ultimately, the value of AI is not to be found in the models themselves, but in organizations' abilities to harness them. Business leaders will need to prioritize and make careful choices about how, when, and where to deploy them.

This paper is part of our continuing research into analytics, automation, and AI technologies, and their effect on business, the economy, and society.[3] It is not intended to serve as a comprehensive guide to deploying AI; for example, we identify but do not elaborate on issues of data strategy, data engineering, governance, or change management and culture

---

[1] Geoffrey H. Tison et al., "Cardiovascular risk stratification using off-the-shelf wearables and a multi-mask deep learning algorithm," *Circulation*, volume 136, supplement 1, November 14, 2017.

[2] We do not identify the companies by name or country, for reasons of client confidentiality.

[3] Previous McKinsey Global Institute reports on these issues include *The age of analytics: Competing in a data-driven world*, December 2016; *A future that works: Automation, employment and productivity*, January 2017; and *Artificial intelligence: The next digital frontier?* June 2017. See a list of our related research at the end of this paper.

that are vital for companies seeking to capture value from AI and analytics.[4] The use cases we examined are not exhaustive; indeed, we continue to identify and examine others, and we may update our findings in due course. Nonetheless, we believe that this research can make a useful contribution to our understanding of what AI can and can't (yet) do, and how much value could be derived from its use. It is important to highlight that, even as we see economic potential in the use of AI techniques, the use of data must always take into account concerns including data security, privacy, and potential issues of bias, issues we have addressed elsewhere.[5]

# 1. MAPPING AI TECHNIQUES TO PROBLEM TYPES

As artificial intelligence technologies advance, so does the definition of which techniques constitute AI (see Box 1, "Deep learning's origins and pioneers").[6] For the purposes of this paper, we use AI as shorthand specifically to refer to deep learning techniques that use artificial neural networks. In this section, we define a range of AI and advanced analytics techniques as well as key problem types to which these techniques can be applied.

## NEURAL NETWORKS AND OTHER MACHINE LEARNING TECHNIQUES

We looked at the value potential of a range of analytics techniques. The focus of our research was on methods using artificial neural networks for deep learning, which we collectively refer to as AI in this paper, understanding that in different times and contexts, other techniques can and have been included in AI. We also examined other machine learning techniques and traditional analytics techniques (Exhibit 1). We focused on specific potential applications of AI in business and the public sector (sometimes described as "artificial narrow AI") rather than the longer-term possibility of an "artificial general intelligence" that could potentially perform any intellectual task a human being is capable of.

**Exhibit 1**

**Artificial intelligence, machine learning, and other analytics techniques that we examined for this research**
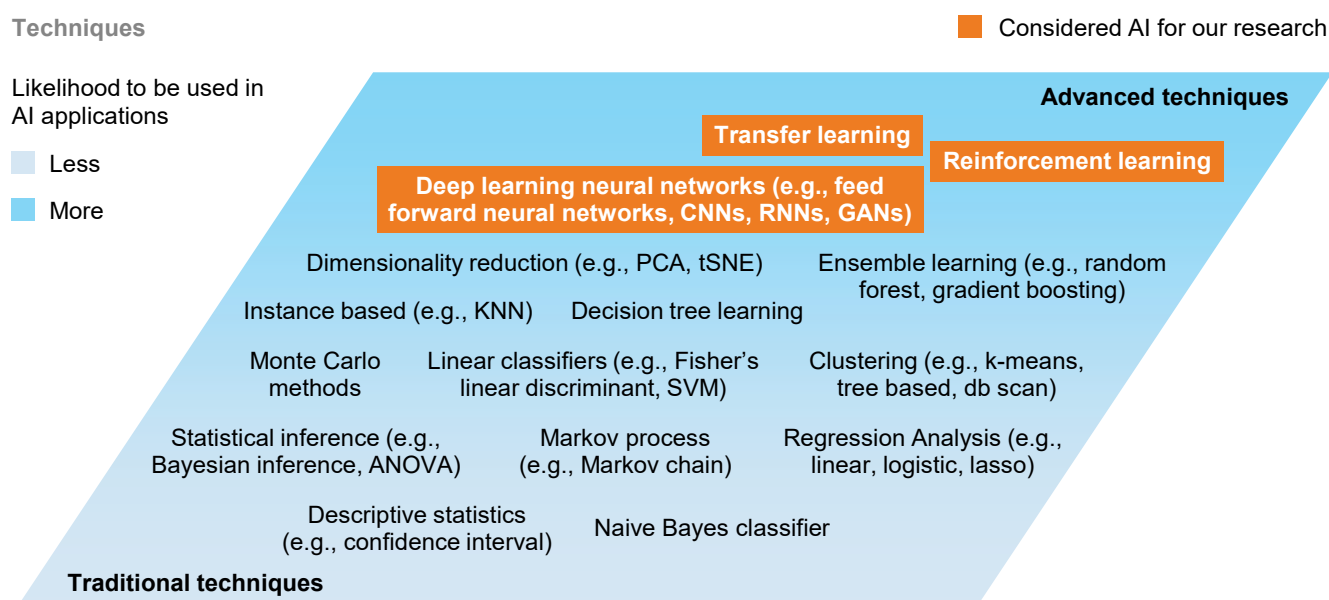


SOURCE: McKinsey Global Institute analysis

[4] See Jacques Bughin, Brian McCarthy, and Michael Chui, "A survey of 3,000 executives reveals how businesses succeed with AI," *Harvard Business Review*, August 28, 2017.

[5] Michael Chui, James Manyika, and Mehdi Miremadi, "What AI can and can't do (yet) for your business," *McKinsey Quarterly*, January 2018.

[6] For a detailed look at AI techniques, see *An executive's guide to AI*, McKinsey Analytics, January 2018. https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai

## Box 1: Deep learning's origins and pioneers

It is too early to write a full history of deep learning—and some of the details are contested—but we can already trace an admittedly incomplete outline of its origins and identify some of the pioneers. They include Warren McCulloch and Walter Pitts, who as early as 1943 proposed an artificial neuron, a computational model of the "nerve net" in the brain.[1] Bernard Widrow and Ted Hoff at Stanford University, developed a neural network application by reducing noise in phone lines in the late 1950s.[2] Around the same time, Frank Rosenblatt, an American psychologist, introduced the idea of a device called the Perceptron, which mimicked the neural structure of the brain and showed an ability to learn.[3] MIT's Marvin Minsky and Seymour Papert then put a damper on this research in their 1969 book "Perceptrons", by showing mathematically that the Perceptron could only perform very basic tasks.[4] Their book also discussed the difficulty of training multi-layer neural networks. In 1986, Geoffrey Hinton at the University of Toronto, along with colleagues David Rumelhart and Ronald Williams, solved this training problem with the publication of a now famous back propagation training algorithm—although some practitioners point to a Finnish mathematician, Seppo Linnainmaa, as having invented back propagation already in the 1960s.[5] Yann LeCun at NYU pioneered the use of neural networks on image recognition tasks and his 1998 paper defined the concept of convolutional neural networks, which mimic the human visual cortex.[6] In parallel, John Hopfield popularized the "Hopfield" network which was the first recurrent neural network.[7] This was subsequently expanded upon by Jurgen Schmidhuber and Sepp Hochreiter in 1997 with the introduction of the long short-term memory (LSTM), greatly improving the efficiency and practicality of recurrent neural networks.[8] Hinton and two of his students in 2012 highlighted the power of deep learning when they obtained significant results in the well-known ImageNet competition, based on a dataset collated by Fei-Fei Li and others.[9] At the same time, Jeffrey Dean and Andrew Ng were doing breakthrough work on large scale image recognition at Google Brain.[10] Deep learning also enhanced the existing field of reinforcement learning, led by researchers such as Richard Sutton, leading to the game-playing successes of systems developed by DeepMind.[11] In 2014, Ian Goodfellow published his paper on generative adversarial networks, which along with reinforcement learning has become the focus of much of the recent research in the field.[12] Continuing advances in AI capabilities have led to Stanford University's One Hundred Year Study on Artificial Intelligence, founded by Eric Horvitz, building on the long-standing research he and his colleagues have led at Microsoft Research. We have benefited from the input and guidance of many of these pioneers in our research over the past few years.

1   Warren McCulloch and Walter Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bulletin of Mathematical Biophysics*, volume 5, 1943.

2   Andrew Goldstein, "Bernard Widrow oral history," *IEEE Global History Network*, 1997.

3   Frank Rosenblatt, "The Perceptron: A probabilistic model for information storage and organization in the brain," *Psychological review*, volume 65, number 6, 1958.

4   Marvin Minsky and Seymour A. Papert, *Perceptrons: An introduction to computational geometry*, MIT Press, January 1969.

5   David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams, "Learning representations by back-propagating errors," Nature, volume 323, October 1986; for a discussion of Linnainmaa's role see Juergen Schmidhuber, *Who invented backpropagation?*, Blog post http://people.idsia.ch/~juergen/who-invented-backpropagation.html, 2014.

6   Yann LeCun, Patrick Haffner, Leon Botton, and Yoshua Bengio, *Object recognition with gradient-based learning*, Proceedings of the IEEE, November 1998.

7   John Hopfield, *Neural networkds and physical systems with emergent collective computational abilities*, PNAS, April 1982.

8   Sepp Hochreiter and Juergen Schmidhuber, "Long short-term memory," *Neural Computation*, volume 9, number 8, December 1997.

9   Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, *ImageNet classification with deep convolutional neural networks*, NIPS 12 proceedings of the 25th International Conference on Neural Information Processing Systems, December 2012.

10  Jeffrey Dean et al., *Large scale distributed deep networks*, NIPS 2012.

11  Richard S. Sutton and Andrew G. Barto, *Reinforcement learning: An introduction*, MIT Press, 1998.

12  Ian J. Goodfellow, *Generative adversarial networks*, ArXiv, June 2014.

Neural networks are a subset of machine learning techniques. Essentially, they are AI systems based on simulating connected "neural units," loosely modeling the way that neurons interact in the brain. Computational models inspired by neural connections have been studied since the 1940s and have returned to prominence as computer processing power has increased and large training data sets have been used to successfully analyze input data such as images, video, and speech. AI practitioners refer to these techniques as "deep learning," since neural networks have many ("deep") layers of simulated interconnected neurons. Before deep learning, neural networks often had only three to five layers and dozens of neurons; deep learning networks can have seven to ten or more layers, with simulated neurons numbering into the millions.

In this paper, we analyzed the applications and value of three neural network techniques:

- **Feed forward neural networks.** One of the most common types of artificial neural network. In this architecture, information moves in only one direction, forward, from the input layer, through the "hidden" layers, to the output layer. There are no loops in the network. The first single-neuron network was proposed in 1958 by AI pioneer Frank Rosenblatt. While the idea is not new, advances in computing power, training algorithms, and available data led to higher levels of performance than previously possible.

- **Recurrent neural networks (RNNs).** Artificial neural networks whose connections between neurons include loops, well-suited for processing sequences of inputs, which makes them highly effective in a wide range of applications, from handwriting, to texts, to speech recognition. In November 2016, Oxford University researchers reported that a system based on recurrent neural networks (and convolutional neural networks) had achieved 95 percent accuracy in reading lips, outperforming experienced human lip readers, who tested at 52 percent accuracy.

- **Convolutional neural networks (CNNs).** Artificial neural networks in which the connections between neural layers are inspired by the organization of the animal visual cortex, the portion of the brain that processes images, well suited for visual perception tasks.

We estimated the potential of those three deep neural network techniques to create value, as well as other machine learning techniques such as tree-based ensemble learning, classifiers, and clustering, and traditional analytics such as dimensionality reduction and regression.

For our use cases, we also considered two other techniques—generative adversarial networks (GANs) and reinforcement learning—but did not include them in our potential value assessment of AI, since they remain nascent techniques that are not yet widely applied in business contexts. However, as we note in the concluding section of this paper, they may have considerable relevance in the future.

- **Generative adversarial networks (GANs).** These usually use two neural networks contesting each other in a zero-sum game framework (thus "adversarial"). GANs can learn to mimic various distributions of data (for example text, speech, and images) and are therefore valuable in generating test datasets when these are not readily available.

