Table of Contents

1. The rise of Artificial Intelligence

A. Defining Artificial Intelligence

The definition of "Artificial Intelligence" is not easy and remains contested,[1] especially given science's inability to nail a definition of "intelligence" accepted by all. Definitions abound and generally overlap by pointing to 'agents' (programs running on computer systems) able to learn, adapt and deploy themselves successfully in dynamic and uncertain environments. Intelligence in that sense intersects with autonomy and adaptability, through the ability to learn from a dynamic environment. The ambiguity which has and still surrounds the notion of "Artificial Intelligence" calls for an initial exercise in pedagogy over its definition, boundaries and dynamics (Part I), if we want then to analyze the significant opportunities it creates (Part II), and the associated socio-economic challenges it poses (Part III).

The intersection of Big Data, machine learning and cloud computing

To understand the current renaissance of what we frame as "Artificial Intelligence," which is as old as computer science, we need to turn to the convergence of three trends: i) Big Data, ii) machine learning and iii) cloud super-computing. In that sense, the rise of AI is really a manifestation of the digital revolution. One of its central laws, predicted in 1965 by *Intel* chip manufacturer co-founder Gordon Moore, tells us that computing power doubles every two years, on an average, at a constant cost.[2] This exponential growth has resulted from continued technoscientific prowess in miniaturization, bringing about the age of micro- and, now, nano-computing with increasing power; and along with it, the possibility of smart phones and the "Internet of Things."

Coupled with the development of Internet communication protocols and machine virtualization, the digital revolution then made possible the availability of highly and easily scalable supercomputing capabilities on the cloud. From that point, the exponentially growing flow of high resolution data[3] produced day after day by connected humans and machines could be processed by algorithms.

These contexts finally made possible the flourishing of an old branch of computer science, called machine learning,[4] where algorithms are capable of automatically sorting out complex patterns out of

---

[1] There is no standardized and globally accepted definition for what AI is. "*The choice of the very name "artificial intelligence" is a perfect example: if the mathematician John McCarthy used these words to propose the Dartmouth Summer Research Project – the workshop of summer 1956 that many consider as the kick-off of the discipline – it was as much to set it apart from related research, such as automata theory and cybernetics, as to give it a proper definition […].There are actually many definitions for artificial intelligence. A first great group of definitions could be called "*essentialist*", aiming at defining the end-goal a system has to show to enter the category […].Besides this – and often complementarily – are the definitions one could call "*analytical*", which means they unfold a list of required abilities to create artificial intelligence, in part or in whole. […]".* Tom Morisse, "AI New Age, *Fabernovel*, February 2017 https://en.fabernovel.com/insights/tech-en/ais-new-new-age ; See also U.K. Government Office for Science, Report on "Artificial Intelligence: opportunities and implications for the future of decision-making", 2016 (page 6). See also https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/566075/gs-16-19-artificial-intelligence-ai-report.pdf

[2] The first processors in the 1970s could carry out about 92,000 instructions per second. The processor in an average modern smartphone can carry out billions of instructions per second.

[3] IBM estimates that 90 percent of the world's data has been created in the last two years. Looking at various application platforms, experts estimate that Spotify has 10 Petabytes in storage (1 Petabyte = 1 million Gigabyte); eBay has 90 PB; Facebook 300 PB; and Google 15 000 PB. For reference, the human brain has 2.5 Petabyte in storage. https://royalsociety.org/topics-policy/projects/machine-learning/machine-learning-infographic/

[4] Short explanatory infographic from the Royal Society: https://royalsociety.org/topics-policy/projects/machine-learning/machine-learning-infographic/

very large data sets, either via supervised or unsupervised learning.[5] The convergence of two branches of machine learning in particular have demonstrated impressive results over the past five years: deep learning[6] and reinforced learning.

AI vs. Robotics

To better understand Artificial Intelligence as an interdisciplinary field, it is useful to draw and analyze its boundary with robotics. In both cases, we refer to 'machines' (since an algorithm is a robot, hence the shortened word 'bot' to refer to conversational computer programs); but while robotics is mostly material in its manifestations, and operates at the intersection of mechanical engineering, electrical engineering, and computer sciences, artificial intelligence is mostly[7] immaterial and virtual in its manifestations. In order to simplify for analytical purposes, one can say that, in an "autonomous machine," the AI is the intelligence, and refers to cognitive functions, while robotics refers to motor functions.

Indeed, the boundary between cognitive and motor functions is porous, since mobility requires sensing/knowing the environment. For example, advances in machine learning have played a crucial role in computer vision. That said, relying on materiality as a differentiating criterion is useful because it carries major industrial consequences affecting the growth potential of autonomous machines: the more complex the motor functions, the slower the growth, and vice versa. The most popular symbols of the convergence between AI and robotics are self-driving cars and humanoid robots.

AI vs. Neurosciences

To then hone our understanding of the state of AI today and where it could go in the future, we need to turn to its relation with the interdisciplinary field of neurosciences. The renaissance of AI since 2011 is mostly attributed to the success of a branch of machine learning called "deep artificial neural networks" (also called deep learning), supported by another branch called "reinforcement learning". Both branches claim to loosely emulate the way the brain processes information, in the way that they learn through pattern recognition.

It is crucial not to exaggerate the current state of convergence between AI and neurosciences. To date, our understanding of the extremely complex biochemical processes that run the human brain remain far beyond the reach of science. In short, the human brain largely remains a "black box," and neuroscience knows how the brain functions mainly by correlating inputs and outputs. As such, there is not much that designers of algorithms can emulate *from*, especially given that machine learning still operates exclusively from the realm of statistics; that too on silicon-based computer systems, which are radically different from biological brains. A more meaningful convergence between the fields of AI and neuroscience is expected to unfold later this century, as we break into the "black box" and seek to understand the human brain in greater depth.

---

[5] *"There are many different kinds of algorithm used in machine learning. The key distinction between them is whether their learning is 'unsupervised' or 'supervised'. Unsupervised learning presents a learning algorithm with an unlabeled set of data – that is, with no 'right' or 'wrong' answers – and asks it find structure in the data, perhaps by clustering elements together – for example, examining a batch of photographs of faces and learning how to say how many different people there are. Google's News service2 uses this technique to group similar news stories together, as do researchers in genomics looking for differences in the degree to which a gene might be expressed in a given population, or marketers segmenting a target audience. Supervised learning involves using a labelled data set to train a model, which can then be used to classify or sort a new, unseen set of data (for example, learning how to spot a particular person in a batch of photographs). This is useful for identifying elements in data (perhaps key phrases or physical attributes), predicting likely outcomes, or spotting anomalies and outliers. Essentially this approach presents the computer with a set of 'right answers' and asks it to find more of the same. Deep Learning is a form of supervised learning".* U.K. Government Office for Science, Report on "Artificial Intelligence: opportunities and implications for the future of decision-making", 2016 (page 6).

[6] Short explanatory video here from the Royal Society: https://www.youtube.com/watch?v=bHvf7Tagt18

[7] AI refers to a program running on a computer, either embedded or on the cloud. It thus carries a very concrete material manifestation which we tend to forget at times.

Owing to the very different evolutionary trajectories followed by artificial intelligence and our biological brains, two consequential differences should be singled out. First, humans can reliably develop pattern recognition and generalize transferable knowledge out of very few occurrences, but in general we struggle to replicate and transfer learning processes across educational subjects. Machines, on the contrary, require very large data sets[8] to achieve pattern recognition, and struggle to generalize knowledge. However, they excel at transferring and replicating pattern recognition at scale once it is achieved. Facial recognition is the most well-known example of this. Second, while autonomous machines that combine the most advanced AI and robotics techniques are still poor at reproducing very basic non-cognitive motor functions mastered by most animals (for example, walking or hand-manipulation), they are increasingly proving very adept at outperforming humans over a number of complex cognitive functions, for example, image recognition in radiology and computationally-intensive tasks.

Artificial 'Narrow' Intelligence vs. Artificial 'General' Intelligence

The penultimate boundary we need to explore to better delineate and understand what we mean by artificial intelligence is the frontier between *Artificial Narrow Intelligence* (ANI, also called "weak" AI) and *Artificial General Intelligence* (AGI, also called "strong" AI). For a majority of experts, AGI refers to an autonomous machine's ability to perform any intellectual tasks that a human can perform. This implies generalizing and abstracting learning across various cognitive functions. Transferring learning autonomously and nimbly from one domain to another has happened only very embryonically thus far.[9]

According to experts, the most advanced artificial intelligence systems available today, such as the famous *IBM Watson*[10] or Google's *AlphaGo*,[11] are still "narrow" (weak), in the sense that they operate strictly within the confine of the scenarios for which they are programmed. Even if they are capable of generalizing pattern recognition, for instance *transferring* knowledge learned in the frame of image recognition into speech recognition,[12] we are still very far away from the versatility of a human mind. This is expected to change with the convergence of machine learning and neurosciences in the coming decades, but experts disagree profoundly over the probability and timeline of the march towards AGI: some say it will never happen; some say it will take one hundred years or more; some say thirty; and some say ten.[13]

Beyond the discord among experts, relying on the frontier between narrow and general artificial intelligence is problematic because of its very benchmark for measurement: human intelligence. Since

---

[8] As a matter of comparison, a child needs to be exposed to five to ten images of elephant to be able to recognize an 'elephant' while a deep neural networks requires over a million images.

