Tallinn
Digital
Summit

digitalsummit.ee

Briefing note: The Promise and Challenge of the Age of Artificial Intelligence

By The McKinsey Global Institute

McKinsey&Company

MCKINSEY GLOBAL INSTITUTE

This briefing note has been prepared by the McKinsey Global Institute for Tallinn Digital Summit, October 2018.

Summary:

The time may have finally come for artificial intelligence (AI) — this after periods of hype followed by several "AI winters" over the past 60 years. Today, announcements of technical breakthroughs of superhuman performance by machines follow one another in rapid succession. AI now powers so many real-world applications ranging from facial recognition, language translators, and assistants like Siri and Alexa that we barely notice it. Along with these consumer applications, companies across sectors are increasingly harnessing AI's power in their operations. Embracing AI promises considerable benefits for businesses and economies through its contributions to productivity growth and innovation. Already, early adopters are starting to reap benefits, but adoption is still at a relatively early stage and uneven, and AI readiness varies widely among companies, sectors, and countries.

At the same time, AI's impact on work is likely to be profound. Some occupations as well as demand for some skills will decline, while others grow, and many change as people work alongside ever evolving and increasingly capable machines. While most scenarios we have developed suggest that more jobs will be created than lost to automation, the transition is likely to be disruptive, and occupations and skill requirements will shift significantly. AI also has the potential to contribute to tackling pressing societal challenges from healthcare to climate change and humanitarian crises, but it could also introduce or exacerbate social challenges, through its use and misuse, and challenges related to bias, privacy, and cybersecurity. Given its economic and other benefits, policy makers and business leaders should embrace AI, but put in place measures to facilitate adoption, soften the disruptive transitions that will accompany it, and place a new focus on responsible use.

This briefing note draws on ongoing research by the McKinsey Global Institute into AI technologies, their uses, limitations, and impact. It concludes with a set of issues that policy makers and business leaders will need to address. A list of the main sources used in the note and further reading is included at the end.

1. AI's time may have finally come, but more progress is needed

The term "artificial intelligence" was popularized at a conference at Dartmouth College in the United States in 1956, which brought together researchers on a broad range of topics, from language simulation to learning machines. Despite periods of significant scientific advances in the six decades since, AI has often failed to live up to the hype that surrounded it. Decades were spent trying to describe human intelligence precisely, and the progress made did not deliver on the earlier excitement.

Since the late 1990s, however, technological progress has gathered pace, especially in the past decade. Machine-learning algorithms have progressed, especially through the development of deep learning and reinforcement-learning techniques based on neural networks. Several other factors have contributed to the recent progress: exponentially more computing capacity has become available to train larger and more complex models; this has come through silicon-level innovation including the use of graphics processing units (GPUs), tensor processing units (TPUs) and more on the way. This compute capacity is being aggregated in hyper-scale clusters, increasingly being made accessible to users through the cloud.

Another key factor is massive amount of data being generated and now available to train AI algorithms. Some of the progress in AI has been the result of system-level innovations. Autonomous vehicles are a good illustration of this: they take advantage of innovations in sensors, LIDAR, machine vision, mapping and satellite technology, navigation algorithms, and robotics all brought together in integrated systems.

Despite the progress, many hard problems remain that will require more scientific breakthroughs. So far, most of the progress has been in what is often referred to as "narrow AI"— where machine-learning techniques are being developed to solve specific problems, for example, in natural language processing. The harder issues are in what is generally referred to as "Artificial General Intelligence", where the challenge is to develop AI that can tackle general problems, much like humans can. Many researchers consider this to be decades away from becoming reality.

Deep learning and machine-learning techniques are driving AI

Much of the recent excitement in AI has been the result of advances in the field known as deep learning, a set of techniques to implement machine learning that is based on artificial neural networks. These AI systems loosely model the way that neurons interact in the brain. Neural networks have many ("deep") layers of simulated interconnected neurons, hence the term "deep learning." Whereas earlier neural networks only had 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.