- **Reinforcement learning.** This is a subfield of machine learning in which systems are trained by receiving virtual "rewards" or "punishments," essentially learning by trial and error. Google DeepMind has used reinforcement learning to develop systems that can play games, including video games and board games such as Go, better than human champions.

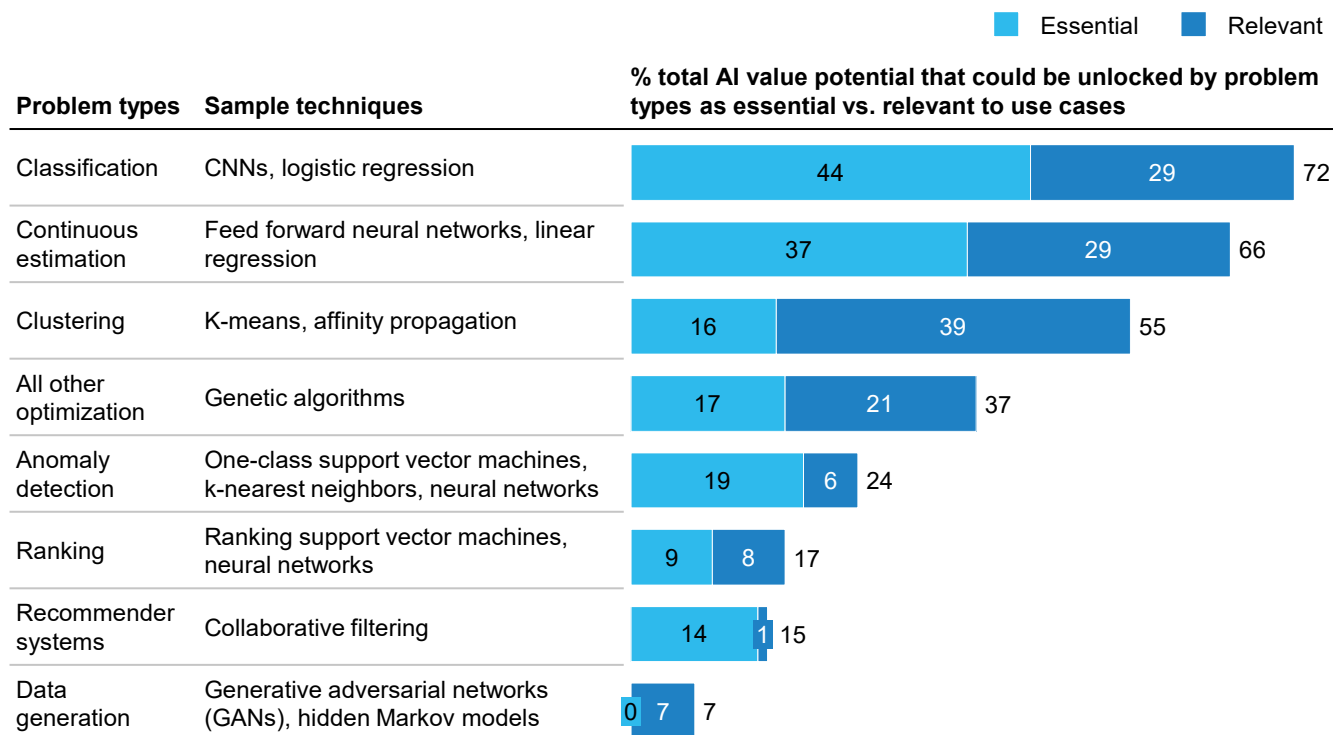## PROBLEM TYPES AND THE ANALYTIC TECHNIQUES THAT CAN BE APPLIED TO SOLVE THEM

In a business setting, those analytic techniques can be applied to solve real-life problems. For this research, we created a taxonomy of high-level problem types, characterized by the inputs, outputs, and purpose of each. A corresponding set of analytic techniques can be applied to solve these problems. These problem types include:

- **Classification.** Based on a set of training data, categorize new inputs as belonging to one of a set of categories. An example of classification is identifying whether an image contains a specific type of object, such as a truck or a car, or a product of acceptable quality coming from a manufacturing line.

- **Continuous estimation.** Based on a set of training data, estimate the next numeric value in a sequence. This type of problem is sometimes described as "prediction," particularly when it is applied to time series data. One example of continuous estimation is forecasting the sales demand for a product, based on a set of input data such as previous sales figures, consumer sentiment, and weather. Another example is predicting the price of real estate, such as a building, using data describing the property combined with photos of it.

- **Clustering.** These problems require a system to create a set of categories, for which individual data instances have a set of common or similar characteristics. An example of clustering is creating a set of consumer segments based on data about individual consumers, including demographics, preferences, and buyer behavior.

- **All other optimization.** These problems require a system to generate a set of outputs that optimize outcomes for a specific objective function (some of the other problem types can be considered types of optimization, so we describe these as "all other" optimization). Generating a route for a vehicle that creates the optimum combination of time and fuel use is an example of optimization.

- **Anomaly detection.** Given a training set of data, determine whether specific inputs are out of the ordinary. For instance, a system could be trained on a set of historical vibration data associated with the performance of an operating piece of machinery, and then determine whether a new vibration reading suggests that the machine is not operating normally. Note that anomaly detection can be considered a subcategory of classification.

- **Ranking.** Ranking algorithms are used most often in information retrieval problems in which the results of a query or request needs to be ordered by some criterion. Recommendation systems suggesting next product to buy use these types of algorithms as a final step, sorting suggestions by relevance, before presenting the results to the user.

- **Recommendations.** These systems provide recommendations, based on a set of training data. A common example of recommendations are systems that suggest the "next product to buy" for a customer, based on the buying patterns of similar individuals, and the observed behavior of the specific person.

- **Data generation.** These problems require a system to generate appropriately novel data based on training data. For instance, a music composition system might be used to generate new pieces of music in a particular style, after having been trained on pieces of music in that style.

Exhibit 2 illustrates the relative total value of these problem types across our database of use cases, along with some of the sample analytics techniques that can be used to solve each problem type. The most prevalent problem types are classification, continuous estimation, and clustering, suggesting that meeting the requirements and developing the capabilities in associated techniques could have the widest benefit. Some of the problem types that rank lower can be viewed as subcategories of other problem types—for example, anomaly detection is a special case of classification, while recommendations can be considered a type of optimization problem—and thus their associated capabilities could be even more relevant.

**Exhibit 2**

**Problem types and sample techniques**

Essential ▪ Relevant

**% total AI value potential that could be unlocked by problem types as essential vs. relevant to use cases**

| Problem types | Sample techniques | Essential | Relevant | Total |
|---|---|---|---|---|
| Classification | CNNs, logistic regression | 44 | 29 | 72 |
| Continuous estimation | Feed forward neural networks, linear regression | 37 | 29 | 66 |
| Clustering | K-means, affinity propagation | 16 | 39 | 55 |
| All other optimization | Genetic algorithms | 17 | 21 | 37 |
| Anomaly detection | One-class support vector machines, k-nearest neighbors, neural networks | 19 | 6 | 24 |
| Ranking | Ranking support vector machines, neural networks | 9 | 8 | 17 |
| Recommender systems | Collaborative filtering | 14 | 1 | 15 |
| Data generation | Generative adversarial networks (GANs), hidden Markov models | 0 | 7 | 7 |

NOTE: Sample techniques include traditional analytical techniques, machine learning, and the deep learning techniques we describe in this paper as AI. Numbers may not sum due to rounding.

SOURCE: McKinsey Global Institute analysis

# 2. INSIGHTS FROM USE CASES

We collated and analyzed more than 400 use cases across 19 industries and nine business functions for this research (see Box 2, "Our AI use cases"). They provided considerable insight into the areas within specific sectors where deep neural networks can potentially create the most value, the incremental performance lift that these neural networks can generate compared with more traditional analytics, and the voracious data requirements—in terms of volume, variety, and velocity—that must be met for this potential to be realized.

Examples of where AI can be used to improve the performance of existing use cases include:

- **Predictive maintenance: the power of machine learning to detect anomalies.** Some existing predictive maintenance systems have analyzed time series data from Internet of Things (IoT) sensors, such as those monitoring temperature or vibration, in order to detect anomalies or make forecasts on the remaining useful life of components. Deep learning's capacity to analyze very large amounts of high dimensional data can take this to a new level. By layering in additional data, such as audio and image data, from other sensors—including relatively cheap ones such as microphones and cameras—neural networks can enhance and possibly replace more traditional methods. AI's ability to predict failures and allow planned interventions can be used to reduce downtime and operating costs while improving production yield. In some industry examples we studied,

## Box 2. Our AI use cases

We define a "use case" as a targeted application of digital technologies to a specific business challenge, with a measurable outcome. We recognize that use cases can be described at different levels of granularity. The use cases we analyzed for this paper correspond to descriptions of specific business challenges which practitioners in the industries and sectors we studied acknowledged as meaningful. For example, recommending the "next product to buy" for e-commerce in the retail industry was considered a use case, whereas "marketing and sales" was not sufficiently granular to be considered a use case, even though it is a relevant business function. We also collated similar use cases across sectors into "domains" (see related section below).

For this research we built a library of use cases across the economy that is as comprehensive as possible. The cases that we identified come from a variety of sources, including thousands of engagements by McKinsey Analytics with clients around the globe. The data incorporate findings from real-life examples of companies and public-sector organizations using a range of analytics techniques, both AI and more traditional approaches, as well as the potential application of these techniques in situations similar to those in which they have already been successfully deployed. For example, where we found a use case in one sector, such as a pricing promotion use case in travel, we looked to see if it could be applied in

other sectors, for example in retail. Where possible, we identified and analyzed multiple instances of individual use cases.

For each use case, we estimated the annual value potential of applying AI and other analytics across the entire economy. This value potential could be captured by companies and organizations themselves, in the form of increased profits, or by their customers, in the form of lower prices or higher quality. For use cases that involve increasing revenue, such as those in sales and marketing, we estimated the economy-wide value potential in terms of the increased potential productivity of sales and marketing expenditures, assuming that overall annual spend in the economy is fixed in a given year (rather than estimating the impact of higher revenues, which would have assumed that overall spending would increase). Our estimates are based on the structure of the global economy in 2016. We did not estimate the value potential of creating entirely new product or service categories, such as autonomous driving.

Our library of use cases, while extensive, is not exhaustive, and thus it may overstate or understate the potential for certain sectors. We realize it may also contain some biases based on the business profile of clients. Our library of examples is a living one and we will continue refining and adding to it.

using remote on-board diagnostics to anticipate the need for service could theoretically generate value of 1 to 2 percent of total sales. In a case involving cargo aircraft, AI can extend the life of the plane beyond what is possible using traditional analytic techniques by combining plane model data, maintenance history, IoT sensor data such as anomaly detection on engine vibration data, and images and video of engine condition.

- **AI-driven logistics optimization can reduce costs through real-time forecasts and behavioral coaching.** Application of AI techniques such as continuous estimation to logistics can add substantial value across many sectors. AI can optimize routing of delivery traffic, thereby improving fuel efficiency and reducing delivery times. One European trucking company has reduced fuel costs by 15 percent, for example. By using sensors that monitor both vehicle performance and driver behavior, drivers receive real-time coaching, including when to speed up or slow down, optimizing fuel consumption and reducing maintenance costs. In another example, an airline uses AI to predict congestion and weather-related problems in order to avoid costly cancellations. For an airline with 100,000 flights per day, a 1 percent reduction in cancellations can make a material difference

- **AI can be a valuable tool for customer service management and personalized marketing.** Improved speech recognition in call center management and call routing by applying AI techniques allow a more seamless experience for customers—and more efficient processing. The capabilities go beyond words alone. For example, deep learning analysis of audio allows systems to assess customers' emotional tone; if a customer is responding negatively to an automated system, the call can be rerouted to human operators and managers. In other areas of marketing and sales, AI techniques can also have a significant impact. "Next product to buy" recommendations that target individual customers—as companies such as Amazon and Netflix have successfully implemented—can lead to a substantial increase in the rate of sales conversions. Individual pricing and promotion also has a broad application across sectors. For example, in auto insurance, premiums can be customized based on data about driving patterns and distances driven. In one use case, a travel company that micro-segmented customers built a 360-degree customer view and offered additional services, such as hotels and airlines, using a recommender system algorithm trained on product and customer data. This led to a 10 to 15 percent increase in ancillary revenue. In retail, AI can use SKU-performance data to optimize weekly and thematic product promotions, including giving daily promotional recommendations.