[9] See here the emerging field of "transfer knowledge" perceived by an increasing number of experts, including Google Deepmind as a potential path of accelerated progress in the coming decades. See here for example https://hackernoon.com/transfer-learning-and-the-rise-of-collaborative-artificial-intelligence-41f9e2950657#.n5aboetnm and https://medium.com/@thoszymkowiak/deepmind-just-published-a-mind-blowing-paper-pathnet-f72b1ed38d46#.6fnivpish

[10] See https://www.ibm.com/cognitive/

[11] See https://deepmind.com/research/alphago/

[12] See https://hackernoon.com/transfer-learning-and-the-rise-of-collaborative-artificial-intelligence-41f9e2950657#.n5aboetnm

[13] A detailed study of AI timeline surveys carried out by *AI Impacts* in 2015 concluded: "*If we collapse a few slightly different meanings of 'human-level AI': median estimates for when there will be a 10% chance of human-level AI are all in the 2020s (from seven surveys); median estimates for when there will be a 50% chance of human-level AI range between 2035 and 2050 (from seven surveys); of three surveys in recent decades asking for predictions but not probabilities, two produced median estimates of when human-level AI will arrive in the 2050s, and one in 2085. One small, informal survey asking about how far we have come rather than how far we have to go implies over a century until human-level AI, at odds with the other surveys. Participants appear to mostly be experts in AI or related areas, but with a large contingent of others. Several groups of survey participants seem likely over-represent people who are especially optimistic about human-level AI being achieved soon*". See http://aiimpacts.org/ai-timeline-surveys/

we still have an imperfect understanding today of the complex processes driving the brain and the way human intelligence and consciousness manifest themselves, excessively relying on that boundary to gauge the transformative impact of the rise of AI could be risky. It could expose us to major blind spots, with supposed "advances" masking major socio-economic externalities which we need to anticipate in order to adapt. We recommend doing more research to delineate that boundary and map its surroundings as well as their evolution more precisely.

Beyond their disagreement, experts broadly agree on two levels. First the socio-economic impacts of the current rise of ANI will bring about serious consequences, generating new opportunities, new risks, and new challenges. Second, the advent of an AGI later this century would amplify these consequences by—at least—an order of magnitude. More research is needed to map and understand what these consequences would be as well as how they would play out socially and economically.

The unresolved question of consciousness; and speculations over the possibility of an intelligence explosion

The final boundary we need to explore to map the future terrain of AI is that of consciousness. Here, there is a broad consensus among experts: neither the most advanced AI systems currently existing, nor the ones that are expected to be developed in the coming decades, exhibit consciousness. Machines (programs running on connected and sensing computer systems) are not aware of themselves, and this "functionality" may never be possible. But, again, a word of caution: since science is still far from having explained the mysteries of animal sentience and human consciousness, that boundary remains more fragile that it seems.

Finally, one speculative but highly consequential long-term scenario which constantly appears in mainstream media and across the expert community: "the technological singularity". According to that hotly contested scenario, popularized by the inventor, futurist, and now Director of Engineering at Google, Ray Kurzweil, the rise of AI could lead to an "intelligence explosion" as early as 2045. It would result from the emergence of an Artificial Super Intelligence (ASI): a self-recursive AI improving exponentially, which could follow relatively quickly (a few decades or less) the advent of an Artificial General Intelligence (AGI). If this scenario were to unfold, it would naturally carry with it potentially existential consequences for mankind and intelligent life.[14] We recommend nurturing a reasonable debate across the expert community, and society at large, over the possibilities and consequences of an ASI, to enable responsible investment choices and risk management. Framing the conversation in the right way will be critical: in this case, transparency and moderation will be key. With its unique identity and history, the OECD could be an important policy leader in this field.

To be clear, the analysis we will carry out in the remainder of this chapter excludes the AGI or ASI scenarios. To narrow the definition even further for practical analytical purpose, "Artificial Intelligence" will henceforth mean machine-learning algorithms, which combine various techniques (e.g. deep learning), and are associated with sensors and other computer programs and algorithms. These sense,[15] comprehend,[16] and act[17] on the world, learning from experience and adapting over time.

[14] For more information, see Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, 2014.

[15] Computer vision and audio processing, for example, are able to actively perceive the world around them by acquiring and processing images, sounds and speech. Facial and speech recognition are two typical applications.

[16] Natural language processing and inference engines can enable analysis of the information collected. Language translation is a typical application.

[17] An AI system can take cognitive action like decision-making (e.g. credit application or tumor diagnostic) or undertake actions in the physical world (e.g. from assisted braking to full auto-pilot in cars).

B. Contemporary dynamics and main players

AI pervasiveness

Unlimited access to supercomputing on the cloud—a market estimated to reach $70 billion in 2015—[18] and continued growth in big data, which has had a compound annual growth rate of more than 50 percent since 2010,[19] are the two key macro-trends powering the rise of Artificial Intelligence. AI systems are already profoundly changing the way we live, work, and socialize. On the market are virtual personal assistants, recommendation engines, self-driving cars, surveillance systems, crop prediction, smart grids, drones, banking and trading, and gene-sequencing machines. More and more multinationals are now shifting their business models to revolve around data and predictive analytics to be able to capture the productivity gains generated by the rise of AI.

This revolution is fueled on the one hand by the quest for technological solutions to address pressing global challenges, including climate change, growth and development, security or demography. On the other hand, it is spurred by the continuing international strategic competition whereby nation-states fund science and early innovation in pursuit of technological dominance, which private global players then scale up, competing with others to become "go-to" platforms. Though the ambiguity of the definitional boundaries of "Artificial Intelligence" constrains the ability to generate a robust classification or ranking of most advanced countries in the field of AI, capabilities in the field of computer sciences and Information & Communication Technologies (ICT) can be used as a proxy. Accordingly, the U.S., China, Russia, Japan, South Korea, the U.K., France, Germany, and Israel are emerging as the dominant players in AI. Given their techno-scientific capabilities and their large market size, India and Brazil should also figure in this leading group, even if they are yet to translate potential into reality.

The role of governments

National governments have historically played, and will continue to play, a key role in spurring the rise of AI through the allocation of higher education, research & development budgets for defense, security, healthcare, science and technology (e.g. computer sciences, neuroscience, ICT), infrastructure (especially transport, energy, healthcare, and finance), and pro-innovation policies. AI is increasingly perceived as a source of technological dominance in the information age where cyber and physical worlds merge as hybrids, so more and more countries have or are in the process of releasing national strategies for AI.

In the U.S., where the term Artificial Intelligence was coined, and which has been a pioneer in the field since its inception in the 1950s, the Obama Administration led an inter-agency initiative last year on "Preparing for the Future of Artificial Intelligence."[20] This high-level initiative culminated with the release of a "National Research & Development Artificial Intelligence Strategic Plan,"[21] as well as two reports.[22] Historically, the U.S. Defense Advanced Research Project Agency (DARPA), and more recently the Intelligence Advance Research Projects Activity (IARPA), have provided long-term high-risk investment in AI, playing an instrumental role in most AI techno-scientific breakthroughs. Last year, the U.S. Department of Defense (DoD) unveiled its "Third Offset" strategy[23] with a total five-year investment of $18 billion[24]. To maintain technological dominance, this macro-strategy plans on bringing

---

[18] https://www.accenture.com/us-en/_acnmedia/PDF-33/Accenture-Why-AI-is-the-Future-of-Growth.pdf
[19] https://www.accenture.com/us-en/_acnmedia/PDF-33/Accenture-Why-AI-is-the-Future-of-Growth.pdf
[20] https://obamawhitehouse.archives.gov/blog/2016/05/03/preparing-future-artificial-intelligence
[21] https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf
[22] Executive Office of the U.S. President, "Preparing for the Future of Artificial Intelligence", October 2016. And "Artificial Intelligence, Automation and the Economy", December 2016.
[23] DEPSECDEF, http://www.defense.gov/News/Speeches/Speech-View/Article/606641/the-third-us-offset-strategyand-its-implications-for-partners-and-allies. The "First Offset Strategy" refers to the development of nuclear weapons, the "Second Offset Strategy" to precision guided munitions.
[24] Mackenzie Eaglen, "What is the Third Offset Strategy", *Real Clear Defense*, February 2016. Note: this $18 billion five-year investment goes far beyond Artificial Intelligence.

AI and autonomous systems to the forefront of all U.S. battle digital networks, operational, planning and support processes. DoD's operational goal is to make such processes faster and more efficient. In January 2017, a report published by a group of elite scientists which advises the U.S. Government on sensitive technoscientific matters confirmed the strategic importance of the rise of AI for defense capabilities[25].

Meanwhile, the Chinese Government unveiled a three-year national AI plan in May 2016. The plan was formulated jointly by the National Development and Reform Commission, the Ministry of Science and Technology, the Ministry of Industry and Information Technology, and the Cyberspace Administration of China. The government envisions creating a $15 billion market by 2018 by investing in research and supporting the development of the Chinese AI techno-industrial base. Anecdotally, the country surpassed the U.S. last year in terms of the number of papers published annually on "deep learning."[26] The rate of increase was remarkably steep, reflecting how quickly China's research priorities have shifted.

Beyond U.S. and China, Japan, South Korea,[27] France,[28] the U.K,[29] and Germany are also in the process of developing specific plans and strategies in AI, robotics, and other complementary sectors.

The platform business

From the business perspective, we seem to be heading towards a global oligopoly dominated by a dozen U.S. (*Google*, *Apple*, *Facebook*, *Amazon*, *Microsoft* and *IBM*) and Chinese (*Baidu*, *Alibaba*, *Tencent*, *Xiaomi*) multinationals controlling AI.

For competition played on the global stage, the key factor for success is no longer the length of computer code, but the size of databases. As of now, AI needs to see millions of pictures of animals or cars to achieve actionable pattern recognition. Facebook has effectively relied on the nearly ten billion images published every day by its users to continuously improve its visual recognition algorithms. Similarly, Google DeepMind has relied heavily on YouTube video clips to train its AI image recognition software. In a way, consumers are used as commodities to train AI systems through their behaviors and interactions.

The efficiency of AI systems has also relied on the use of specific microprocessors, which are playing an increasing role in the IT infrastructure on the cloud. For example, the training phase of the deep neural networks has tended to rely on so-called "Graphic Processing Units" (GPUs), processors which were initially designed for video games and have become more powerful over the years[30]. For the implementation phase, digital giants tend to develop dedicated processors. Google, for instance, developed the "Tensor Processing Unit" (TPU), while Microsoft has repurposed "Field Programmable Gate Array" (FPGA).