There are several types of machine learning: supervised learning, unsupervised learning, and reinforcement learning, with each best suited to certain use cases. Most current practical examples of AI are applications of supervised learning. In supervised learning, often used when labeled data are available and the preferred output variables are known, training data are used to help a system learn the relationship of given inputs to a given output, for example, to recognize objects in an image, or transcribe human speech. Unsupervised learning is a set of techniques used without labeled training data, for example, to detect clusters or patterns in a set of existing data, such as images of buildings that have similar architectural styles. In reinforcement learning, systems are trained by receiving virtual "rewards" or "punishments," often through a scoring system, essentially learning by trial and error. Needless to say, as a result of ongoing work, these techniques are evolving.

Limitations remain, although new techniques show promise

AI still faces many practical challenges, though new techniques are emerging to address them. Machine learning can require large amounts of human effort to label the training data necessary for supervised learning. In-stream supervision, in which data can be labeled in the course of natural usage, and other techniques, could help alleviate this issue. Obtaining data sets that are sufficiently large and comprehensive to be used for training is also often challenging, for example, creating or obtaining sufficient clinical-trial data to predict health-care treatment outcomes more accurately.

The "black box" complexity of deep-learning techniques creates the challenge of "explainability"—that is, showing which factors led to a decision or prediction, and how. This is particularly important in applications where trust matters and there are societal implications of such predictions, as in criminal justice applications or financial lending. Some nascent approaches, including local-interpretable-model-agnostic explanations (LIME), aim to increase model transparency. Another challenge is that of building generalized learning techniques, since AI techniques continue to have difficulties in carrying their experiences from one set of circumstances to another. 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.

2. Businesses stand to benefit from AI

While AI is increasingly pervasive in consumer applications, businesses are beginning to adopt it across their operations, at times with striking results.

AI's potential cuts across industries and functions

AI can be used to improve business performance in areas including predictive maintenance, where deep learning's ability to analyze large amounts of high dimensional data from audio and images can effectively detect anomalies in factory assembly lines or aircraft engines. In logistics, AI can optimize routing of delivery traffic, improving fuel efficiency and reducing delivery times. In customer service management, AI has become a valuable tool in call centers, thanks to improved speech recognition. In sales, combining customer demographic and past transaction data with social media monitoring can help generate individualized "next product to buy" recommendations, which many retailers now use routinely.

Such practical AI use-cases and applications can be found across all sectors of the economy and multiple business functions from marketing to supply chain operations. In many of these use-cases, deep learning techniques primarily add value by improving on traditional analytics techniques. Our analysis of more than 400 use cases across 19 industries and nine business functions found that AI improved on traditional analytics techniques in 69 percent of potential use cases. In only 16 percent of AI use cases did we find a "greenfield" AI solution that was applicable where other analytics methods would not be effective. Our research estimated that deep learning techniques based on artificial neural networks could generate as much as 40 percent of the total potential value that all analytics techniques could provide by 2030. Further, we estimate that several of the deep-learning techniques could enable up to $6 trillion in value annually.

So far, adoption is uneven across companies and sectors

Although many organizations have begun to adopt AI, the pace and extent of adoption has been uneven. Nearly half of respondents in a 2018 McKinsey survey on AI adoption say their companies have embedded at least one AI capability in their business processes, and another 30 percent are piloting AI. Still, only 21 percent say their organizations have embedded AI in several parts of the business and barely 3 percent of large firms have integrated AI across their full enterprise workflows.

Other surveys show that early AI adopters tend to think about these technologies more expansively, to grow their markets or increase market share, while companies with less experience focus more narrowly on reducing costs. Highly digitized companies tend to invest more in AI and derive greater value from its use. At the sector level, too the gap between digitized early adopters and others is widening. Sectors highly ranked in MGI's Industry Digitization Index, such as high tech and telecoms or financial services, are leading AI adopters and have the most ambitious AI investment plans. As these firms expand AI adoption and acquire more data and AI capabilities, laggards may find it harder to catch up.