### MAPPING ANALYTICS TECHNIQUES TO USE CASES

For each use case, we catalogued the specific analytical techniques that could be applied, including traditional analytics techniques, and various forms of machine learning, including deep learning. We found that the applicability of different techniques varied across sectors and business functions. Exhibits 3 and 4 are heatmaps that show the extent to which applicable techniques can be used by industry and function, based on our library of use cases. A select set of traditional techniques, including clustering, regression, and tree-based models, are broadly applicable across many industries and functions. Similarly, the more advanced AI techniques, such as RNNs, CNNs, and feed forward neural networks, are relevant in many of the same industries and functions. The variations in the heat maps relate to the different types of problems to be solved in the use cases, and the most applicable techniques with which to do that.

Exhibit 3

**Heat map: Technique relevance to industries**

Number of use cases    Low ▢▢▢▢■ High

| | Focus of report | | | | | Traditional analytics techniques | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Reinforcement learning | Feed forward networks | Recurrent neural networks | Convolutional neural networks | Generative adversarial networks | Tree-based ensemble learning | Dimensionality reduction | Classifiers | Clustering | Regression analysis | Statistical inference | Monte Carlo | Markov processes | Other optimization |
| Advanced electronics/ semiconductors | | | | | | | | | | | | | | |
| Aerospace and defense | | | | | | | | | | | | | | |
| Agriculture | | | | | | | | | | | | | | |
| Automotive and assembly | | | | | | | | | | | | | | |
| Banking | | | | | | | | | | | | | | |
| Basic materials | | | | | | | | | | | | | | |
| Chemicals | | | | | | | | | | | | | | |
| Consumer packaged goods | | | | | | | | | | | | | | |
| Health-care systems and services | | | | | | | | | | | | | | |
| High tech | | | | | | | | | | | | | | |
| Insurance | | | | | | | | | | | | | | |
| Media and entertainment | | | | | | | | | | | | | | |
| Oil and gas | | | | | | | | | | | | | | |
| Pharmaceuticals and medical products | | | | | | | | | | | | | | |
| Public and social sector | | | | | | | | | | | | | | |
| Retail | | | | | | | | | | | | | | |
| Telecommunications | | | | | | | | | | | | | | |
| Transport and logistics | | | | | | | | | | | | | | |
| Travel | | | | | | | | | | | | | | |

SOURCE: McKinsey Global Institute analysis

Exhibit 4

**Heat map: Technique relevance to functions**

Number of use cases    Low ▢▢▢ High

| | Focus of report | | | | | Traditional analytics techniques | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Reinforcement learning | Feed forward networks | Recurrent neural networks | Convolutional neural networks | Generative adversarial networks | Tree-based ensemble learning | Dimensionality reduction | Classifiers | Clustering | Regression analysis | Statistical inference | Monte Carlo | Markov processes | Other optimization |
| Finance and IT | | | | | | | | | | | | | | |
| Human resources | | | | | | | | | | | | | | |
| Marketing and sales | | | | | | | | | | | | | | |
| Other operations | | | | | | | | | | | | | | |
| Product development | | | | | | | | | | | | | | |
| Risk | | | | | | | | | | | | | | |
| Service operations | | | | | | | | | | | | | | |
| Strategy and corporate finance | | | | | | | | | | | | | | |
| Supply-chain management and manufacturing | | | | | | | | | | | | | | |

SOURCE: McKinsey Global Institute analysis

Use cases are often associated with multiple problem types: in fact, solving an average of three problem types is either required for or applicable to a use case. Truck route optimization provides one example. While an optimization solution is required to generate most of the value by ensuring that vehicles take the most efficient routes, point estimation could add incremental value. For example, this could be a model that predicts how long a traffic light will stay red. With this knowledge, the route optimizer can determine whether to suggest speeding up or slowing down to minimize vehicle stop time, which can be costly from a fuel consumption standpoint. Thus, we catalogued problem types that are of primary importance, and others that are relevant, for individual use cases.

The techniques with the broadest scope of applicability across industries and functions include traditional analytic techniques, such as regression, tree-based ensemble learning, classifiers, clustering, and other forms of statistical inference. That said, the neural network-based techniques that we identify with the current generation of AI. AI also demonstrates wide potential applicability across sectors and functions. However, their usage is not yet widespread, partly because of the relative immaturity of the technology and the organizational challenges of deploying these techniques. Among business functions, these techniques are for now mostly to be found in marketing and sales and in supply-chain management and manufacturing. In particular, feed forward neural networks feature as the main technique deployed in these two functions.

We have also collected individual use cases into broader "domains" that address closely related problems. For example, the predictive maintenance domain aggregates use cases from across multiple sectors and types of machinery in which the use of analytics and AI can help to identify when repairs can be made to prevent a piece of equipment from breaking down (Exhibit 5).

Overall, a deep understanding of use cases and how they are associated with particular problem types, analytical techniques, and data types can help guide organizations regarding where to invest in the technical capabilities and data that will provide the greatest impact.

**Exhibit 5**

**Use case mapping to data types**

Number of use cases    Low ▇▇ High

| Use case domains | Structured/ semi-structured | Time series | Text | Audio | Video | Image |
|---|---|---|---|---|---|---|
| Analytics-driven accounting and IT | | | | | | |
| Analytics-driven hiring and retention | | | | | | |
| Channel management | | | | | | |
| Churn reduction | | | | | | |
| Customer acquisition/lead generation | | | | | | |
| Customer service management | | | | | | |
| Fraud and debt analytics | | | | | | |
| Inventory and parts optimization | | | | | | |
| Logistics network and warehouse optimization | | | | | | |
| Marketing budget allocation | | | | | | |
| Next product to buy/individualized offering | | | | | | |
| Predictive maintenance | | | | | | |
| Predictive service/intervention | | | | | | |
| Pricing and promotion | | | | | | |
| Procurement and spend analytics | | | | | | |
| Product development cycle optimization | | | | | | |
| Product feature optimization | | | | | | |
| Risk modeling | | | | | | |
| Sales and demand forecasting | | | | | | |
| Smart capital expenditures | | | | | | |
| Task automation | | | | | | |
| Workforce productivity and efficiency | | | | | | |
| Yield optimization | | | | | | |

SOURCE: McKinsey Global Institute analysis

## TWO-THIRDS OF THE OPPORTUNITIES TO USE AI ARE IN IMPROVING THE PERFORMANCE OF EXISTING ANALYTICS USE CASES

In 69 percent of the use cases we studied, deep neural networks can be used to improve performance beyond that provided by other analytic techniques. Cases in which only neural networks can be used, which we refer to here as "greenfield" cases, constituted just 16 percent of the total. For the remaining 15 percent, artificial neural networks provided limited additional performance over other analytics techniques, among other reasons because of data limitations that made these cases unsuitable for deep learning.

Greenfield AI solutions are prevalent in business areas such as customer service management, as well as among some industries in which the data are rich and voluminous and at times integrate human reactions. A key differentiator that often underpins higher AI value potential is the possibility of applying large amounts of audio, video, image, and text data to these problems. Among industries, we found many greenfield use cases in health care, in particular. Some of these cases involve disease diagnosis and improved care, and rely on rich data sets incorporating image and video inputs, including from MRIs.

On average, our use cases suggest that modern deep learning AI techniques have the potential to provide a boost in value above and beyond traditional analytics techniques ranging from 30 percent to 128 percent, depending on industry (Exhibit 6).

In many of our use cases, however, traditional analytics and machine learning techniques continue to underpin a large percentage of the value creation potential in industries including insurance, pharmaceuticals and medical products, and telecommunications, with the potential of AI limited in certain contexts. In part this is due to the way data are used by those industries and to regulatory issues, as we discuss later in this paper.

## DATA REQUIREMENTS FOR DEEP LEARNING ARE SUBSTANTIALLY GREATER THAN FOR OTHER ANALYTICS, IN TERMS OF BOTH VOLUME AND VARIETY

Making effective use of neural networks in most applications requires large labeled training data sets alongside access to sufficient computing infrastructure. As the size of the training data set increases, the performance of traditional techniques tends to plateau in many cases. However, the performance of advanced AI techniques using deep neural networks, configured and trained effectively, tends to increase. Furthermore, these deep learning techniques are particularly powerful in extracting patterns from complex, multi-dimensional data types such as images, video, and audio or speech. The data will need to be collected in a way that addresses society's concerns about issues of privacy.

### Data volume is essential for neural networks to achieve a high level of accuracy in training algorithms

Deep learning methods require thousands of data records for models to become relatively good at classification tasks and, in some cases, millions for them to perform at the level of humans. By one estimate, a supervised deep-learning algorithm will generally achieve acceptable performance with around 5,000 labeled examples per category and will match or exceed human level performance when trained with a data set containing at least 10 million labeled examples.[7] In some cases in which advanced analytics is currently used, so much data are available—millions or even billions of rows per data set—that AI usage is the most appropriate technique. However, if a threshold of data volume is not reached, AI may not add value to traditional analytics techniques.
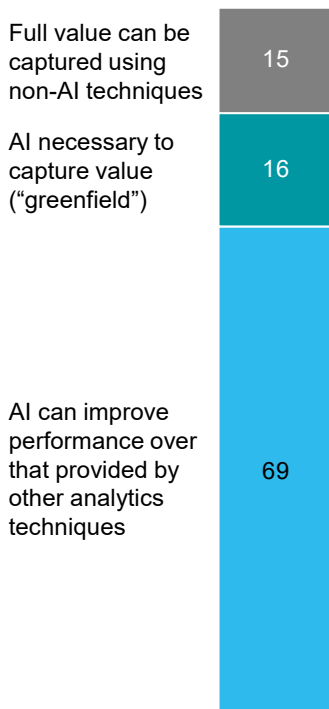
---

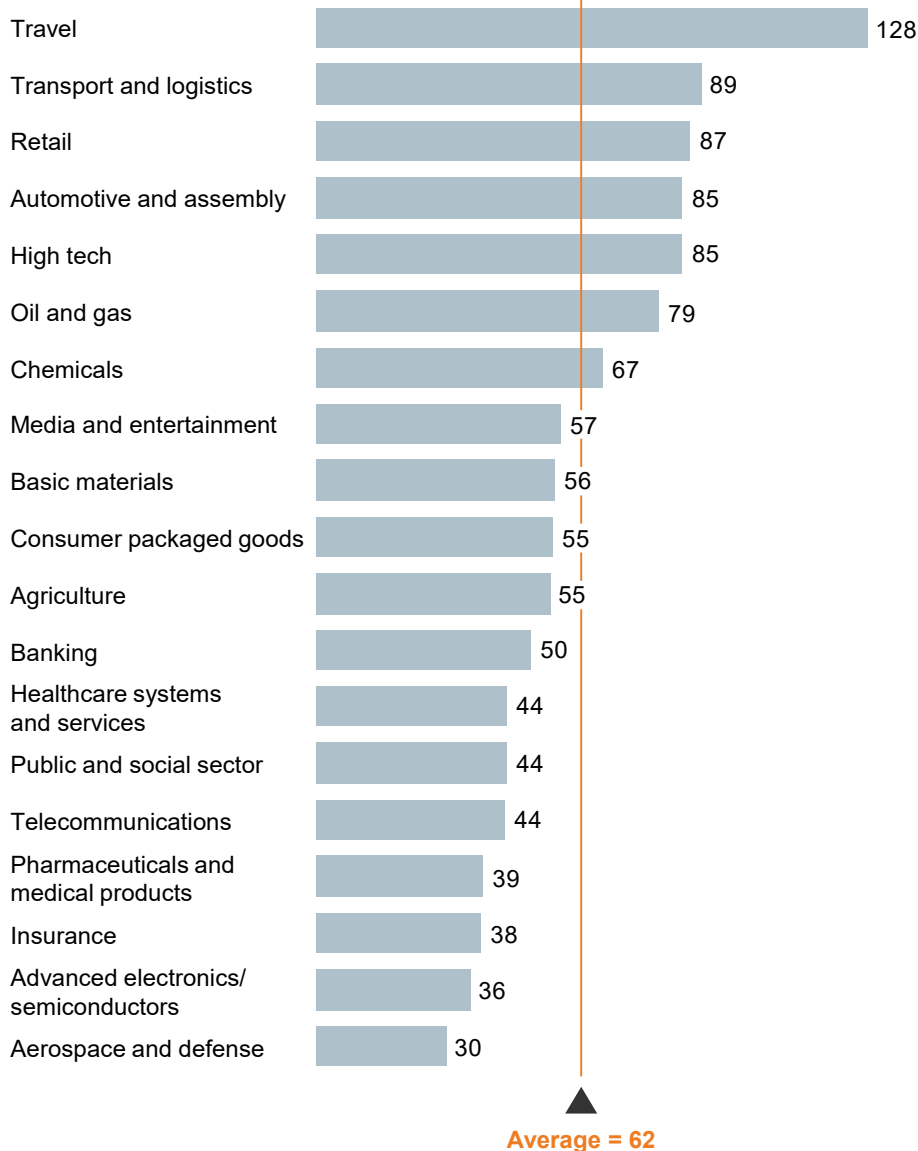[7]   Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep learning*, MIT Press, 2016.