---

http://www.realcleardefense.com/articles/2016/02/16/what_is_the_third_offset_strategy_109034.html

[25] JASON, The MITRE Corporation, *Report on Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD*, January 2017. https://fas.org/irp/agency/dod/jason/ai-dod.pdf

[26] https://www.washingtonpost.com/news/the-switch/wp/2016/10/13/china-has-now-eclipsed-us-in-ai-research/

[27] South Korea government announced in March last year a $863 million five-year R&D investment in AI. http://www.nature.com/news/south-korea-trumpets-860-million-ai-fund-after-alphago-shock-1.19595

[28] France's government announced in January 2017 it is working on a National AI Strategy to be published in March 2017. http://www.gouvernement.fr/en/franceia-the-national-artificial-intelligence-strategy-is-underway

[29] UK Government announced in January that AI would be at the center of its post-Brexit "Modern Industrial Strategy". http://www.cbronline.com/news/verticals/central-government/modern-industrial-strategy-theresa-may-bets-ai-robotics-5g-uks-long-term-future/. See also U.K. Government Office for Science, Report on "Artificial Intelligence: opportunities and implications for the future of decision-making", 2016 (page 6) https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/566075/gs-16-19-artificial-intelligence-ai-report.pdf

[30] http://www.nvidia.com/object/what-is-gpu-computing.html . See also JASON, *Report on Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD*, (p. 7 & 15). Ibid.

These digital giants are building ecosystems around an "AI tap" that they control, and an intense competition is on to become the "go to" AI platforms which host consumers' and businesses' data. Selling AI through the "software-as-a-service" (SAAS) business model seems to be the route which Google and IBM have adopted. The start-up landscape is also very active in this area. According to CB Insight, the value of AI Mergers & Acquisitions (M&A) has increased from $160 million in 2012 to over $658 million in 2016, while disclosed funding rose from $589 million to over $5 billion over the same time period.[31] Nearly 62 percent of the deals in 2016 went to U.S. start-ups, down from 79 percent in 2012,[32] with U.K., Israeli, Indian, and Canadian start-ups following respectively. The AI market is expected to represent from $40 to $70 billion by 2020, depending on definitional boundaries.[33]

Machine-learning algorithms require a vast amount of data to achieve efficient pattern recognition, so consumer markets' critical mass appears to be a crucial enabler of the establishment of AI techno-industrial bases, in tandem with technoscientific capabilities.

---

[31] CB Insights, "The 2016 AI Recap: Startups See Record High In Deals And Funding", January 2017, https://www.cbinsights.com/blog/artificial-intelligence-startup-funding/ . Important note: these figures don't include the Chinese market.

[32] Ibid.

[33] http://techemergence.com/valuing-the-artificial-intelligence-market-2016-and-beyond/ ; and https://www.bofaml.com/content/dam/boamlimages/documents/PDFs/robotics_and_ai_condensed_primer.pdf

2. Economic, Social and Public Policy Opportunities enabled by AI

A. Efficiency of public and private management

Planning, allocation and monitoring of resources

The rise of AI could first and foremost translate into a revolution in the efficiency of decision-making processes for all actors, both public and private. This, in turn, could give rise to new forms of public-private partnerships. The ability of advanced machine-learning algorithms to mine the growing stocks and flows of data related to the planning and operations of complex systems at the micro or macro levels is likely to trigger a wave of optimization across domains—energy, agriculture, finance, transportation, healthcare, construction, defense, retail and many more—and production factors, including the weather, labor, capital, innovation, information and, of course, the environment.

AI can be essentially analyzed as a "prediction technology,"[34] the diffusion of which could drastically bring down the cost of processing historical data and therefore of making prediction for a wide array of crucial tasks such as risk profiling, inventory management, and demand forecasting. Such a cost decrease would in turn favor reliance on prediction for a growing number of tasks and activities, including and not limited to banking and insurance, preventative health care for patients, predictive maintenance for all types of equipment and complex infrastructure, and crop efficiency through the analysis of satellite or drone imagery.

The optimization potential in terms of resource consumption in complex dynamics is highly significant. Consider the case of energy and its associated carbon emissions. Google DeepMind has already demonstrated how its advanced machine-learning algorithms can be used to reduce energy consumption in data centers. Concluding a two-year experiment cross-analyzing over 120 parameters in a Google data center, DeepMind's artificial neural network worked out the most efficient and adaptive method of cooling and overall power usage. The outcome of the experiment went far beyond traditional formula-based engineering and human intuition. DeepMind claims that this method resulted in a net fifteen percent reduction in overall power consumption, potentially translating into hundreds of millions of dollars worth of savings per year.[35] And the company qualified this as a "phenomenal step forward" given how sophisticated its data centers already are in the field of energy consumption optimization. DeepMind claims that "possible applications of this technology include improving power plant conversion efficiency [...], reducing semiconductor manufacturing energy and water usage, or helping manufacturing facilities in general increase throughput.[36]

Similar predictive approaches are already applied to banking,[37] for product recommendations, advisory services and risk profiling, trading,[38] transportation, traffic management and logistics, healthcare, and meteorology. Firms like Ocado and Amazon are already relying on AI to optimize their storage and distribution networks, planning the most efficient routes for delivery, and making best use of their

---

[34] Ajay Agrawal, Joshua Gans, and Avi Goldfarb, "The Simple Economics of Machine Intelligence", Harvard Business Review, November 2016. https://hbr.org/2016/11/the-simple-economics-of-machine-intelligence

[35] Considering that Google used over 4 million MWh of electricity in 2014 (equivalent to the amount of energy consumed by 366,903 US households), this 15 percent will translate into savings of hundreds of millions of dollars over the years. https://deepmind.com/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-40/

[36] For instance the industrial robotics company Fanuc has teamed up with Cisco to develop a platform to reduce factory downtime—estimated at one major automotive manufacturer to cost US$20,000 per minute. Based on machine learning, Fanuc Intelligent Edge Link and Drive (FIELD) captures and analyzes data from the manufacturing process to improve efficiency. Tantzen, B., "Connected Machines: Reducing Unplanned Downtime and Improving Service," October 6, 2015; and *FANUC*, "Manufacturing Automation Leaders Collaborate: Optimizing Industrial Production Through Analytics," April 18, 2016.

[37] https://thefinancialbrand.com/63322/artificial-intelligence-ai-banking-big-data-analytics/

[38] https://www.wired.com/2016/01/the-rise-of-the-artificially-intelligent-hedge-fund/

warehousing capacity. In healthcare, data from smart phones and fitness trackers can be analyzed to improve management of chronic conditions—including mental illnesses—as well as predicting and preventing acute episodes. IBM Watson is researching the development of automated speech analysis tools running on mobile device to predict the onset of neurological (Huntington's, Alzheimer's, Parkinson's, etc.) and mental (depression or psychosis) diseases for earlier intervention and better treatment planning.[39] The field of "affective computing" aims more broadly at enabling computers to understand and simulate emotions.

Detecting criminal and fraudulent behaviors

Machine-learning has also started to be used to detect early criminal and fraudulent behaviors, and to ensure compliance in innovative ways. One of the first uses of AI in banking was precisely for fraud detection through a continuous monitoring review of accounts activity patterns, with aberrations being flagged for review. With advances in machine-learning, we are now moving towards near real-time monitoring.

Last year, the banking multinational Credit Suisse Group AG launched an AI joint venture with Silicon Valley firm Palantir Technologies, whose solutions are widely used for surveillance and security, to detect unauthorized trading.[40] Credit Suisse started working with Palantir in 2011 after it suffered a $2.3 billion loss on unauthorized trading by Kweku Adoboli. The Zurich-based bank declared its objective is to adapt Palantir AI systems to monitor all employee behavior, so that it can catch breaches of conduct rules. Eventually, it aims to offer this service to other banks.

Besides trading, AI technologies are increasingly being used in the fight against terrorism, and for policing. The U.S. Intelligence Advanced Research Projects Activity is working on a host of programs relying on AI to enhance face recognition for identification[41] based on contextual information –spatial and temporal; or even to automatically detect and geo-localize untagged suspicious videos published online.[42]

Finally, the impact of fake news campaigns on recent elections has prompted Facebook to start working on using AI to help analyze the veracity of the trillions of posts made on the social network.[43] Facebook has started to rely on AI to detect words or patterns of words that might indicate fake news stories.[44]

B. A new wave of productivity gains and growth

Like other great technological revolutions in the past,[45] the largest set of opportunities created by the march of AI technologies results in their ability to trigger a new wave of productivity gains across

[39] https://www.ibm.com/blogs/research/2017/01/ibm-5-in-5-our-words-will-be-the-windows-to-our-mental-health/

[40] https://www.bloomberg.com/news/articles/2016-03-22/credit-suisse-cia-funded-palantir-build-joint-compliance-firm

[41] https://www.iarpa.gov/index.php/research-programs/janus

[42] https://www.iarpa.gov/index.php/research-programs/aladdin-video

[43] http://www.forbes.com/sites/jasonbloomberg/2017/01/08/fake-news-big-data-and-artificial-intelligence-to-the-rescue/#db541e07a214

[44] Peter Kafka, "Facebook has started to flag fake news stories", *Recode*, March 2017. https://www.recode.net/2017/3/4/14816254/facebook-fake-news-disputed-trump-snopes-politifact-seattle-tribune

[45] Elizabeth Eisenstein, *The printing press as an agent of change*, Cambridge University Press, 1980; Robert Hoe, *A short history of the printing press and of the improvements in printing machinery from the time of Gutenberg up to the present day*, 1902. *And Growth and renewal in the United States: Retooling America's economic engine*, McKinsey Global Institute, February 2011.

domains. In this technological revolution, the lynchpins will be machine autonomy and automation.[46] The impacts will be seen on factory shop-floors, service centers, and offices, through the automation of an increasing number of complex cognitive and physical tasks. The rise of AI also means new and more economically efficient forms of collaboration and complementarity between humans and machines. AI can be seen as potentially a new factor of production, enhancing the efficiency of the traditional factors of labor and capital, and creating a hybrid that is capable of creating entirely new workforces. In many cases, AI will be capable of outperforming humans in terms of scale and speed, and it will be capable of self-improvement.