Several challenges to adoption

Many companies and sectors lag in AI adoption. Developing an AI strategy with clearly defined benefits, finding talent with the appropriate skill set, overcoming functional silos that constrain end-to-end deployment, and lacking ownership and commitment to AI on the part of leaders, are among the barriers to adoption most often cited by executives. On the strategy side, companies will need to develop an enterprise-wide view of compelling AI opportunities, potentially transforming parts of their current business processes. For data, organizations will need to have robust data capture and governance processes, modern digital capabilities, and be able to build or access the requisite infrastructure.

Even more challenging will be overcoming the "last mile" problem of making sure that the superior insights provided by AI are inculcated in the behavior of the people and processes of an enterprise. On the talent front, much of the construction and optimization of deep neural networks remains an art requiring real expertise. Demand for these skills far outstrips supply; according to some estimates fewer than 10,000 people have the skills necessary to tackle serious AI problems, and competition for them is fierce. Companies considering the option of building their own AI solutions will need to consider whether they have the capacity to attract and retain these specialized skills.

3. Economies also stand to benefit from AI primarily through increased productivity and innovation

Deployment of AI and automation technologies can do much to lift the global economy and increase global prosperity. At a time of aging and falling birth rates, productivity growth becomes critical for long term economic growth. Even in the near term, productivity growth has been sluggish in developed economies, dropping to an average of 0.5 percent in 2010—14 from 2.4 percent a decade earlier in the United States and major European economies. AI has the potential to contribute to productivity growth much like previous general-purpose technologies.

AI could contribute to economic impact through a variety of channels

The largest economic impacts of AI will likely be on productivity growth through labor market effects including substitution, augmentation, and contributions to labor productivity. Our research suggests that labor substitution could account for less than half of the total benefit. AI will augment human capabilities, freeing up workers to engage in more productive and higher-value tasks, as well as increase demand for jobs associated with AI technologies. AI can also boost innovation, enabling companies to improve their top line by reaching underserved markets more effectively with existing products, and over the longer term, creating entirely new products and services.

AI will also create positive externalities, facilitating more efficient cross-border commerce and enabling expanded use of valuable cross-border data flows. Such increases in economic activity and incomes can be reinvested into the economy, contributing to further growth. The deployment of AI will also bring some negative externalities that could lower, although not eliminate, the positive economic impacts. On the economic front, these include increased competition that shifts market share from nonadopters to front-runners, the costs associated with managing labor-market transitions and potential loss of consumption for citizens during periods of unemployment, as well the transition and implementation costs for deploying AI systems.

All in all, these various channels net out to significant positive economic growth, assuming businesses and governments proactively manage the transition. One simulation we conducted using McKinsey survey data suggests that AI adoption could raise global GDP by as much as $13 trillion by 2030, about 1.2 percent additional GDP growth per year. This effect will only build up through time, however, given that most of the implementation costs of AI may be ahead of the revenue potential.

The AI-readiness of countries varies considerably

Countries vary on leading enablers of their potential AI-driven economic growth, such as investment and research activity, digital absorption, connectedness, and labor-market structure and flexibility. Our research suggests that the ability to innovate and acquire the necessary human capital skills will be among the most important enablers—and that AI competitiveness will likely be an important factor influencing future GDP growth.

Countries leading the race to supply AI have unique strengths that set them apart. Scale effects enable more significant investment, and network effects enable these economies to attract the talent needed to make the most of

AI. For now, China and the United States are responsible for most AI-related research activities and investment. A second group of countries that includes Germany, Japan, Canada, and the United Kingdom have a history of driving innovation on a major scale and may accelerate the commercialization of AI solutions. Smaller, globally connected economies such as Belgium, Singapore, South Korea, and Sweden also score highly on their ability to foster productive environments where novel business models thrive. A third group of countries, that includes but is not limited to Brazil, India, Italy, and Malaysia, are in a relatively weaker starting position, but they exhibit comparative strengths in specific areas on which they may be able to build. India, for instance, produces around 1.7 million graduates a year with STEM degrees—more than the total of STEM graduates produced by all G7 countries. Other countries with relatively underdeveloped digital infrastructure, innovation and investment capacity, and digital skills, risk falling behind their peers.