Exhibit 6

**In more than two-thirds of our use cases, AI can improve performance beyond that provided by other analytics techniques**

%

**Breakdown of use cases by applicable techniques**

| | |
|---|---|
| Full value can be captured using non-AI techniques | 15 |
| AI necessary to capture value ("greenfield") | 16 |
| AI can improve performance over that provided by other analytics techniques | 69 |

**Potential incremental value from AI over other analytics techniques**

| | |
|---|---|
| Travel | 128 |
| Transport and logistics | 89 |
| Retail | 87 |
| Automotive and assembly | 85 |
| High tech | 85 |
| Oil and gas | 79 |
| Chemicals | 67 |
| Media and entertainment | 57 |
| Basic materials | 56 |
| Consumer packaged goods | 55 |
| Agriculture | 55 |
| Banking | 50 |
| Healthcare systems and services | 44 |
| Public and social sector | 44 |
| Telecommunications | 44 |
| Pharmaceuticals and medical products | 39 |
| Insurance | 38 |
| Advanced electronics/ semiconductors | 36 |
| Aerospace and defense | 30 |

**Average = 62**

NOTE: Numbers may not sum due to rounding.

SOURCE: McKinsey Global Institute analysis

These massive data sets can be difficult to obtain or create for many business use cases, and labeling remains a challenge. For example, teaching an autonomous vehicle to navigate urban traffic requires enormous image data sets in which all critical objects (other vehicles of all types, traffic signs, pedestrians, road markings, and so on) are labeled under all weather and lighting conditions. Most current AI models are trained through "supervised learning", which requires humans to label and categorize the underlying data. However promising new techniques are emerging to overcome these data bottlenecks, such as reinforcement learning, generative adversarial networks, transfer learning, and one-shot learning.[8] Unlike reinforcement learning, in which systems learn to perform tasks by trial and error, one-shot learning allows a trained AI model to learn about a subject based on a small number of real-world demonstrations or examples—and sometimes just one.

Nonetheless, our research shows that almost three-quarters of the impact from advanced analytics is tied to use cases requiring millions of labeled data examples. This means that organizations will have to adopt and implement strategies that enable them to collect, integrate, and process data at scale. They will need to ensure that as much of their data as possible is captured electronically, typically by progressing in their overall digitization journeys. Even with large datasets, they will have to guard against "overfitting," in which a model too tightly matches the "noisy" or random features of the training set, resulting in a corresponding lack of accuracy in future performance, and against "underfitting," in which the model fails to capture all of the relevant features. Linking data across customer segments and channels and, where possible, to production data, rather than allowing different sets of data to languish in silos, is especially important to create value. To achieve this linkage, companies will need to create meta-data models, and manage both the internal challenges as well as the regulatory risks of sharing the data across a range of business functions.

### Realizing AI's full potential requires a diverse range of data types including images, video, and audio

Neural AI techniques excel at analyzing image, video, and audio data types because of their complex, multi-dimensional nature, known by practitioners as "high dimensionality." Just a single training instance can have many different features, often requiring multiple levels of analysis. These systems can thus require orders of magnitude more training data than other machine learning systems with lower dimensionality data.

However, while images, video, and audio data are particularly well-suited for use with modern deep learning techniques, even more value can be extracted from mining insights from traditional structured and time series data types. Exhibit 7 highlights the range of value from our use cases based on the types of data used.

In the past few years, advances in deep learning have pushed AI-driven image classification tasks beyond human performance in some cases. An image of a face, for example, consists of thousands of pixels that can be combined to form eyes, ears, and other features. When arranged in the right way, they form the face of a specific person. Neural networks are good at dealing with high dimensionality, as multiple layers in a network can learn to represent the many different features present in the data. Thus, for facial recognition, the first layer in the network could focus on raw pixels, the next on edges and lines, another on generic facial features, and the final layer might identify the face. Unlike previous generations of AI, which often required human expertise to do "feature engineering," these neural network techniques are often able to learn to represent these features in their simulated neural networks as part of the training process.
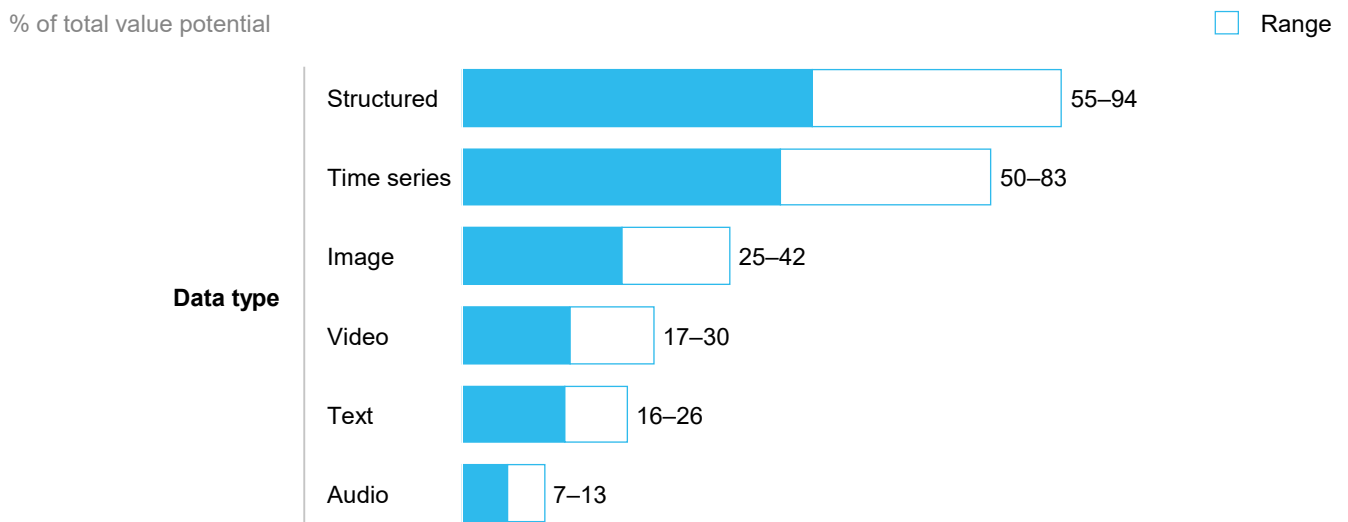
An example of the progress that has been made can be found in the results of the annual ImageNet Large Scale Visual Recognition Challenge (Exhibit 8).[9] Since 2010 machine learning researchers have competed by submitting algorithms for detecting objects within a public database of 14 million labeled images. The increase in accuracy observed in 2012 is widely considered the harbinger of the deep learning revolution; for the first time, a deep neural net was used to address the problem and showed dramatic improvement over previous efforts. The accuracy of the best-performing algorithms (all now using deep learning) now exceeds "human level performance"—the accuracy level expected of a human performing the same task.

---

Exhibit 7

**Range of potential AI value impact by data type**

% of total value potential                                      ☐ Range

**Data type**

| Structured | 55–94 |
| Time series | 50–83 |
| Image | 25–42 |
| Video | 17–30 |
| Text | 16–26 |
| Audio | 7–13 |

SOURCE: McKinsey Global Institute analysis

---

Exhibit 8

**The ability of AI systems to recognize objects has improved markedly to the point where the best systems now exceed human performance**

**Accuracy**
%
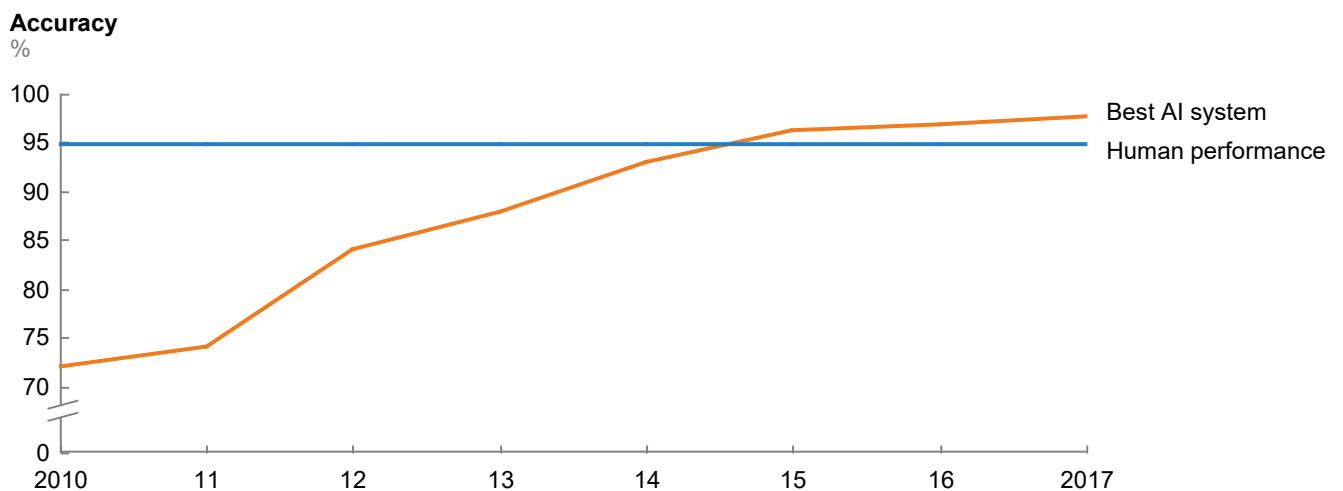


SOURCE: ImageNet Large Scale Visual Recognition Challenge; McKinsey Global Institute analysis

---

[9]  Olga Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision*, volume 115, issue 3, December 2015.

These increases in performance using deep learning techniques enabled many of the consumer products we already take for granted, including services such as Siri, Alexa, and Cortana. Improved image-processing technology underpins some of the capabilities that are being tested in self-driving cars today.

**Ongoing data acquisition for retraining AI systems is necessary; one out of three use cases requires model refreshes at least monthly and sometimes daily**
An analysis of our use cases shows that, along with issues around the volume and variety of data, velocity is also a requirement: AI techniques require models to be retrained to match potential changing conditions, so the training data must be refreshed frequently. In one-third of the cases, the model needs to be refreshed at least monthly, and almost one in four cases requires a daily refresh; this is especially true in marketing and sales and in supply-chain management and manufacturing (Exhibit 9).