Artificial intelligence can automatize and prioritize routine administrative and operational tasks by training conversational robot software ('bots'), which can then plan and manage interactions. Google's Smart Reply software can already draft messages to respondents based on previous responses to similar messages.[47] Newsrooms are increasingly using machine learning to produce reports and to draft articles.[48] Similar technology can produce financial reports and executive briefings. Robots using lasers, 3D depth-sensors, advanced computer vision, and deep neural networks, can navigate safely and work alongside warehouse and factory workers.

Artificial Intelligence can also generate significant productivity gains by drastically reducing the cost of searching large sets of data manually. This is particularly useful for the legal sector, for instance, where companies like *ROSS*, *Lex Machina*, *H5* and *CaseText* already rely on machine learning for natural language processing, combing through legal documents for case-relevant information. Thousands of legal documents can now be reviewed in a matter of days, as opposed to the traditional method which might take months.[49] In another vein, natural language processing can offer a way of interacting effectively with specialized domain-specific datasets, answering factual questions like IBM Watson Virtual Agent claims it can do.[50]

Productivity gains will not reside solely in the replacement of humans with machines, but also through the advent of new forms of collaboration between humans and machines harnessing the complementarity of biological intelligence with digital intelligence. It is sometimes referred to as "intelligence augmentation." Such novel forms of human-machine teaming are likely to open up a wealth of opportunities for creativity and innovation, translating into higher productivity. One notable example concerns the use of radiology to detect breast cancer, where deep-learning algorithms combined with human pathologists' inputs lowered the error rate to 0.5 per-cent, representing a 85 per-cent reduction in error rates achieved by human pathologists alone (3.5 per-cent) or machines alone (7.5 per-cent)[51].

In terms of economic impact, Accenture published a report in 2016 analyzing twelve developed economies, and claimed that AI has the potential to double their annual growth rates, and increase the productivity of labor by up to 40 per-cent by 2035.[52] In January 2017, the McKinsey Global Institute published its own report on the future of automation. Their definitional boundaries differ from that of

---

[46] "*Autonomy refers to the ability of a system to operate and adapt to changing circumstances with reduced or without human control. For example, an autonomous car could drive itself to its destination. Despite the focus in much of the literature on cars and aircraft, autonomy is a much broader concept that includes scenarios such as automated financial trading and automated content curation systems. Autonomy also includes systems that can diagnose and repair faults in their own operation, such as identifying and fixing security vulnerabilities.*
Automation occurs when a machine does work that might previously have been done by a person. The term relates to both physical work and mental or cognitive work that might be replaced by AI. Automation, and its impact on employment, have been significant social and economic phenomena since at least the Industrial Revolution". See Report on "Preparing for the Future of AI", Executive Office of the President, NSTC, October 2016 (page 10).

[47] https://www.blog.google/products/gmail/smart-reply-comes-to-inbox-by-gmail-on-the-web/

[48] https://www.theguardian.com/media/2016/apr/03/artificla-intelligence-robot-reporter-pulitzer-prize

[49] ABA Journal, "How artificial intelligence is transforming the legal profession", April 1, 2016.

[50] https://www.ibm.com/watson/whitepaper/solutions-guide/

[51] Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad, Andrew H. Beck, "Deep Learning for Identifying Metastatic Breast Cancer," June 18, 2016, https://arxiv.org/pdf/1606.05718v1.pdf

[52] Mark Purdy and Paul Daugherty, *Why Artificial Intelligence if the future of growth*, Accenture, October 2016. www.accenture.com/futureofAI

Accenture's report, and include robotics. Whilst McKinsey's estimate of automation's pace and consequences[53] is more modest, it still offers a very positive vision: automation could raise global productivity by as much as 0.8-1.4 per-cent annually.

Economists had been preoccupied with falling productivity growth rate[54] in recent decades. Attributed to a deficit in innovation, declining working-age population, flagging education attainment and wealth inequality, this productivity growth slowdown has had serious consequences, contributing to slower growth in real wages, and increasing long-run fiscal challenges.[55] According to the McKinsey Global Institute, the expected impact of automation technologies has the potential to match the imperative of high productivity growth needed globally to balance declining birthrates and aging, thereby enabling continued GDP growth.[56] That said, countries will react and absorb the automation wave unequally depending on demography, wage levels, productivity and socio-political appetite for growth and inequality. In principle, advanced economies that have been aging would absorb the impacts of automation more easily and rapidly than emerging economies with an aging workforce.[57]

C. A revolution in healthcare?

In the field of healthcare, advances in AI are setting up what many consider to be truly transformative and life-changing technologies for addressing and tackling human diseases and medicine in general. In 2015, the fast evolution of machine learning and particularly deep neural networks in the field of image recognition enabled computers to surpass human-level performance[58] in accuracy, performance and scaling, and interpreting images. That year at the ImageNet Challenge, an AI developed by the Chinese firm Baidu was wrong only 4.58 percent of the time, beating human image recognition (at an average of 5.1% error rate) for the first time.

These breakthroughs have allowed the field of image and objects recognition to augment radiologists', dermatologists' and oncologists' work in early cancer detections. In fact, all the main AI players are entering the healthcare market. AI is perceived by practitioners as the enabler of the precision and preventive medical revolution. The work of DeepMind[59] is particularly notable, with impressive advances in early detection and treatment of eye diseases and radiotherapy treatments.

Advances in AI are also significantly enhancing the efficiency of Computer-Aided Diagnosis (CAD) and of Content Based Image Retrieval (CBIR), which enhance possibilities for the early detection and

---

[53]James Manyika, Michael Chui, Mehdi Miremadi, Jacques Bughin, Katy George, Paul Willmott, and Martin Dewhurst, *Harnessing Automation for a Future that Works*, McKinsey Global Institute, January 2017. http://www.mckinsey.com/global-themes/digital-disruption/harnessing-automation-for-a-future-that-works

[54] Measured productivity growth has slowed in 30 of the 31 advanced economies, slowing from a 2 percent average annual growth rate from 1994 to 2004 to a 1 percent average annual growth rate from 2004 to 2014.
Jason Furman, "Is this time different? The opportunities and challenges of artificial intelligence," remarks at AI Now: The Social and Economic Implications of Artificial Intelligence Technologies in the Near conference in New York, July 7, 2016

[55] 55 James Manyika, Michael Chui, Mehdi Miremadi, Jacques Bughin, Katy George, Paul Willmott, and Martin Dewhurst, *Harnessing Automation for a Future that Works*, McKinsey Global Institute, January 2017 (p. 95-103).

[56] *Research from the McKinsey Global Institute has shown that even if global productivity growth maintains its 1.8 percent annual rate of the past half century, the rate of GDP growth will fall by as much as 40 percent over the next 50 years. On a per capita basis, the GDP growth decline is about 19 percent. In order to compensate for slower employment growth, productivity would need to grow at a rate of 3.3 percent annually, or 80 percent faster than it has grown over the past half century.*
Global growth: Can productivity save the day in an aging world? McKinsey Global Institute, January 2015.

[57] James Manyika, Michael Chui, Mehdi Miremadi, Jacques Bughin, Katy George, Paul Willmott, and Martin Dewhurst, *Harnessing Automation for a Future that Works*, McKinsey Global Institute, January 2017 (p. 95-103).

[58] https://www.technologyreview.com/s/537436/baidus-artificial-intelligence-supercomputer-beats-google-at-image-recognition/

[59] https://deepmind.com/applied/deepmind-health/research/

identification of medical conditions. This makes possible the use of increasing amount of data (e.g., radiography, ultrasonography, computed tomography, magnetic resonance imaging) which require to be captured, stored, indexed, retrieved and analyzed. These advances enabled, for instance, IBM Watson and doctors from the University of Tokyo's success last year in detecting a rare form of leukemia in a Japanese patient, a diagnosis which had escaped her doctors. Beyond oncology, IBM Watson is now used to mine and make sense of a wealth of unstructured data in heart disease, brain diseases, eye health and diabetes.

Advances in machine learning are considered crucial for the future prospects of the revolution in genome sequencing and editing. Machine learning is poised to transform the cross-analyzation of a wealth of possible combinations of DNA mutations and phenotypic expressions. This will give full potential to new editing techniques already widely used in synthetic biology and clinical research, such as CRISPR -Cas/9. These machine learning advances could prove to be immensely valuable in oncology, by enhancing cancer immunotherapy as well as agronomy.[60]

Other uses of AI in healthcare include:
- Drug inventions, creations and discoveries have hitherto been a fairly random process, too often relying on serendipity (the discovery of penicillin by Alexander Fleming in 1928 is a good example). With the vast amount of data and research publications available for mining, the use of AI can help realize the serendipity potential available through cross-analysis and cross-pollination of existing and emerging discoveries. This could reveal new therapeutic paths in a more systematic way, significantly cutting down the time and expenses involved in the process, including through clinical trials and their approval by regulators.
- Dulight[61] is a small wearable camera, created by Baidu, which uses image recognition to identify objects to help the visually-impaired navigate and better understand their surroundings.
- Personalized healthcare services[62] and life-coaches: from medical appointments on a smartphone to better understanding and integrating various personal health tracking data. Google Now, Cortana (Microsoft) and Siri (Apple) all have enormous potential to make leeway as personalized healthcare assistants.
- Healthcare bots, where patients can interact anytime with a bot for tailored recommendations and services from healthcare providers.
- Elderly care: developments in in-home health monitoring, personal health management, smartphone apps and even robotics are all fueled by AI and Natural Language Processing technologies. The same advances can be seen in emerging visual and hearing assistance technologies, as well as assistive devices relying on AI and machine learning to better understand diverse environments (exoskeleton, electric wheel chairs, intelligent walkers…).