4. AI and automation will have a profound impact on work

Even as AI and automation bring benefits to business and the economy, major disruptions to work can be expected.

About half of the current work activities (not jobs) are technically automatable

Our analysis of the impact of automation and AI and work shows that certain categories of activities are technically more easily automatable than others. They include physical activities in highly predictable and structured environments, as well as data collection and data processing, which together account for roughly half of the activities that people do across all sectors in most economies. The least susceptible categories include managing others, providing expertise, and interfacing with stakeholders. The density of these highly automatable activities varies across occupations, sectors and to a lesser extent across countries. Our research finds that about 30 percent of the activities in 60 percent of all occupations could be automated—but that in only about 5 percent of occupations are nearly all activities automatable. In other words, more occupations will be partially automated than wholly automated.

Three simultaneous jobs effects: Jobs lost, jobs gained, jobs changed

The pace and extent to which automation will be adopted and impact actual jobs will depend on several factors besides technical feasibility of automation. These include the cost of deployment and adoption, and the labor-market dynamics, including labor supply quantity, quality, and the associated wages—this factor leads to wide differences across developed and developing economies. The business benefits beyond labor substitution—often involving use of AI for beyond human capabilities—which contribute to business cases for adoption are another factor. Social norms and acceptance and various regulatory factors will also determine the timing. How all these factors play out across sectors and countries will vary, and for countries will largely be driven by the labor-market dynamics. For example, in advanced economies with relatively high wage levels, such as France, Japan, and the United States, jobs affected by automation could be more than double the rate in India.

Given the interplay of all these factors, it is difficult to make predictions, but possible to develop various scenarios. First on jobs lost: Our midpoint adoption scenario of the period 2016 to 2030 suggests that about 15 percent of the global workforce (400 million workers worldwide), could be displaced by automation. Second, jobs gained: We developed scenarios for labor demand to 2030 based on anticipated economic growth through productivity and by considering several drivers of demand for work. These included rising incomes, especially in emerging economies, increased spending on healthcare for aging populations, investment in infrastructure and buildings, energy transition spending, and spending on technology development and deployment. The number of jobs gained through these and other catalysts could range from 555 million to 890 million, or 21 to 33 percent of the global workforce. This suggests that the growth in demand for work, barring extreme scenarios, would thus more than offset the number of jobs lost to automation. However, it is important to note that in many emerging economies with young populations, there will already be a challenging need to provide jobs to workers entering the workforce and that, in developed economies, the approximate balance between jobs lost and those created in our scenarios is also a consequence of aging, and thus fewer people entering the workforce

No less significant are the jobs that will change as machines increasingly complement human labor in the workplace. The notion of jobs changed will come about as a result of the partial automation observation made above, and will affect many more occupations than jobs lost. Skills for workers complemented by machines, as well as work design, will need to adapt to keep up with rapidly evolving and increasingly capable machines.

Four major workforce transitions

Even if there will be enough work for people in 2030, as most of our scenarios suggest, the transitions that will accompany automation and AI adoption will be significant.

First, millions of workers will likely need to change occupations. Some of these shifts will happen within companies and sectors, but many will occur across sectors and even geographies. While occupations requiring physical activities in highly structured environments or in data processing will decline, others that are difficult to automate will grow. These could include managers, teachers, nursing aides, and tech and other professionals, but also gardeners and plumbers, who work in unpredictable physical environments. These changes may not be smooth and could lead to temporary spikes in unemployment. Second, workers will need different skills to thrive in the workplace of the future. Demand for social and emotional skills such as communication and empathy will grow almost as fast as demand for many advanced technological skills. Basic digital skills have been increasing in all jobs. Automation will also spur growth in the need for higher cognitive skills, particularly critical thinking, creativity, and complex information processing. Demand for physical and manual skills will decline, but these will remain the single largest category of workforce skills in 2030 in many countries. The pace of skill shifts has been accelerating, and may lead to excess demand for some skills and excess supply for others.