---

Exhibit 9

**For about one-third of use cases, the models require frequent updating: three-quarters of those cases require monthly refreshes, while nearly one-quarter are at least weekly**

Share of use cases

Frequency of refresh required



At least monthly refreshes — 34

Less frequent refreshes — 66

| At least monthly | At least weekly |
| --- | --- |
| 77 | 23 |
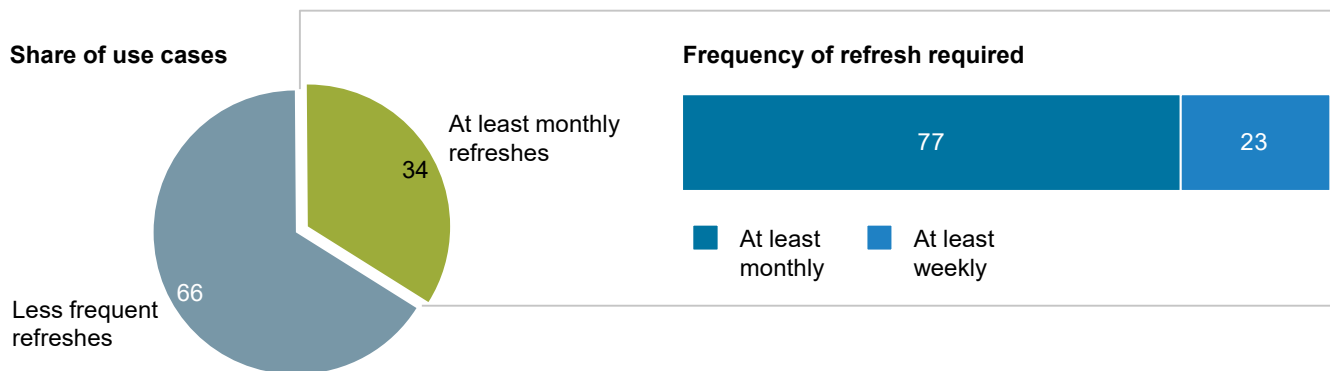
NOTE: Numbers may not sum due to rounding.

SOURCE: McKinsey Global Institute analysis

---

# 3. SIZING THE POTENTIAL VALUE OF AI

An analysis of the value derived from our use cases suggests that AI can generate considerable value. We estimate that the AI techniques we cite in this report—feed forward neural networks and convolutional neural networks—together have the potential to create between $3.5 trillion and $5.8 trillion in value annually across nine business functions in 19 industries. The value as measured by percentage of industry revenue varies significantly among industries, depending on the specific applicable use cases, the availability of abundant and complex data, and regulatory and other constraints. While we found patterns in the potential of AI to create value within industries and functions, every company will need to look at the details of its business to estimate the opportunities to create value enabled by AI.

## AI COULD POTENTIALLY CREATE $3.5 TRILLION TO $5.8 TRILLION IN ANNUAL VALUE IN THE GLOBAL ECONOMY

We estimated a range of annual value to the global economy for both AI and other analytics techniques, based on the value we observed being created in current use cases and the potential value in projected future ones. The total annual value potential of AI alone across 19 industries and nine business functions in the global economy came to between $3.5 trillion and $5.8 trillion. This constitutes about 40 percent of the overall $9.5 trillion to $15.4 trillion annual impact that could potentially be enabled by all analytical techniques (Exhibit 10).

Per industry, we estimate that AI's potential value amounts to between 1 and 9 percent of 2016 revenue. Even the industry with the smallest potential value at stake, aerospace and defense (less than $50 billion) could enable the annual creation of value that is equivalent to the GDP of Lebanon.

These figures are not forecasts for a particular period in time, but they are indicative of the considerable potential for the global economy that advanced analytics represents. The percentage of this potential that individual organizations, sectors, and functions can achieve will be the product of competitive and market dynamics, as well as of many choices and decisions—and indeed business model choices—by organizations and others. They include the developers of technology, policy makers who set the overall context, and customers who choose what to purchase. Some of this value will be captured in a variety of ways, for example it may result in more valued products and services, revenue growth, cost savings, or indeed consumer surplus. While the aggregate numbers may appear modest, in some use cases the advancements amount to radical transformation.

For AI to realize a substantial portion of the potential value we have estimated will require companies to deploy these techniques comprehensively in areas where they can most effectively harness their ability to make the complexity of data an advantage. Take the travel industry (by which we refer to all aspects of commercial passenger travel, from travel agents and airlines to hotels and online providers), in which the potential AI impact can more than double what is achievable using traditional analytic methods, amounting to between 7 and almost 12 percent of total revenue for the industry. To achieve that potential will require AI deployment in top-line revenue-related marketing and sales use cases, and bottom-line operations use cases, where in both cases the data are rich and the value of each incremental percentage point performance increase is significant. As an example of one single AI use case of the many that would be required for this industry to achieve its full potential, Hawaii's state tourism authority, working with a major online travel company, uses facial recognition software to monitor travelers' expressions through their computer webcams and deliver personalized offers.[10]
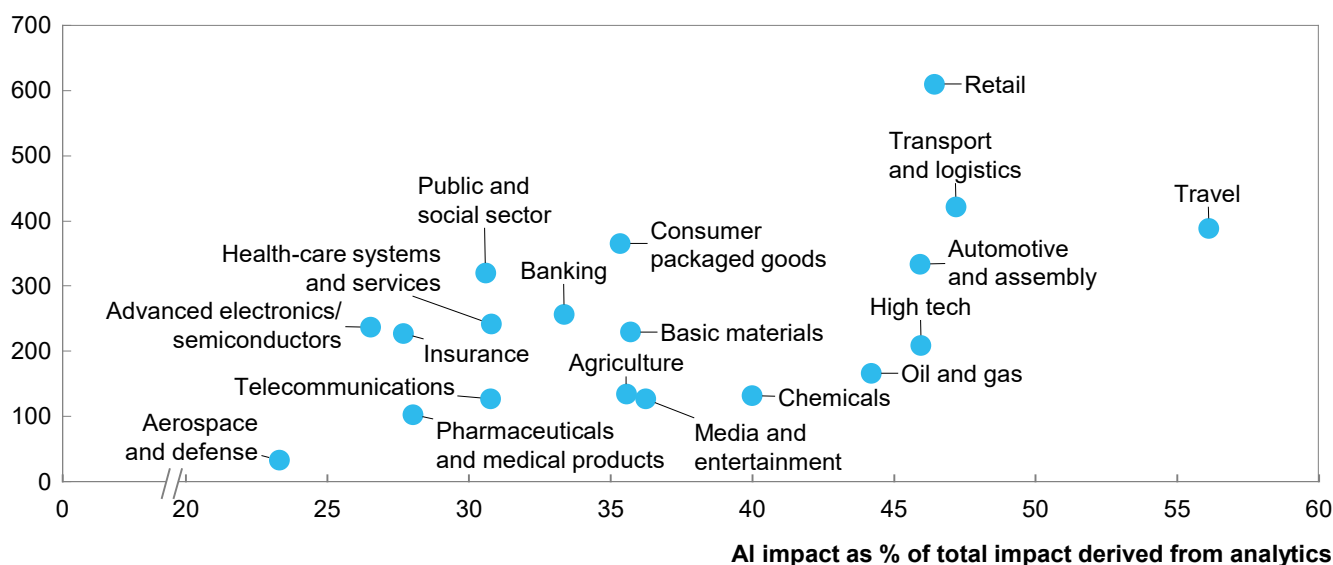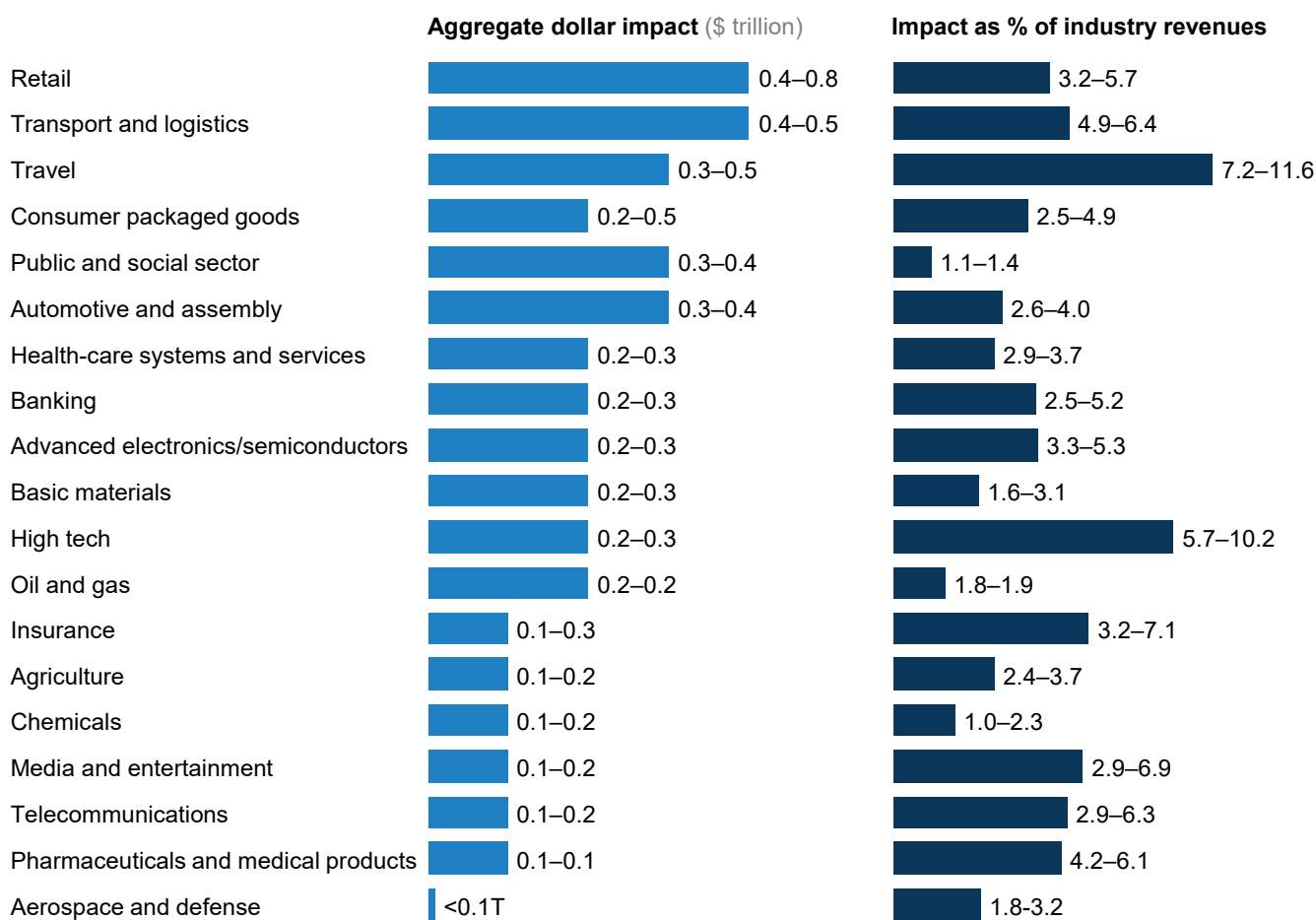
---

**Exhibit 10**

**AI has the potential to create annual value across sectors totaling $3.5 trillion to $5.8 trillion, or 40 percent of the overall potential impact from all analytics techniques**

**AI impact**
$ billion



AI impact as % of total impact derived from analytics

**The potential value of AI by sector**

| | **Aggregate dollar impact** ($ trillion) | **Impact as % of industry revenues** |
|---|---|---|
| Retail | 0.4–0.8 | 3.2–5.7 |
| Transport and logistics | 0.4–0.5 | 4.9–6.4 |
| Travel | 0.3–0.5 | 7.2–11.6 |
| Consumer packaged goods | 0.2–0.5 | 2.5–4.9 |
| Public and social sector | 0.3–0.4 | 1.1–1.4 |
| Automotive and assembly | 0.3–0.4 | 2.6–4.0 |
| Health-care systems and services | 0.2–0.3 | 2.9–3.7 |
| Banking | 0.2–0.3 | 2.5–5.2 |
| Advanced electronics/semiconductors | 0.2–0.3 | 3.3–5.3 |
| Basic materials | 0.2–0.3 | 1.6–3.1 |
| High tech | 0.2–0.3 | 5.7–10.2 |
| Oil and gas | 0.2–0.2 | 1.8–1.9 |
| Insurance | 0.1–0.3 | 3.2–7.1 |
| Agriculture | 0.1–0.2 | 2.4–3.7 |
| Chemicals | 0.1–0.2 | 1.0–2.3 |
| Media and entertainment | 0.1–0.2 | 2.9–6.9 |
| Telecommunications | 0.1–0.2 | 2.9–6.3 |
| Pharmaceuticals and medical products | 0.1–0.1 | 4.2–6.1 |
| Aerospace and defense | <0.1T | 1.8-3.2 |

NOTE: Artificial Intelligence here includes neural networks only. Numbers may not sum due to rounding.