D. A revolution in transportation?

AI is already very significantly impacting transportation[63] with the introduction of autonomous driving capabilities. Advances in deep neural networks are one of the main drivers behind the impressive progress achieved in autonomous vehicles in the past decade, thanks to computer vision in particular.[64]

---

[60] https://www.broadinstitute.org/blog/machine-learning-approach-improves-crispr-cas9-guide-pairing and https://www.mskcc.org/blog/crispr-genome-editing-tool-takes-cancer-immunotherapy-next-level
[61] http://dulight.baidu.com/
[62] https://www.babylonhealth.com/ and https://www.kickstarter.com/projects/1050572498/vi-the-first-true-artificial-intelligence-personal
[63] Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg, Shivaram Kalyanakrishnan, Ece Kamar, Sarit Kraus, Kevin Leyton-Brown, David Parkes, William Press, AnnaLee Saxenian, Julie Shah, Milind Tambe, and Astro Teller. "Artificial Intelligence and Life in 2030." One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA, September 2016. Doc: http://ai100.stanford.edu/2016-report.
[64] See here https://devblogs.nvidia.com/parallelforall/deep-learning-self-driving-cars/

Used in combination with many other types of algorithms, deep neural networks are able to make the most out of the complex combination sensors used for navigation,[65] and can learn how to drive in complex environments. The ongoing transformation may shift mobility into a service industry, with fewer people needing to possess a car, with people able to spend more working or resting time while commuting. This will also impact urban and suburban landscapes, with the possibility that this could address spatial segregation, which has often gone hand-in-hand with social marginalization. Disrupted by the arrival of new actors such as Google, Baidu,[66] Tesla and Uber, the automobile industry is now investing billions of dollars to acquire promising start-ups,[67] and companies are forging alliances and developing in-house capabilities.

While uncertainty prevails regarding the shape and timeline of the industry landscape restructuring, the Boston Consulting Group[68] predicts that connected and autonomous vehicles could help save a very large share of the over 33,000 lives lost every year on U.S. roads (1.3 million globally), which disproportionately affects the youth.[69] There could also be an end to most of the 3.9 million non-fatal injuries (20-50 million globally) seen on the roads. According to a KPMG[70] study published in 2015, accidents should go down by 80 percent by 2040, which will inevitably be a game-changing outcome for the insurance industry and society as a whole.

In a 2015 study, BCG also estimated that the consolidated cost of road insecurity in the U.S. (including the 24 million vehicles damaged every year and the associated overall loss of productivity) comes to almost a trillion dollars per annum.[71] This surplus has the potential to fund the penetration of self-driving technologies, both in vehicles and infrastructure. This will take time: BCG estimates that a 25 percent market penetration globally should take 20-25 years, with high initial prices slowing down adoptions, even if the rise of mobility-as-as-service mitigates the impact of price.[72] This projected adoption curve is also related to the uncertainties surrounding the regulatory environment. Currently, many regulators are in the consultation and/or implementation phase for regulation; however, some markets still fail to address these anticipated technological advances.[73] BCG's report concluded that the global market for autonomous vehicles should grow from $42 billion to $72 billion per annum during the 2025-2035 period.[74]

Other uses of AI in Transportation include:

---

[65] LIDAR (Light Detection and Ranging) to accurately measure distances with a laser, radars (detecting objects using radio waves), ultrasonic sensors and visible cameras.

[66] See https://techcrunch.com/2016/11/17/baidus-self-driving-cars-begin-public-test-in-wuzhen-china/

[67] See the example of Ford Motors: http://money.cnn.com/2017/02/10/technology/ford-argo-self-driving-cars/

[68] The Boston Consulting Group and The Motor & Equipment Manufacturers Association, *A Roadmap to Safer Driving through advanced driver assistance systems*, 2015 (p. 5).

[69] Road crashes are the leading cause of death among young people ages 15-29, and the second leading cause of death worldwide among young people ages 5-14. See http://asirt.org/initiatives/informing-road-users/road-safety-facts/road-crash-statistics

[70] KPMG, *Marketplace of change: automobile insurance in the era of autonomous vehicles*, 2015 https://assets.kpmg.com/content/dam/kpmg/pdf/2016/06/id-market-place-of-change-automobile-insurance-in-the-era-of-autonomous-vehicles.pdf

[71] The Boston Consulting Group and The Motor & Equipment Manufacturers Association, *A Roadmap to Safer Driving through advanced driver assistance systems*, 2015 (intro).

[72] BCG estimates a twenty-year adoption curve. A 25 percent market penetration globally will take 20-25 years with high initial prices slowing down adoptions even if new mobility-as-as service should mitigate the impact. See The Boston Consulting Group, *Revolution in the driver's seat: the road to autonomous vehicle,* 2015. (page 16-17); See also Raj Rajkumar, co-director of the General Motors/Carnegie Mellon Autonomous Driving Collaborative Research Lab: true Level 5 cars won't be on the road for at least a decade or more. http://gas2.org/2016/10/10/level-5-autonomous-cars-decade-away/

[73] http://www.itf-oecd.org/automated-and-autonomous-driving-regulation-under-uncertainty

[74] The Boston Consulting Group, *Revolution in the driver's seat: the road to autonomous vehicle,* 2015. (p. 18)

- Mapping and directions: use of AI (Google Maps) to suggest best itinerary based on traffic data generated by cellphone geo-localization. Similarly, Uber uses Machine Learning[75] to estimate ride times;
- Use of AI by city and transportation planning departments to dynamically manage the best traffic routes for the transportation of people and goods;[76]
- AI for drones and planes.[77]

E. AI and personalized education

Though education is, in the short term, a sector less susceptible to job substitution by AI, with an estimated automation potential of 27 percent,[78] the tools that AI brings to educators and students will prove to be highly valuable in terms of the efficiency of training and retention. The potential for AI in education is geared towards bringing a lifelong learning companion to everyone. One of the most promising areas is in the diffusion of Intelligent Tutoring Systems (ITS) into the market.

As MOOCs[79] (Massive Open Online Courses) and SPOCs (Small Private Online Courses) gain in popularity by giving access to the best course content for all, natural language processing, crowd-sourcing, and machine learning are entering the field, for tasks such as giving assignments and grading them. The field of EdTech is growing, with the market estimated to be worth $250 billion by 2020. AI will help move EdTech in the direction of self-paced, individual learning, with hyper-tailored tutoring (through virtual personal assistants), which will eventually bring personalization of teaching at scale. The example of the Cognitive Tutor[80] tool developed at Carnegie Mellon University is a poignant example of how AI can help shape and deliver the best teaching strategies for individual students, thanks to learning analytics and intelligent feedback to identify and monitor weaknesses, which then informs curricula adaptation.

For educators and universities, plagiarism checkers are already an immensely valuable tool to detect instances of fraud. Machine learning algorithms are now used to move investigations across several languages, to include non-digitized sources, and to make it much smarter than brute force methods. “Robo-readers” or “robo-graders” exhibit a lot of potential for scoring and providing feedback on essay grading, although progress is still needed to reach a sophisticated level of semantic processing. Expected advances in natural language processing in the coming decades should be instrumental. Simulations and gamification have the potential to better engage students through fun “trial and error” tools. Learning environments such as Virtual Reality will allow immersive, high-resolution, hyper-realistic training, and are poised to profoundly affect teaching methods (see the fold-it[81] and mozak[82] human-based computation games, for instance).

Google DeepMind has recently pointed out that the unorthodox masterstroke AphaGo innovated to beat the world champion last year has since allowed Go players to improve their game tactics, by thinking in previously unknown directions. And, as *Wired* Magazine creator Kevin Kelly summarized in his 2016 book *The Inevitable*:[83] “*if AI can help humans become better chess players, it stands to reason that it can help us become better pilots, better doctors, better judges, better teachers*.”

---

[75] http://techemergence.com/examples-of-artificial-intelligence/
[76] ibid ai100 report
[77] AI powering plane autopilots which nowadays allow to manually fly a Boeing 777 for an average of seven minutes only. https://www.nytimes.com/2015/04/07/science/planes-without-pilots.html?_r=0
[78] Ibid McKinsey Global Institute report
[79] EdX: https://www.edx.org/ Coursera: https://www.coursera.org/ Udacity: https://www.udacity.com/ Khan Academy: https://www.khanacademy.org/
[80] http://ctat.pact.cs.cmu.edu/
[81] http://fold.it/portal/
[82] http://www.mozak.science/landing
[83] Kevin Kelly, The Inevitable: Understanding the 12 Technological Forces That Will Shape Our Future, 2016

While education is a sector largely ripe for disruption by AI, barriers of entry are still high, cycles longer and slower, and conservatism more entrenched. As a result, no large actor has yet brought upon a transformational offer. If mass education's intrinsically governmental nature is a great asset to foster diffusion of scaled innovations to all, it is also a liability due to excessive rigidity in most countries. Public-private partnerships offer an interesting path to capture the revolution in education.

F. A safer world?

By understanding norms and variations in user behaviors, AI has already proven to be an efficient tool to counter cyber-attacks and identity theft. AI is now used as a defense against hackers, and engages at the same time in more pro-active measures to prevent hacking, with just-in-time responses available in an ever more sophisticated cyber environment. The world witnessed the game-changing potential of AI in cybersecurity during the DARPA Cyber Grand Challenge competition last summer, with attacks and defense relying on machine learning software systems. An important milestone according to DARPA, the Grand Challenge "*validated the concept of automated cyber defense, bridging the gap between the best security software and cutting-edge program analysis research*."[84]

Beyond cybersecurity, AI will get an increasing weight in defense and security,[85] with its involvement in a range of military platforms, systems and tools. AI will help national security personnel be better informed through smarter battlefield management systems and simulation tools. The extensive use of AI for computer vision and sensing, as well as for in-flight control, has and will continue to enable greater precision strikes. In robotics, advances in AI radically impact tactics (swarm formations), sizes (from Northrop Grumman's RQ-4 Global Hawk High Altitude Long Endurance UAV to micro-surveillance drones), ranges, and mission capabilities of autonomous platforms (land, sea and air; lethal or non-lethal), which are poised to shape and influence the way conflicts and security operations are conducted in the future. Lethal Autonomous Weapon Systems or LAWS have also been at the center of a debate led by the United Nations Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons (CCW)[86] since 2014. The crux of the debate revolves around trade-offs between loss of human control on the one hand, and critically enhanced efficiency and velocity decision-making in the battlefield, on the other.