Third, workplaces and workflows will change as more people work alongside machines. As self-checkout machines are introduced in stores, for example, cashiers will shift from scanning merchandise themselves to help-ing answer questions or troubleshoot the machines. Finally, automation will likely put pressure on average wages in advanced economies. Many of the current middle-wage jobs in advanced economies are dominated by highly automatable activities, such as in manufacturing or in accounting, which are likely to decline. High-wage jobs will grow significantly, especially for high-skill medical and tech or other professionals. However, a large portion of jobs expected to be created typically have lower wage structures, such as teachers and nursing aides.

In tackling these transitions, many economies, especially in the OECD, start in a hole, given the already existing skill shortages and challenged educational systems, as well as the trends towards declining expenditures on on-the-job training and worker transition supports. Many economies are already experiencing income inequality and wage polarization.

5. AI will also bring both societal benefits and challenges

Alongside the economic benefits and challenges, AI will impact society in a positive way, as it helps tackle societal challenges ranging from health and nutrition to equality and inclusion. However, it is also creating pitfalls that will need to be addressed, including unintended consequences and misuse.

AI could help tackle some of society's most pressing challenges

By automating routine or unsafe activities or those prone to human error, AI could allow humans to be more productive and work and live more safely. One study looking at the United States estimates that replacing human drivers with more accurate autonomous vehicles could save thousands of lives per year by reducing accidents. AI can also reduce the need for humans to work in unsafe environments such as offshore oil rigs, or coal mines. DARPA, for example, is testing small robots that could be deployed in disaster areas to reduce the need for humans to be put in harm's way. Several AI capabilities are especially relevant. Image classification on photos of skin taken via a mobile phone app could evaluate whether moles are cancerous or not, facilitating early-stage diagnosis for individuals with limited access to dermatologists. Object detection can help visually impaired people navigate and interact with their environment by identifying obstacles such as cars and lamp post. Natural language processing could be used to monitor disease outbreaks by monitoring and analyzing text messages in local languages.

Our work and that of others has highlighted numerous use cases across many domains where AI could be applied for social good. For these AI-enabled interventions to be effectively applied, several barriers must be overcome. These include the usual challenges of data, computing, and talent availability faced by any organization trying to apply AI, as well as more basic challenges of access, infrastructure, and financial resources that are particularly acute in remote or economically-challenged places and communities.

AI will need to address societal concerns including unintended consequences, misuse, algorithmic bias, and questions about data privacy

In economic terms, difficult questions will need to be addressed around the widening economic gaps across individuals, firms, sectors and even countries that might emerge as an unintended consequence of deployment. Other areas of concern include the use and misuse of AI. These range from use in surveillance and military applications, to use in social media and politics, or where the impact has social consequences such as in criminal justice systems. We must also consider the potential for users with malicious intent, including in areas of cybersecurity. Multiple research efforts are currently under way to identify best practices and address such issues in academic, non-profit, and private-sector research.

Some concerns are directly related to the way algorithms and the data used to train them may introduce new biases or perpetuate and institutionalize existing social and procedural biases. For example, facial recognition models trained on a population of faces corresponding to the demographics of artificial intelligence developers may not reflect the broader population. Data privacy and use of personal information are also critical issues to address if AI is to realize its potential. Europe has led the way here with the General Data Protection Regulation (GDPR), which introduced more stringent consent requirements for data collection, gives users the right to be forgotten and the right to object, and strengthens obligations on organizations that gather, control, and process data, with significant fines for failures to comply. Cybersecurity and "deep fakes" that could manipulate election results or perpetrate large-scale fraud are also a concern.

6. Some priorities for achieving good outcomes

The potential benefits of AI to business, the economy, and some of the societal challenges should encourage business and policy-makers to embrace and adopt AI. At the same time, the potential challenges to adoption including the workforce impacts, and other social concerns cannot be ignored. Key challenges to be addressed include:

The deployment challenge. We have an interest in embracing AI, given its likely contributions to business value, economic growth, and social good, at a time when many economies need to boost productivity. Businesses and countries have a strong incentive to keep up with the rate of progress in global leaders such as the United States and China. Increased and broad deployment will require accelerating the progress being made on the technical challenges, as well making sure that all potential users have access to AI and can benefit from it. Among measures that may be needed:

- Investing in and continuing to progress AI research and innovation in a manner that ensures benefits can be shared by all.
- Expanding available data sets, especially in areas where their use would drive wider benefits for the economy and society.
- Investing in AI-relevant human capital and infrastructure to broaden the talent base capable of creating and executing AI solutions to keep pace with global AI leaders.
- Encouraging further AI literacy among business leaders and policy makers to guide informed decision making.
- Supporting existing digitization efforts that form the foundation for eventual AI deployment for both organizations and countries.