SOURCE: McKinsey Global Institute analysis

By contrast, as a percentage of overall revenues, the potential annual value of the AI use cases we identified in telecommunications is between about 3 and 6 percent of industry revenue (still amounting to over $100 billion in potential annual impact). While telecommunications operators have a large volume of customer data, much of that data is in structured forms for which firms can just as easily use more conventional analytics and machine learning techniques to leverage that data. The public sector also has a large volume of data, and a range of data types and use cases that make it a ripe area for AI applications. However, regulations and requirements for data privacy and interpretability, particularly in developed markets, present significant challenges to its use, and thus constrain the value potential.

Even with those caveats, industries for which we have estimated lower AI impact in terms of percentages of overall revenue compared with other sectors, such as aerospace, defense, and the public sector, nonetheless have the potential to create billions or even hundreds of billions of dollars of value from its deployment where appropriate. A key benefit of this exercise across industries is to provide executives and leaders in each industry with a perspective for where the largest potential opportunities lie. This will help them set priorities as well as identifying where they can benefit from AI.

## THE BIGGEST VALUE OPPORTUNITIES FOR AI ARE IN MARKETING AND SALES AND IN SUPPLY-CHAIN MANAGEMENT AND MANUFACTURING

From the use cases we have examined, we find that the greatest potential value impact from using AI are both in top-line-oriented functions, such as marketing and sales, and in bottom-line-oriented operational functions, including supply-chain management and manufacturing (Exhibit 11). At a company level, every firm will need to examine its mix of functions to find the most attractive opportunities to use AI, and determine where it makes most sense to invest in AI deployment.

From the analysis of our use cases we can see broad applications of AI in marketing and sales across industries including consumer packaged goods, banking, retail, telecommunications, high tech, travel, insurance, and media and entertainment. Indeed, marketing and sales and supply-chain management together constitute some of the biggest areas of opportunity for AI (Exhibit 12). As noted earlier, AI is a powerful tool for personalizing product recommendations including through analyzing aggregated user data to understand individual customer preferences. Companies will nonetheless still need to think through business models that involve data use in such cases.

Consumer industries such as retail and high tech will tend to see more potential from marketing and sales AI applications because frequent and digital interactions between business and customers generate larger data sets for AI techniques to tap into. E-commerce platforms, in particular, stand to benefit. This is because of the ease with which these platforms collect customer information, such as click data or time spent on a web page, and can then customize promotions, prices, and products for each customer dynamically and in real time. For their part, brick-and-mortar retailers can implement AI applications to improve product assortment and inventory management per store, and to optimize their supply chains end-to-end. Indeed, they can take advantage of the Internet of Things to generate data usable by AI techniques to both improve the performance of their supply chains and apply some of the top-line innovations from the online world to the offline world, for example, by viewing dwell time in front of a physical display as analogous to spending more time viewing web or mobile content.[11]

---

[11] See *The Internet of Things: Mapping the value beyond the hype*, McKinsey Global Institute, June 2015.
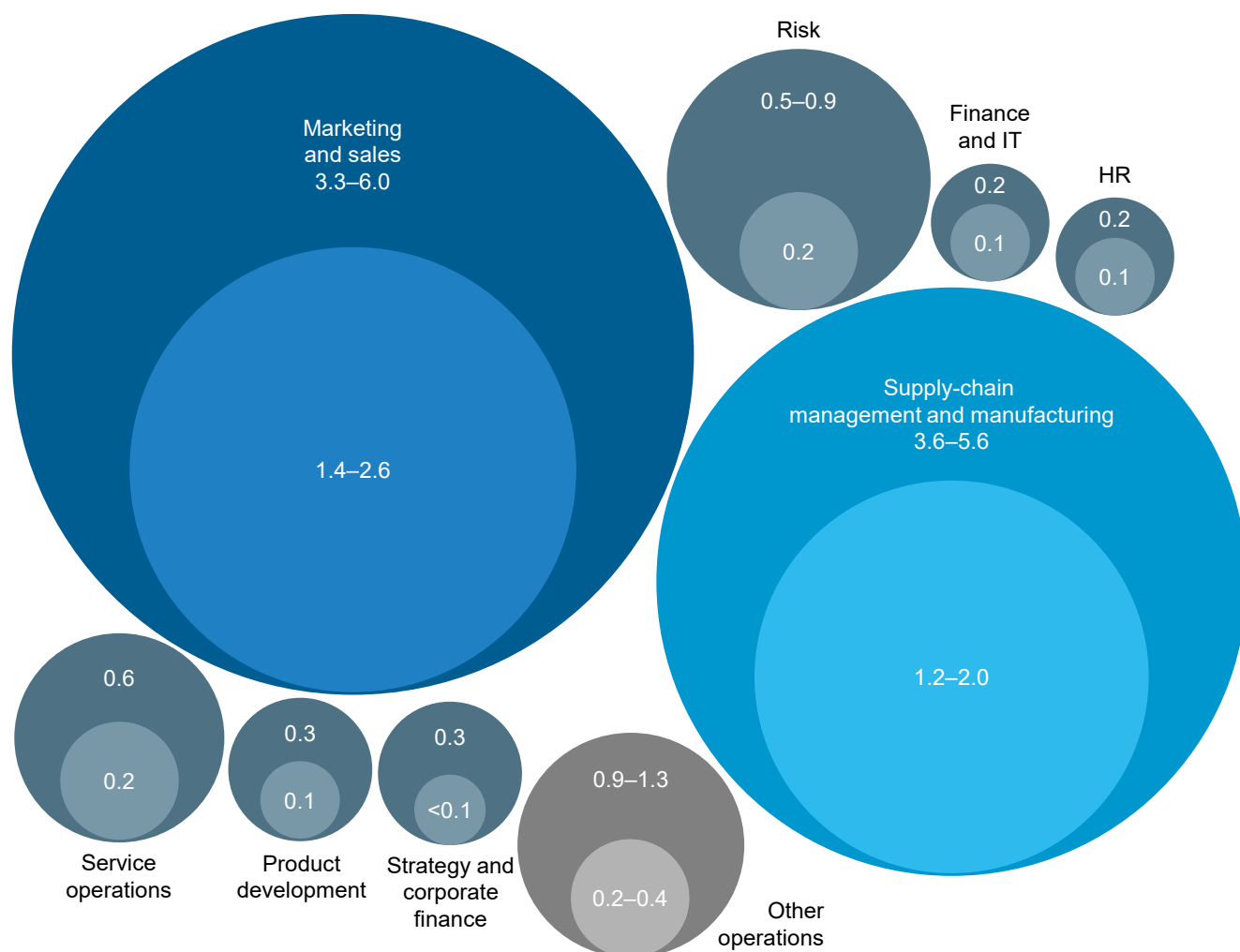
Exhibit 11

**AI's potential impact is greatest in marketing and sales and supply-chain management and manufacturing, based on our use cases**

Value potential
$ trillion

Value potential

**By all analytics** (darker color)
$9.5 trillion–15.4 trillion

**By AI** (lighter color)
$3.5 trillion–5.8 trillion

Risk
0.5–0.9
0.2

Finance
and IT
0.2
0.1

HR
0.2
0.1

Marketing
and sales
3.3–6.0

1.4–2.6

Supply-chain
management and manufacturing
3.6–5.6

1.2–2.0

0.6
0.2
Service
operations

0.3
0.1
Product
development

0.3
<0.1
Strategy and
corporate
finance

0.9–1.3
0.2–0.4
Other
operations

NOTE: Numbers may not sum due to rounding.

SOURCE: McKinsey Global Institute analysis

Exhibit 12

**Marketing and sales and supply-chain management and manufacturing are among the functions where AI can create the most incremental value**

Highest potential impact business problems per functional area
Impact size comparison by chart area)
$ trillion

**Marketing and sales**

**Supply-chain management and manufacturing**

**Other**

Customer service management
0.4–0.8

Next product to buy (NPTB) individual-ized offering
0.3–0.5

Price and promotion
0.3–0.5

Predictive maintenance
0.5–0.7

Yield optimization
0.3–0.6

Workforce productivity and efficiency
0.1–0.2

Predictive service/intervention
0.2

Task automation
0.1–0.2

Product develop-ment cycle
0.1

Fraud and debt analytics
0.1

Risk
0.1

Product feature optimiza-tion
<0.1

Smart capex
<0.1

Analytics-driven accounting and IT
0.1–0.2

Analytics-driven HR
0.1

Customer acquisition/lead generation
0.1–0.3

Churn reduction
0.1–0.2

Channel manage-ment
0.1–0.2

0.1

Procurement and spend analytics
0.1–0.2

Inventory and parts optimization
0.1–0.2

0.1

0.1

Marketing budget allocation

Logistics network and warehouse optimization

Sales and demand forecast

NOTE: Numbers may not sum due to rounding.

SOURCE: McKinsey Global Institute analysis

In our use cases, for example, we found that using real-time data to predict hyper-regional demand trends can increase sales by 0.25 percent to 0.75 percent, with margin improvements from lower waste and spoilage amounting to as much as half of one percentage point of sales. The impact can be considerably larger in pharmaceutical and medical products, in which predicting hyper-regional product demand and relevant health trends to inform inventory levels and reduce spoilage has the potential to raise sales by 5 to 10 percent.

AI's ability to conduct preventive maintenance and field force scheduling, as well as optimizing production and assembly processes, means that it also has considerable application possibilities and value potential across sectors including advanced electronics and semiconductors, automotive and assembly, chemicals, basic materials, transportation and logistics, oil and gas, pharmaceuticals and medical products, aerospace and defense, agriculture, and consumer packaged goods. In advanced electronics and semiconductors, for example, harnessing data to adjust production and supply-chain operations can minimize spending on utilities and raw materials, cutting overall production costs by 5 to 10 percent in our use cases.

## THE OPPORTUNITIES TO CREATE VALUE WITH AI CORRESPOND TO THE VALUE DRIVERS WITHIN INDUSTRIES

Identifying the opportunity for AI deployment depends on a range of factors that are specific to individual sectors and, within those sectors, to different types of businesses. Given the wide range of applicability of AI techniques, broadly speaking, if you want to know where AI can create the most value, you need to follow the money. For industries in which the main drivers of value are related to marketing and sales, including many consumer-facing industries, that is where the greatest value from deploying AI can be found. However, for industries in which the key driver of value is operational excellence, such as in advanced manufacturing and oil and gas, functions such as supply chain and manufacturing are those in which AI can create the most value.

It is instructive to compare industry sectors in terms of where AI has the greatest value potential. A detailed breakdown by sector is available on our interactive chart online. Here, by way of illustration, we take a snapshot of three sectors—retail, consumer packaged goods, and banking—where we have seen AI's impact (Exhibits 13–15). In retail, marketing and sales is the area with the most significant potential value from AI, and within that function, pricing and promotion and customer service management are the main value areas. Our use cases show that using customer data to personalize promotions, for example, including tailoring individual offers every day, can lead to a 1 to 2 percent increase in incremental sales for brick-and-mortar retailers alone. In packaged goods, supply-chain management is the key function that could benefit from AI deployment. Among the examples in our use cases, we see how forecasting based on underlying causal drivers of demand rather than prior outcomes can improve forecasting accuracy by 10 to 20 percent, which translates into a potential 5 percent reduction in inventory costs and revenue increases of 2 to 3 percent. In banking, particularly retail banking, AI has significant value potential in marketing and sales, much as it does in retail. Although the banking industry has been among the fastest to adopt AI and capture the value in risk management, the benefits could be even greater for other industries. For example, AI is helping insurers price risk more accurately, pharmaceutical companies are using AI to reduce risk in clinical trials, and mining firms have deployed the technologies to anticipate disruptions to production.