In policing, AI is used as a powerful identification method (e.g. facial recognition harnessing large networks of surveillance cameras), and increasingly as a predictive instrument pointing where and when crimes can happen. Like in defense, these applications will need to be very seriously governed to avoid abuses, which could have devastating effects. As far as emergency and disaster management are concerned, the impact of machine learning advances' on prediction, planning and management of critical and time-sensitive resources are poised to enhance the performance of emergency relief work. AI technologies will also help aggregate and mine in-time relevant information streams on disaster areas (satellites, drones, social media, geotagging…) to give a clearer picture of where efforts have to be prioritized. In parallel, structuring and making sense of the vast amount of data accumulated on previously disaster-stricken areas will help reduce response time for future event and improve efficiency of aid[87].

[84] http://archive.darpa.mil/cybergrandchallenge/
[85] See AI and the Law of Armed Conflicts co-written by Prof. Thomas Wingfield, Lydia Kostopoulos, Cyrus Hodes and Nicolas Miailhe http://www.thefuturesociety.org/ai-initiative.html
[86] http://www.unog.ch/80256EE600585943/(httpPages)/4F0DEF093B4860B4C1257180004B1B30?OpenDocument
[87] See One Concern and their work on analytical disaster assessment and damage estimates calculus. http://www.oneconcern.com/

3. Policy Challenges

A. Balance of power and global regulation

Because of its business dynamics, the rise of AI has been characterized by the formation of what we could call a "dissymmetric global oligopoly" which raises important issues in terms of wealth and power distribution but could also stifle innovation by preventing the rise of new actors[88]. The rapid evolution of AI technology is a challenge to adjust pro-competition regulations and policies, in terms of understanding the market scope but also possibilities for disruptions or the contestability of monopoly positions. As argued by Jason Furman, former Chairman of President Obama's *Council of Economic Advisers* last year: "*One might imagine that when a large incumbent has access to most of the customer data in the market, it is able to use AI to refine its products better than any potential entrant could hope, and can thereby effectively foreclose entry*"[89].

We refer to an oligopoly because the AI market is dominated by a few multinationals from the U.S. (GAFAMI[90]) and China (BATX[91]) which market capitalization represents over $3.3 trillion dollars today[92]. The AI market exhibits strong "winner-takes most" characteristics because of the particular prevalence of network effects[93] and scale effects[94]. This is not dissimilar to what has been observed in other digital and data markets (such as online search platforms, social media platforms, media delivery platforms) and other platforms in the physical world (such as Internet of Things, data, healthcare, and travel and transportation).

We refer to a "dissymmetric" oligopoly since the latter exhibits serious imbalances between, on the one hand, highly innovative digital multinationals whose disruptive business models unfold transnationally and, on the other hand, nation-states. The latter have so far not been really able to develop global coordination, let alone common ruling mechanisms to favor appropriate levels of competition or on antitrust policies -by looking at data as "critical resource" for competition for instance. This impression of a "governance lag" is probably the result of the high velocity with which the digital revolution has unfolded[95] since the turn of the millennium.

The inability of national governments to create efficient global governance frameworks over the flows and stocks of data, or the reliance on algorithms is also due to profound differences in the value-systems that underpin the way national communities react to the socio-economic transformation (e.g. privacy vs. free speech vs. development vs. national security). That dissymmetry which concerns the digital revolution in general is further compounded by the rise of AI since multinationals are now in a position

---

[88] The Nobel Prize-winning economist Kenneth Arrow (1962) argued that a monopoly translates into weak incentives to innovate. Besides competition pushes firms to invest in new technologies that help to lower costs, and also to invest in innovations that can lead to improvements in the quality of existing products.

[89] Jason Furman, "Is this time different? The opportunities and challenges of artificial intelligence," remarks at AI Now: The Social and Economic Implications of Artificial Intelligence Technologies in the Near conference in New York, July 7, 2016, p. 12.

[90] Google, Apple, Facebook, Amazon, Microsoft, IBM.

[91] Baidu, Alibaba, Tencent, Xiaomi.

[92] Market capitalizations of Google: $582 Bn; Apple: $711 Bn; Facebook: $387 Bn; Amazon: $403 Bn; Microsoft: $499 Bn; IBM: $172 Bn; Baidu: $64 Bn; Alibaba: $246 Bn; Tencent: $256 Bn; Xiaomi: $40 Bn.

[93] "*The network effect is a phenomenon whereby a good or service becomes more valuable when more people use it. The internet is a good example. Initially, there were few users of the internet, and it was of relatively little value to anyone outside of the military and a few research scientists. As more users gained access to the internet, however, there were more and more websites to visit and more people to communicate with. The internet became extremely valuable to its users*". http://www.investopedia.com/terms/n/network-effect.asp

[94] Put very simply, scale effects are when the average unit cost declines at production volume increases.

[95] Google, Facebook, Amazon, Alibaba, Baidu, Tencent and many others multinationals which today rule digital market did not exist in the 1990s.

to capture and deliver key governance functions (e.g. information or knowledge) which were up to know essentially the purview of Nation-States.

The “dissymmetric global oligopoly” of AI could be a source of severe tensions in the coming decades within countries and between themselves. Without regulation and redistribution, the unbalanced accumulation of wealth and power in the hands of a few private actors could favor the rise of populisms and extremisms globally. An unregulated “dissymmetric global oligopoly” could therefore undermine democratic governance or even the development of inclusive global governance mechanisms which we urgently need.

Besides, the anticipation of such a “winner-takes-most” paradigm coupled with the strategic potential of AI at the intersection of defense, security and cyber could lead to an AI arms race. Naturally, this would have dangerous -potentially catastrophic- consequences accelerating the march towards more advanced forms of AI without the adequate governance. In that regard, the increasing tensions between the U.S., Europe, Russia, China and many others around military and paramilitary cyber operations creates a worrisome background when analyzed from an AI perspective. Countries will need to evolve confidence-building measures and coordination mechanisms bilaterally and multilaterally.

B. The challenges of an increasing delegation to autonomous agents

Algorithms’ governance & oversight

Another crucial body of policy challenges associated with the rise of AI relates to the governance of the increasing delegation of competences to autonomous agents. What kind of oversight and accountability norms can we, and should we, enforce over machine learning algorithms? The question already surfaces in a number of critical areas such as determining priorities in line of care in hospitals, social networks feeds, automatic vehicles’ emergency response procedures, citizens’ risks-profiling in the frame of criminal justice procedures, preventive policing or of access to credit and insurance. And there is a clear tension here between the quest for productivity and access on the one hand, and with some of our cornerstone values such as justice, pluralism, collective intelligence, fairness and accountability on the other hand.

The challenge of algorithmic governance is compounded in the case of advanced machine learning techniques by what we call the “black box problem”: the fact that the detailed processes by which deep neural networks self-organize layers of parallel processing to align intended output with actual output remain opaque even to those who design and train these algorithms[96]. The black box problem is not new to computer science but the rise of advanced AI in the age of big data has caused a cardinal shift in its manifestation. According to the most renowned experts, tracing and understanding in details the complex decision-making mechanisms of AI algorithms will be difficult[97]. Researchers have started working on it, but results are far from certain or mature[98].

This question is not trivial on many levels. First, because recent research suggests that the black-box problem might be worse than expected with deep neural networks proving to be surprisingly easy to fool with images that to people look like random noise[99]. Bad-intentioned hackers could potentially learn to

---

[96] Davide Castelvecchi, “Can we open the black box of AI?”, October 2016. http://www.nature.com/news/can-we-open-the-black-box-of-ai-1.20731

[97] Report on “Preparing for the Future of AI”, Executive Office of the President, NSTC, October 2016 (p. 30-34).

[98] Moritz Hardt Eric Price Nathan Srebro, Equal Opportunity in Supervised Learning, October 2016. https://arxiv.org/pdf/1610.02413.pdf ; see also see also Nicholas Diakopoulos and Sorelle Friedler, “How to hold algorithms accountable, MIT Tech Review, November 2016, https://www.technologyreview.com/s/602933/how-to-hold-algorithms-accountable/

[99] Davide Castelvecchi, “Can we open the black box of AI?”, October 2016. http://www.nature.com/news/can-we-open-the-black-box-of-ai-1.20731

exploit these weaknesses, which are now called "adversarial examples attacks"[100]. Second, the European Union's new *General Data Protection Regulation* (GDPR) mandates a *right to explanation* in its article 22[101] which has caused anxiety in the machine learning community[102].

Because of its complexity, developing and implementing algorithmic accountability solutions at scale is going to be costly[103]. Who should bear the associated costs? That is another non-trivial question given that the unregulated global dissymmetric oligopoly pushes actors towards the path of lower cost potentially leading to abuses. And policy-makers will have to work in close symbiosis with AI researchers and engineers to develop balanced mechanisms that address the tension between the need for transparency, and legitimate requirements of secrecy driven by commercial considerations.

Then comes the challenges of deciding over the most adapted institutional frameworks and mechanisms: the route of technical standardization agencies and independent authorities seems preferable wherever technical designs and business models can be reasonably aligned with socially established value hierarchies[104]. But that's more difficult than it seems. And more critically, the epistemic crisis our societies are going through is set to undermine the legitimacy of independent authorities as the public loses trust in experts. The processes determining algorithmic governance will certainly have to pass through more political - i.e. elected - bodies, where the legitimate debates over trade-offs and priorities between values can take place. To be effective, given the stakes and scale, the conversation will have to unfold at the global level, through the U.N. and other institutions and processes including new ones[105].