The future of work challenge. A starting point for addressing the potential disruptive impacts of automation will be to ensure robust economic and productivity growth, which is a pre-requisite for job growth and increasing prosperity. Governments will also need to foster business dynamism, since entrepreneurship and more rapid new business formation will not only boost productivity, but also drive job creation. Addressing the skills, jobs, and wages issues will require more focused measures. These include:

- Evolving education systems and learning for a changed workplace by focusing on STEM skills as well as creativity, critical thinking, and life-long learning.
- Stepping up private and public-sector investments in human capital, perhaps aided by incentives and credits analogous to those available for R&D investments.
- Improving labor market dynamism by supporting better credentialing and matching, as well as enabling diverse forms of work, including the gig economy.
- Rethinking incomes by considering and experimenting with programs that would provide not only income for work, but also meaning and dignity.
- Rethinking transition support and safety nets for workers affected, by drawing on best practices from around the world and considering new approaches.

The responsible AI challenge. AI will not live up to its promise if the public loses confidence in it as a result of privacy violations, bias, or malicious use, or if much of the world comes to blame it for exacerbating inequality. Establishing confidence in its abilities to do good, at the same time as addressing misuses, will be critical. Approaches could include:

- Strengthening consumer, data, and privacy and security protections.
- Establishing a generally-shared framework and set of principles for the beneficial and safe use of AI.
- Best practice sharing and ongoing innovation to address issues such as safety, bias, and explainability.
- Striking the right balance between the business and national competitive race to lead in AI to ensure the benefits of AI are widely available and shared.

Authors

James Manyika is chairman and director of the McKinsey Global Institute (MGI) and a senior partner at McKinsey & Company based in San Francisco; Jacques Bughin is an MGI director based and McKinsey senior partner based in Brussels. Michael Chui, an MGI partner in San Francisco, Jake Silberg in the San Francisco Office, and Peter Gumbel, MGI's editorial director, contributed to this briefing note.

References

McKinsey Global Institute research reports are available on www.mckinsey.com/mgi. For this briefing note, we have drawn on the following reports:

A labor market that works: Connecting talent with opportunity in the digital age, June 2015.

Digital America: A tale of the haves and the have-mores, December 2015.

Digital Europe: Pushing the frontier, capturing the benefits, June 2016.

A future that works: Automation, employment, and productivity, January 2017.

Jobs lost, jobs gained: Workforce transitions in a time of automation, January 2018.

Skill shift: Automation and the future of the workforce, May 2018.

Artificial intelligence: The next digital frontier?, June 2017.

Notes from the AI frontier: Insights from hundreds of use cases, April 2018.

Notes from the AI frontier: Modeling the impact of AI on the world economy, September 2018.

Further reading

Daron Acemoglu and Pascual Restrepo, *Artificial intelligence, automation and work*, NBER, working paper 24196, January 2018.

Ajay Agrawal, John McHale, and Alex Oettl, *Finding needles in haystacks: Artificial intelligence and recombinant growth*, NBER working paper 24541, April 2018

Erik Brynjolfsson, Daniel Rock, and Chad Syverson, *Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics*, NBER working paper 24001, December 2017.

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

Carlos Guestrin, Marco Tulio Ribeiro, and Sameer Singh, "Introduction to local interpretable model-agnostic explanations (LIME)," August 12, 2016, O'Reilly, oreilly.com.

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, January 2014.

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.

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.

Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep learning," *Nature*, volume 521, May 2015.

Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, third edition, Upper Saddle River, NJ: Prentice Hall Press, 2009.