Exhibit 13

**In retail, AI has the most potential impact in pricing and promotion and other marketing and sales areas**

Retail examples
$ trillion
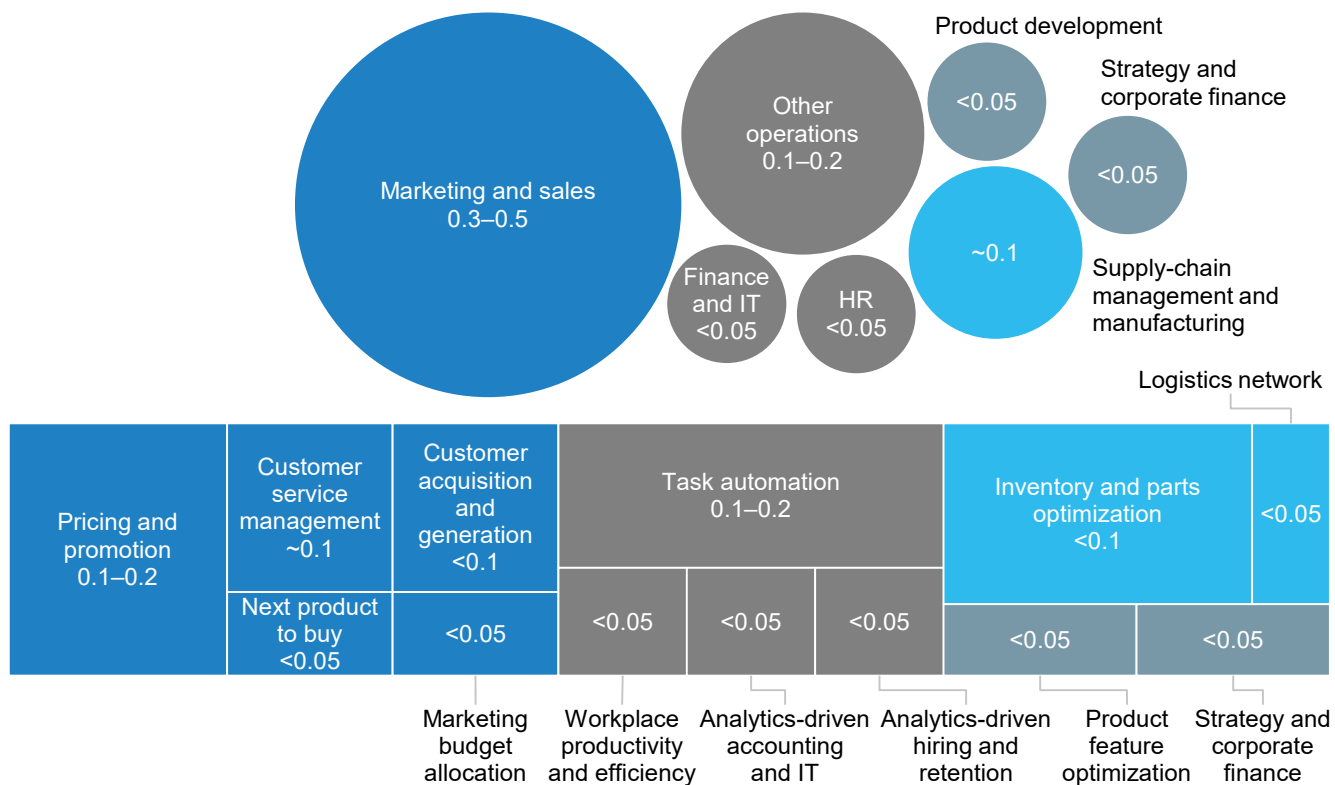


NOTE: Numbers may not sum due to rounding. Not to scale.

SOURCE: McKinsey Global Institute analysis

Exhibit 14

**In consumer packaged goods, AI's greatest potential is in supply-chain management and manufacturing, including predictive maintenance**

Packaged goods examples
$ trillion
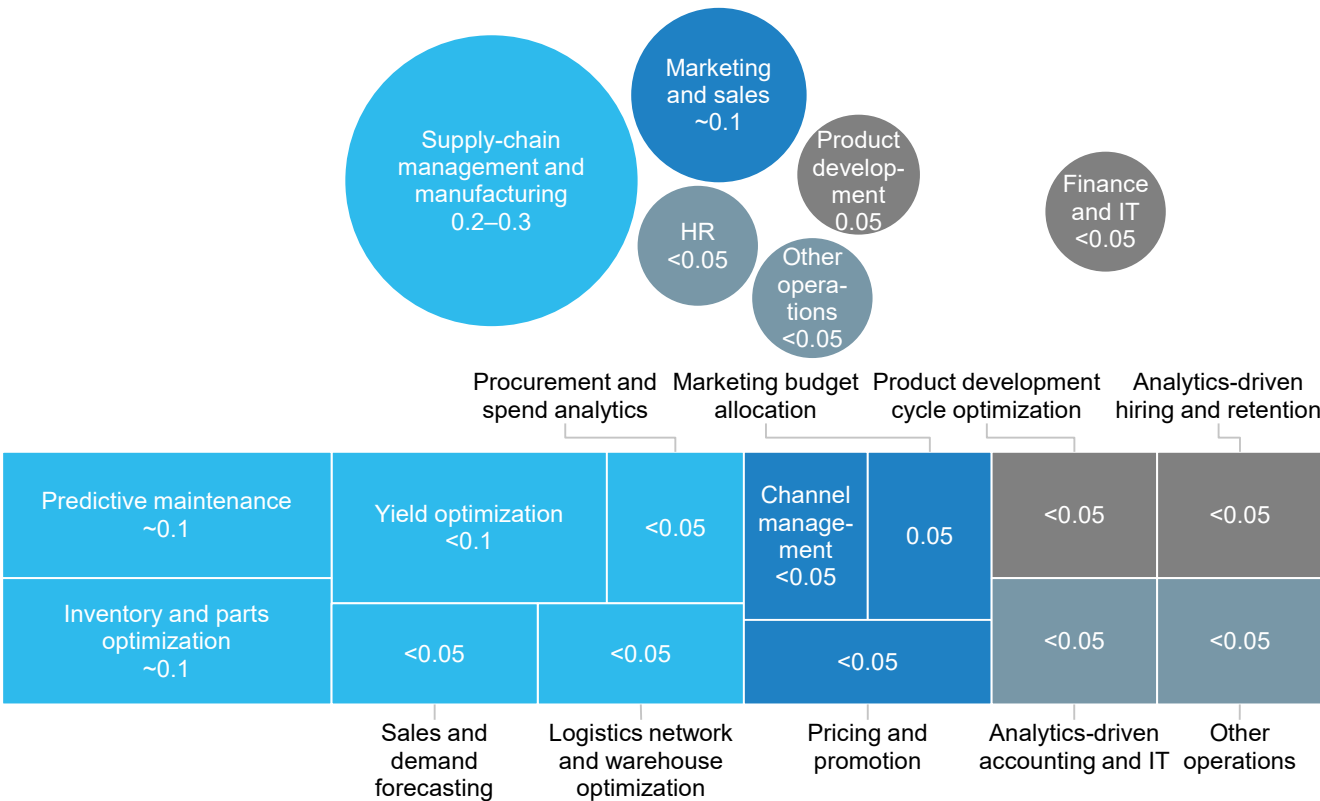


NOTE: Numbers may not sum due to rounding. Not to scale.

SOURCE: McKinsey Global Institute analysis

As our library of use cases evolves, we expect to see significant value in some areas that is not fully captured in these figures. One example is risk in banking. We believe that this could be one of the biggest areas of impact for deep learning using neural networks. For example, incorporating AI into underwriting models could allow banks to underwrite entirely new types of customers, such as the unbanked or semi-banked, and capture significant value through improved fraud detection. The application of neural network techniques to these new use cases in risk will need to be further tested before we can size the impact.

Exhibit 15

**In banking, marketing and sales and risk are among the areas with the most potential value from AI**

Banking example
$ trillion
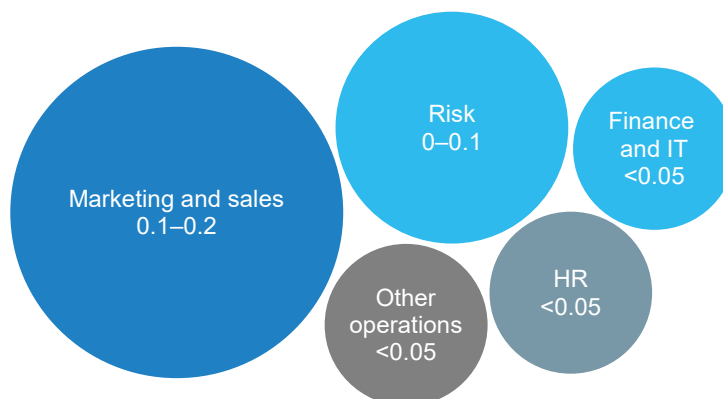


| Channel management 0.1 | Customer service management 0–0.1 | Customer acquisition/ lead generation <0.05 | Fraud and debt analytics 0–0.1 | | |
|---|---|---|---|---|---|
| | | Next product to buy (NPTB) <0.05 | Analytics-driven finance and IT <0.05 | Analytics-driven hiring and retention <0.05 | Workforce productivity and efficiency <0.05 |
| | Churn reduction 0.05 | Pricing and promotion <0.05 | | | |

NOTE: Numbers may not sum due to rounding. Not to scale.

SOURCE: McKinsey Global Institute analysis

# 4. THE ROAD TO IMPACT AND VALUE

Artificial intelligence is attracting growing amounts of corporate investment, and as the technologies develop, the potential value that can be unlocked is likely to grow. So far, however, only a few pioneering firms have adopted AI at scale. Prior research suggests that even among AI-aware firms, only about 20 percent are using one or more of the technologies in a core business process or at scale.[12] For all their promise, AI technologies have plenty of limitations that will need to be overcome, including not just data-related issues but also regulatory obstacles, and social and user acceptance. Yet the potential value to be harnessed provides a clear incentive for technology developers, companies, policy-makers, and users to try to tackle these issues.

## WHILE AI IS PROMISING, ITS USE STILL FACES LIMITATIONS AND CHALLENGES

As discussed, factors that could limit AI use include the requirements around the volume, type, and labeling of data. Other limitations are also significant, for now, although the evolving technologies themselves are starting to provide some solutions.

### Limitations include the need for massive data sets, difficulties in explaining results, generalizing learning, and potential bias in data and algorithms

Among the limitations we have identified, five stand out.[13] First is the challenge of labeling training data, which often must be done manually and is necessary for supervised learning. Ironically, machine learning often requires large amounts of manual effort; in supervised learning, the set of machine learning techniques that is most often used, machines are taught; they don't learn "by themselves." Promising new techniques are emerging to address this challenge, such as reinforcement learning (discussed earlier) and in-stream supervision, in which data can be labeled in the course of natural usage.[14] Second is the difficulty of obtaining data sets that are sufficiently large and comprehensive to be used for training; for many business use cases, creating or obtaining such massive data sets can be difficult—for example, limited clinical-trial data to predict health-care treatment outcomes more accurately. Third is the difficulty of explaining in human terms results from large and complex models: why was a certain decision reached? Product certifications in health care, as well as in the automotive, chemicals, and aerospace industries, for example, can be an obstacle; among other constraints, regulators often want rules and choice criteria to be clearly explainable. Some nascent approaches to increasing model transparency, including local-interpretable-model-agnostic explanations (LIME), may help resolve this explanation challenge in many cases.[15] Fourth is the generalizability of learning: AI models continue to have difficulties in carrying their experiences from one set of circumstances to another. That means companies must commit resources to train new models even for use cases that are similar to previous ones. Transfer learning—in which an AI model is trained to accomplish a certain task and then quickly applies that learning to a similar but distinct activity, is one promising response to this challenge.[16]

The fifth limitation concerns the risk of bias in data and algorithms. Unlike the other limitations listed, which may eventually be resolved by technical advances, this issue of bias touches on concerns that are more social in nature and which could require broader steps to resolve, such as understanding how the processes used to collect training data can

---

[12] Ibid. McKinsey Global institute, *Artificial intelligence*, June 2017.