Conversely, technical standardization bodies and independent authorities will also have to acknowledge that they are not only technical in nature but also political. From that perspective, the *Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems*[106] led by the *Institute of Electronical and Electronics Engineer* – the world's largest technical professional organization and one of the oldest standardization bodies - is a laudable initiative. It appears to genuinely harness IEET's experience in complex standardization processes as well as the strength and inclusion potential of its transnational reach with its global community of over 400 000 practitioners and experts in 160 countries.

And the AI industry has started to self-organize to study, discuss and formulate best practices: the *Partnership on AI*, launched in September 2016 is particularly interesting as it gathers all leading the U.S. players, namely Apple[107], Amazon, Facebook, Google/DeepMind, IBM, and Microsoft. Industry self-policing will not be sufficient as demonstrated by pitfalls in other fields such as finance, but it is

---

[100] Ian Goodfellow, Nicolas Papernot, Sandy Huang, Yuan Duan, Pieter Abeel, Jack Clark, "attacking machine learning with adversarial examples", OpenAI, February 2017. https://openai.com/blog/adversarial-example-research/

[101] https://blog.acolyer.org/2017/01/31/european-union-regulations-on-algorithmic-decision-making-and-a-right-to-explanation/

[102] https://www.wired.com/2016/07/artificial-intelligence-setting-internet-huge-clash-europe/

[103] Andrew Moore, the Dean of Computer Science at Carnegie Mellon University argued in 2016 that the most effective way to minimize the risk of unintended outcomes with AI algorithms is through extensive testing— "*essentially to make a long list of the types of bad outcomes that could occur, and to rule out these outcomes by creating many specialized tests to look for them"*. Report on "Preparing for the Future of AI", Executive Office of the President, NSTC, October 2016 (page 31).

[104] Like we have seen in transportation (safety) or data protection for instance (privacy).

[105] From that perspective the *IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems* is a particularly interesting endeavor. See: http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html

[106] See the first edition of the Initiative's report: *Ethically Aligned Design: A Vision for Prioritizing Human Wellbeing with Artificial Intelligence and Autonomous Systems (AI/AS).* The report was co-created by over 100 thought leaders and experts in AI, ethics and related issues. It is built as a revolving work in progress to be enhanced in time through the participation of the widest possible community globally. http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html

[107] Not a founding member though. Apple joined in January 2017. See https://www.partnershiponai.org/2017/01/partnership-ai-update/

certainly a step needed to move in the right direction, especially if industry bodies succeed in transcending national boundaries to consolidate representation to a certain extent.

Governing the flow of data

As we have seen in the case of the European General Data Protection Reregulation (GDPR), the governance of AI intersects fundamentally with the regulation of data collection, storage, processing, ownership and monetization. In many ways, enabling the potential of AI for growth, development and public good will require agreeing on technical standards and governance mechanisms maximizing the free flow of data and investments in data-intensive services. And this will be difficult because of differing national and regional priorities in terms of value systems and of course because of the competition over the domination and protection of digital markets. It will also be challenging because of the value premium – and associated risks - generated by the high level of uncertainty over how present and future AI technologies can help create, analyze and use data in radically new ways not previously imagined by consumers, firms and governments.

Besides, there could be tensions between the need for stringent data regulations to uphold privacy protection, informed consent and the fight against discrimination on the one hand, and the market conditions required to favor innovation – through lower costs of entry - and to develop robust AI techno-industrial bases able to compete globally on the other hand. The increasing dominance of U.S. and Chinese companies in the global AI oligopoly can also be analyzed as resulting in a way from more flexible regulations governing the flows and stocks of data over time, including "high resolution" personal data. And the European new GDPR which mandates "data portability"[108] as a remedy to enhance competition—an aspect which was also evoked last year by the Chairman of President Obama's Council of Economic Advisers—[109] and suggests that consumers should be compensated for the value that their data generation patterns create should also be analyzed as a response driven to create the conditions of a European "catch up" in the digital economy.

The danger of algorithmic biases and discrimination

Research analyzing the rise of AI has raised severe concerns over the risk of machine learning algorithms that can amplify social biases, and thus become source of discrimination. Similar concerns have been voiced in the context of "Big Data."[110] Such concerns have gained only increased as algorithms grow more complex, autonomous and powerful.[111] AI requires good data, and if the data is incomplete or bias, AI can exacerbate expressions of biases present in society.

The case of "Tay", the teenage AI conversational bot developed my Microsoft, is emblematic of the risks: the firm released Tay on Twitter in March 2016 as an experiment to improve its understanding of

---

[108] "*Article 20 of the GDPR creates a new right to data portability, which is closely related to but differs from the right of access in many ways. It allows for data subjects to receive the personal data, which they have provided to a controller, in a structured, commonly used and machine-readable format, and to transmit them to another data controller. The purpose of this new right is to empower the data subject and give him/her more control over the personal data concerning him or her".* Article 29 WP, *Guidelines on the right to data portability*, WP 242, December 2016
http://ec.europa.eu/information_society/newsroom/image/document/2016-51/wp242_en_40852.pdf

[109] Jason Furman, "Is this time different? The opportunities and challenges of artificial intelligence," remarks at AI Now: The Social and Economic Implications of Artificial Intelligence Technologies in the Near Term, conference in New York, July 7, 2016. See page 12.

[110] The White House, "Big Data: Seizing Opportunities, Preserving Values," May 2014, https://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf; and The White House, "Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights," May 2016, https://www.whitehouse.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf

[111] Report on "Preparing for the Future of AI", Executive Office of the President, NSTC, October 2016 (p. 30-34)

the language used by 18-24 year olds online. The bot had to be shut down within hours, as it started to use racial slurs, defended white supremacist propaganda, and supported genocide.[112]

An often-cited example in the U.S. is the use of apparently biased "risk prediction" tools employed by some judges in criminal sentencing and bail hearings.[113] Other research has questioned the fairness and efficacy of some predictive policing tools[114] or credit scoring tools.[115] Similar issues could also affect hiring practices,[116] and even political campaigns.[117] The use of other data that recruiters consider to be relevant will likely see the need for additional statutory or regulatory protections. For instance, it is important to avoid situations in which algorithms are trained on less diverse workforce samples, as existing group biases may unfairly exclude new talent if and when the algorithms are subsequently integrated into firms' employment practices. Codes of ethics for research, development and commercialization will have to be developed to prevent machine learning algorithms from exacerbating biases.

C. Adapting Social Security and redistributive mechanisms

Creative destruction or destructive creation?

The most discussed set of policy challenges associated with the rise of AI refers to the impact of automation on jobs and inequalities, with some scholars positing the potential "hollowing out" of the middle classes. Experts agree that the automation wave fueled by AI will profoundly impact employment patterns and business processes. How is this time different from previous waves of technological disruption? Whether this "Schumpeterian wave" proves to be a *creative destruction* like those that have come before—resulting in higher average incomes and generating previously unimagined jobs to replace those that get automatized—or turns out to be a *destructive creation*, leading to mass unemployment, depends on the velocity of the development and diffusion of AI technologies over the coming decade. Here, there is significant uncertainty amongst scholars.

Along with President Obama's former Council of Economic Advisers Chairman, Jason Furman's, paper published in July 2016,[118] and the White House report on *Artificial Intelligence, Automation and the Economy* published in December 2016,[119] the McKinsey Global Institute report on *Harnessing Automation for a Future that Works*,[120] released in January 2017, concluded that the fundamental shifts in the labor force caused by automation technologies would be "*of a scale not without precedent*.". In their 2014 book entitled *The Second Machine Age*, Eric Brynjolfsson and Andrew McAfee had on their part argued that we are facing an unprecedented inflection point between the first machine age, based

---

112 James Vincent, Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day", The Verge, March 2016. http://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist

113 Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner, "Machine Bias," *ProPublica*, May 23, 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

114 David Robinson and Logan Koepke, "Stuck in a Pattern: Early evidence on 'predictive policing' and civil rights," *Upturn*, August 2016, http://www.stuckinapattern.org

115 Moritz Hardt Eric Price Nathan Srebro, Equal Opportunity in Supervised Learning, October 2016. https://arxiv.org/pdf/1610.02413.pdf

116 If the data used to train the model reflects past decisions that are biased, the result could be to perpetuate past bias.

117 https://www.theguardian.com/technology/2016/dec/16/google-autocomplete-rightwing-bias-algorithm-political-propaganda

118 Jason Furman, "Is this time different? The opportunities and challenges of artificial intelligence," remarks at AI Now: The Social and Economic Implications of Artificial Intelligence Technologies in the Near Term , conference in New York, July 7, 2016.

119 *Artificial intelligence, automation, and the economy*, Executive Office of the President, December 2016. https://obamawhitehouse.archives.gov/blog/2016/12/20/artificial-intelligence-automation-and-economy

120 James Manyika, Michael Chui, Mehdi Miremadi, Jacques Bughin, Katy George, Paul Willmott, and Martin Dewhurst, *Harnessing Automation for a Future that Works*, McKinsey Global Institute, January 2017 (p.97).

on the automation of physical tasks through mechanization, and a second machine age, based on the automation of cognitive tasks through digital technologies.[121]

Results of studies on the impact of job automation conducted over the past five years have differed quite radically in their assessment and projections: a report from the OECD published in June 2016—[122] focused on its 21 Member countries and centered around "tasks" as a unit of analysis–concluded that a modest average of 9 percent of tasks are automatable. There are predicted to be notable differences between different countries' trends.[123] The 2013 study of Frey and Osborne on the future of employment,[124] which focused on the broader concept of "occupations," had raised alarm bells with its conclusion that about 47 percent of jobs in the U.S. were susceptible to automation over the next two decades. Another report by Citibank,[125] building on the Frey and Osborne study as well as on data from the World Bank, focused on 50 countries and concluded that, on average in OECD countries, 50 percent of the jobs were susceptible to automation. This number was particularly high in India (69% susceptibility) and China (77% susceptibility). Analyzing more than 2,000 work activities across 800 occupations, McKinsey's most recent report concluded that "*about half the activities people are paid almost $15 trillion in wages to do in the global economy have the potential to be automated. […]. While less than 5 percent of all occupations can be automated entirely, about 60 percent of all occupations have at least 30 percent of constituent. More occupations will change than will be automated away*."[126] The report also concluded that activities most exposed include "physical activities in highly structured and predictable environments, as well as the collection and processing of data."