[13] Ibid. Michael Chui, James Manyika, and Mehdi Miremadi, "What AI can and can't do (yet) for your business," *McKinsey Quarterly*, January 2018.

[14] Eric Horvitz, "Machine learning, reasoning, and intelligence in daily life: Directions and challenges," *Proceedings of Artificial Intelligence Techniques for Ambient Intelligence*, Hyderabad, India, January 2007.

[15] LIME attempts to identify which parts of input data a trained model relies on most to make predictions.

[16] For an early example application, see John Guttag, Eric Horvitz, and Jenna Wiens, "A study in transfer learning: Leveraging data from multiple hospitals to enhance hospital-specific predictions," *Journal of the American Medical Informatics Association*, volume 21, number 4, 2014.

influence the behavior of models they are used to train. In certain instances, when applied incorrectly, AI models risk perpetuating existing social and cultural biases. For example, unintended biases can be introduced when training data are not representative of the larger population to which an AI model is applied. Thus, data collected from a city-wide video camera network will likely show more instances of all kinds of signals, from crime to outdoor advertising engagements, in areas where the density of deployed cameras is higher, which might not be representative of the number of instances of these occurrences in a city as a whole. Similarly, facial recognition models trained on a population of faces corresponding to the demographics of artificial intelligence developers could struggle when applied to populations with more diverse characteristics.[17]

Moreover, a recent report on the malicious use of AI highlights a range of security threats, from sophisticated automation of hacking to hyper-personalized political disinformation campaigns.[18] Among the risks it lists are threats associated with privacy invasion and social manipulation from the use of AI to automate tasks involved in surveillance, such as analyzing mass-collected data, persuasion, including creating targeted propaganda, and deception, for example with manipulated videos. Multiple research efforts are under way to identify best practices and address such issues in academic, non-profit, and private-sector research.

## Organizational challenges around technology, processes, and people can slow or impede AI adoption

Organizations planning to adopt significant deep learning efforts will need to consider a spectrum of options about how to do so. The range of options includes building a complete in-house AI capability either gradually in an organic way or more rapidly through acquisitions, outsourcing these capabilities, or leveraging AI-as-a-service offerings.

Given the importance of data, it is vital for organizations to develop strategies for the creation and/or acquisition of training data. But the effective application of AI also requires organizations to address other key data challenges, including establishing effective data governance, defining ontologies, data engineering around the "pipes" from data sources, managing models over time, building the data pipes from AI insights to either human or machine actions, and managing regulatory constraints.

Given the significant computational requirements of deep learning, some organizations will maintain their own data centers, because of regulations or security concerns, but the capital expenditures could be considerable, particularly when using specialized hardware. Cloud vendors offer another option.

Process can also become an impediment to successful adoption unless organizations are digitally mature. On the technical side, organizations will have to develop robust data maintenance and governance processes, and implement modern software disciplines such as Agile and DevOps. Even more challenging, in terms of scale, is overcoming the "last mile" problem of making sure the superior insights provided by AI are instantiated in the behavior of the people and processes of an enterprise.

On the people front, much of the construction and optimization of deep neural networks remains something of an art requiring real experts to deliver step-change performance increases. Demand for these skills far outstrips supply at present; according to some estimates fewer than 10,000 people have the skills necessary to tackle serious AI problems

---

[17]  Joy Buolamwini and Timnit Gebru. "Gender shades: Intersectional accuracy disparities in commercial gender classification," *Proceedings of Machine Learning Research*, 2018. http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

[18]  Miles Brundage et al., *The malicious use of artificial intelligence: Forecasting, prevention, and mitigation*, Future of Humanity Institute, February 2018.

and competition for them is fierce amongst the tech giants.[19] Companies wanting to build their own AI solutions will need to consider whether they have the capacity to attract and retain these specialized skills.

### AI can seem an elusive business case for now

Where AI techniques and data are available and the value is clearly proven, organizations can already pursue the opportunity. But in some areas, the techniques today may be mature and the data available, but the cost and complexity of deploying AI may simply not be worthwhile, given the value that could be generated. For example, an airline could use facial recognition and other biometric scanning technology to streamline aircraft boarding, but the value of doing so may not justify the cost and issues around privacy and personal identification. A recent Stanford University study found that deep neural networks can make highly accurate bond price predictions, but took hours to come up with the answer, whereas other "simpler" techniques produced an answer that was only slightly less accurate but very rapid—just four seconds.[20] For a bond trader that timing difference is critical.

Similarly, we can see potential cases in which the data and the techniques are maturing, but the value is not yet clear. For example, in mining, AI could potentially play a significant role in providing ore body insights, thereby allowing for more efficient exploration, drilling, and mine planning. Given the high capital expenditure costs usually involved in this sector, the benefits could be very significant, but for now are unmeasurable. In other sectors, including banking, opportunities that may exist are sometimes in areas closely guarded as competitive secrets.

In the most unpredictable scenario of all, either the data (both the types and volume) or the techniques are simply too new and untested to know how much value they could unlock. For example, in health care, if AI could build on the superhuman precision we are already starting to see with X-ray analysis and broaden that to more accurate diagnoses and even automated medical procedures, the economic value could be very significant. At the same time, the complexities and costs of arriving at this frontier are also daunting. Among other issues, it would require flawless technical execution and resolving issues of malpractice insurance and other legal concerns.

### Societal concerns and regulations can constrain AI use

Societal concerns and regulations can affect potential value capture from AI, including in use cases related to personally identifiable information. This is particularly relevant at a time of growing public debate about the use and commercialization of individual data on some online platforms.[21] Use and storage of personal information is especially sensitive in sectors such as banking, health care, and pharmaceutical and medical products, as well as in the public and social sector. Alongside these questions about privacy, issues of fairness and equity may arise from bias in data, as well as concerns about transparency and accountability in the use of massively complex algorithms. In addition to addressing these issues, businesses and other users of data for AI will need to continue to evolve business models related to data use in order to address societies' concerns. Furthermore, regulatory requirements and restrictions can differ from country to country, as well from sector to sector. In the European Union, for example, automated individual decision making—that is, algorithms that make decisions based on user-level predictors—will be shaped by the EU-wide general data protection regulation taking effect in 2018, which provides for a right to an explanation for some decisions made by machines. This could affect the insurance industry,

---

[19] Cade Metz, "Tech giants are paying huge salaries for scarce AI talent," *The New York Times*, October 22, 2017.

[20] "Neural networks face unexpected problems in analyzing financial data," *MIT Technology Review*, May 10, 2017.

[21] Julia Angwin, *Dragnet Nation: A quest for privacy, security, and freedom in a world of relentless surveillance*, Times Books, 2014; see also Anna Bernasek and D. T. Mongan, *All you can pay: How companies use our data to empty our wallets*, Nation Books, 2015.

for example, in which underwriters could be required to explain how machines derived their answers about whether to underwrite the risk or explain why the premium was set at a certain level.

## IMPLICATIONS FOR STAKEHOLDERS

As we have seen, it is a company's ability to execute against AI models, rather than the models themselves, that creates value. In prior publications, we have detailed some of the general implications for companies as they adopt AI and analytics.[22] In this final section, we sketch out some of the high-level implications of our study of AI use cases for providers of AI technology, appliers of AI technology, and policy makers who set the context for both.

### For AI technology provider companies

Many companies that develop or provide AI to others have considerable strength in the technology itself and the data scientists needed to make it work, but they can lack a deep understanding of end markets. This is a challenge, since our research has shown that most of the potential impact of AI comes from improving the performance in existing use cases—in other words, the fundamental drivers of the businesses of their potential customers. Furthermore, many of these companies are asking how they should prioritize their resources, whether it is R&D, marketing and sales, or other functions to take advantage of AI opportunities.

Understanding the value potential of AI across sectors and functions can help shape the portfolios of these AI technology companies. That said, they shouldn't necessarily only prioritize the areas of highest potential value. Instead, they can combine that data with complementary analyses of the competitor landscape, of their own existing strengths including technology and data strengths, their sector or function knowledge, and their customer relationships, to shape their investment portfolios.

On the technical side, the mapping of problem types and techniques to sectors and functions of potential value can guide a company with specific areas of expertise as to where to focus. Similarly, for technology companies that have access to certain types of data, this mapping can help guide them to where their access to data can provide them with the most leverage, and toward data gaps needing to be filled. Finally, more innovation is needed to continue to advance the frontier of AI and address some ongoing technical limitations.

### For companies adopting AI to transform and power their own businesses

Many companies seeking to adopt AI in their operations have started machine learning and AI experiments across their business—and are likely to be bombarded by technology companies trying to sell them "AI solutions." Before launching more pilots or testing solutions, it is useful to step back and take a holistic approach to the issue, moving to create a prioritized portfolio of initiatives across the enterprise, including AI and the wider analytic and digital techniques available. For a business leader to create an appropriate portfolio, it is important to develop an understanding about which use cases and domains have the potential to drive the most value for a company, as well as which AI and other analytical techniques will need to be deployed to capture that value. This portfolio ought to be informed not only by where the theoretical value can be captured, but by the question of how the techniques can be deployed at scale across the enterprise. The question of how analytical techniques are scaling is driven less by the techniques themselves and more by a company's skills, capabilities, and data. Companies will need to consider efforts on the "first mile," that is, how to acquire and organize data and efforts, as well as on the "last mile", or how to integrate the output of AI models into work flows, ranging from clinical trial managers and sales force managers to procurement officers. Previous MGI research suggests that

---

[22] See for example, *Artificial intelligence: The next digital frontier?* McKinsey Global Institute, June 2017 and *The age of analytics: Competing in a data-driven world*, McKinsey Global Institute, December 2016.

AI leaders invest heavily in these first- and last-mile efforts.[23] In addition, given growing user and societal concerns, companies deploying AI will need to think through their safe and responsible data use and the business models they employ that make use of the user or customer data.

### For policy makers

Policy makers will need to strike a balance between supporting the development of AI technologies and managing any risks from bad actors, as well as irresponsible use of AI techniques and the data they employ. They have an interest in supporting broad adoption, since AI can lead to higher labor productivity, economic growth, and societal prosperity. Tools to help them include public investments in research and development as well as support for a variety of training programs, which can help nurture AI talent. On the issue of data, governments can spur the development of training data directly through open data initiatives. Opening up public-sector data can spur private-sector innovation. Setting common data standards can also help.

AI is also raising new questions for policy makers to grapple with for which historical tools and frameworks may not be adequate.[24] Therefore, some policy innovations will likely be needed to cope with these rapidly evolving technologies. But given the scale of the beneficial impact on business, the economy, and society, the goal should not be to constrain the adoption and application of AI, but rather to encourage its beneficial and safe use.

• • •

The several hundred use cases we examined underscore the value and performance enhancement that adoption of AI technologies can bring, and provide some indication of where AI can most usefully be deployed. To capture that value, CEOs and other C-suite executives will need to ramp up and staff their analytics capabilities; as our use cases show, AI can most often be adopted and create value where other analytics methods and techniques are also creating value. For companies seeking to deploy AI, that means the same basics will need to be in place, especially moving forward on digitizing their enterprises. As we have seen, abundant volumes of rich data from images, audio, and video, and large-scale text are the essential starting point and lifeblood of creating value with AI. Above all, our analysis of use cases suggests that successful AI adoption will require focus and setting of priorities. Its value is tremendous—and looks set to become even more so as the technologies themselves advance. Identifying where and how that value can be captured looks likely to become one of the key business challenges of our era.

---

[23]  *The age of analytics: Competing in a data-driven world*, McKinsey Global Institute, December 2016.

[24]  Bank risk management is one example. See Ignacio Crespo, Pankaj Kumar, Peter Noteboom, and Marc Taymans, *The evolution of model risk management*, McKinsey & Company, February 2017.

# ACKNOWLEDGMENTS