Moving forward, it is paramount that more research is conducted to understand the factors of job automation at more a granular level, in particular across timeframes, sectors, wage levels, education levels, job types, and locations. Reports have hitherto mainly pointed to a continuation, if not an accentuation,[127] of the skill-biased displacement trend,[128] mitigated by the ability of AI and automation technologies to replace high-skill cognitive tasks which exhibit high degree of routine.[129] Some low-skilled tasks requiring advanced hand-dexterity will also remain in demand, at least in the short term. Studies have also highlighted the loss of jobs for some workers in the short-run, but to a substantial degree the time-frame of displacement depends on institution-specific policy responses.

Policy matters: Making the AI revolution work for everyone

---

[121] Erik Brynjolfsson and Andrew McAfee, *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*, W. W. Norton & Company, 2014.

[122] Melanie Arntz, Terry Gregory, and Ulrich Zierahn, *The risk of automation for jobs in OECD countries: A comparative analysis*, OECD Social, Employment and Migration working paper number 189, OECD, May 2016

[123] For instance the share of automatable jobs is 6% in Korea vs. 12% in Austria.

[124] Carl Benedikt Frey and Michael A. Osborne, *The future of employment: How susceptible are jobs to computerisation?*, Oxford Martin School, September 17, 2013.

[125] *Technology at Work v2.0: The future is not what it used to be*, Citibank, January 2016.

[126] *Harnessing Automation for a Future that Works*, McKinsey Global Institute, January 2017 (p. vi). MGI scenarios suggest that half of today's work activities could be automated by 2055 or 20 years earlier or later depending on the various factors, in addition to other wider economic conditions.

[127] It's what Erik Brynjolfsson and Andrew McAfee have called "super-star biased technological change" in their book *The Second Machine Age*. "*It's the fact that technologies can leverage and amplify the special talents, skill, or luck of the 1% or maybe even the 100th of 1% and replicate them across millions or billions of people. In those kinds of markets, you tend to have winner-take-all outcomes and a few people reap enormous benefits and all of us as consumers reap benefits as well, but there's a lot less need for people of just average or above-average skills*". http://www.businessinsider.com/erik-brynjolfsson-2014-1

[128] For instance, the OECD 2016 study estimates that 44 percent of American workers with less than a high school degree hold jobs made up of highly-automatable tasks while 1 percent of people with a bachelor's degree or higher hold such a job. Melanie Arntz, Terry Gregory, and Ulrich Zierahn, *The risk of automation for jobs in OECD countries: A comparative analysis*. Ibid. See also *Artificial intelligence, automation, and the economy*, Executive Office of the President (p. 13 and 14)

[129] *Harnessing Automation for a Future that Works*, McKinsey Global Institute, Ibid. Also see *Artificial intelligence, automation, and the economy*, White House Report. Ibid. (page 23).

Societies' ability to shape the AI revolution into a *creative destruction* and diffuse its benefit to all mainly depends on how they collectively react to it. Technology is certainly not destiny, and policy as well as institutional choices will matter greatly. Our analysis of the most recent literature points to the likely need for progressive tax policies to rebalance the labor to capital shift that is likely to be seen in the AI revolution, in order to protect the most vulnerable from socio-economic exclusion, as well as to avoid an explosion in inequalities of wealth and opportunities. We believe, however that "taxing robots" per se[130] may not be the best option, and could be counterproductive if implemented narrowly, potentially slowing growth and triggering legal challenges.

Systemic policy responses will be required, including reform, and potential reinvention of, Social Security and redistributive tax. Education and skill development systems will also need reforming to enable for repeated and viable professional transitions. Given the difficulty in predicting areas of greater impact and to disaggregate AI-driven automation from other factors (e.g. other technological changes, globalization, reduction in market competition, workers' bargaining power, past public policy choices), policy responses will initially have to target the whole economy, until targeted strategies become more effective, and monitoring and evaluation practices have been designed.

As large swathes of the workforce will be exposed to significant insecurity in the anticipated transition, the reform and enhancement of safety nets has often been suggested as a priority. However, raising minimum wages might paradoxically accelerate automation trends, if used indiscriminately. The opportunity to provide a Universal Basic Income (UBI)—in essence providing a regular, unconditional cash grant—which would revamp social welfare programs in a "post-secure-wage society" driven by automation, is now a feature of political agendas on the Left[131] and on the Right.[132], of course with different contours and degrees.

Economists are archly divided on the matter. Proponents, including Thomas Piketty,[133] see UBI as a way of simplifying the current bureaucratic system, and making it more efficient and fair. UBI is seen as a solution to address the looming automation wave, by favoring work as opposed to unemployment, which has been demonstrated to favor dangerous spirals of marginalization. Attacking mainly the "unconditionality" criteria, opponents[134] denounce an excessively radical and unrealistic approach to reforming existing safety nets. They argue that unconditionality could be counter-productive, resulting in increased, not decreased, income inequality. This camp also argues that UBI could potentially de-incentivize work which they also see as a pillar of social integration.

Interestingly, Finland recently announced a UBI experiment this year,[135] which should provide valuable evidence to move the debate forward. An experiment is also on the cards in the Netherlands though,

---

[130] In a very recent interview Bill Gates advocated for it as a way to slow down the pace of automation and fund professional transitions. See http://fortune.com/2017/02/25/bill-gates-robot-tax-automation-jobs/

[131] In the recent French Socialist Party primary elections held in January 2017, the large victory of Benoit Hamon has essentially been credited to his ability to bring the progressive deployment of a universal basic income as a his flagship measure. See also Andy Stern and Lee Kravitz, *Raising the Floor: How a Universal Basic Income Can Renew Our Economy and Rebuild the American Dream*, June 2016.

[132] Charles Murray, "A guaranteed income for every American", WSJ, June 2016. https://www.wsj.com/articles/a-guaranteed-income-for-every-american-1464969586 ; See also Matt Zwolinsky, « The Libertarian Case for Basic Income », December 2013. https://www.libertarianism.org/columns/libertarian-case-basic-income

[133] Provided UBI targets low wages. See « Pour un revenu universel crédible et audacieux », Le Monde, 25 Janvier 2017. http://piketty.blog.lemonde.fr/2017/01/25/pour-un-revenu-universel-credible-et-ambitieux/

[134] Jason Furman, "Is this time different? The opportunities and challenges of artificial intelligence," remarks at AI Now: The Social and Economic Implications of Artificial Intelligence Technologies in the Near Term, conference in New York, July 7, 2016.

[135] in which 2,000 unemployed people between the ages of 25 and 58 will receive a guaranteed sum – a "basic income" – of €560 a month for two years whether or not they find work. See Sonia Soda, "Is Finland's basic universal income a solution to automation, fewer jobs and lower wages?", Guardian, February 2017. https://www.theguardian.com/society/2017/feb/19/basic-income-finland-low-wages-fewer-jobs?CMP=share_btn_tw

interestingly, the mention of "universal income" has recently been abandoned[136]. In a 2016 referendum, Switzerland rejected, with a 77 percent majority, plans to deploy a monthly $2,555 universal income for all adults.[137]

Besides the UBI, a large variety of more moderate policy options are also on the table. These include strategies to tighten labor markets and pressure wages upward; and the possibility of strengthening workers' collective bargaining power, thus creating new and innovative ways to make workers' voices hear. Such policies aim to rebalance AI-driven concentration, which it is argued could lead to a monopolistic or oligopolistic market. Other available policy options seek a shift in scale rather than in nature of existing societal safety nets. They include the strengthening of existing unemployment insurance schemes to be more targeted or attuned to local environments, or to have their boundaries extended to include, for instance, people who decide on their own to quit their job[138] to pursue training or transition to entrepreneurship. Work-based reform options also include modernizing overtime, employment contracts, with work-sharing programs and temporary work-based training being encouraged.[139]

Finally, to make the AI revolution work for everyone, policy-responses will have to find innovative ways to ensure that a more mobile and insecure workforce has fair access to credit as well as healthcare and retirement benefits. Such benefits have been hitherto largely dependent on employers' contributions, either directly or indirectly – with limited portability when people seek professional transitions, or choose to adapt their employment contracts (e.g. including part-time jobs and entrepreneurship). Challenges to existing employment practices will involve systemic policy change.

D. The case for 21st century education and skill development systems

Reinventing active labor market programs

As we have written above, the wave of automation caused by the AI revolution will displace a very large amount of jobs across domains and value chains. The U.S. "automated vehicle" case study analyzed in the White House 2016 report on *Artificial intelligence, Automation, and the Economy* is emblematic of what's at stake: 2.2 to 3.1 million existing part- and full-time U.S. jobs are exposed[140] over the next two decades, though the timeline remains uncertain. In particular, between1.3 and 1.7 million heavy truck drivers are threatened. And this is not trivial, for the profession has symbolized in the collective imaginary the manifestation of the American dream of empowerment, liberty and social ascension whereby less-educated people could make it into the middle class[141].

---

[136] Deployed in 20 Dutch municipalities, the experiment will allow small groups of benefit claimants to be paid $825 a month while continuing to earning what they make from work. See Daniel Boffey, "Dutch city plans to pay citizens a 'basic income', and Greens say it could work in the UK", Guadian, December 2016. https://www.theguardian.com/world/2015/dec/26/dutch-city-utrecht-basic-income-uk-greens

[137] See http://www.bbc.com/news/world-europe-36454060

[138] That's what French presidential candidate Emmanuel Macron has proposed for instance. See https://www.contrepoints.org/2016/11/13/271472-macron-veut-autoriser-chomage-aux-salaries-demissionnent

[139] A more detailed description can be found in the White House report *Artificial intelligence, automation, and the economy*, Executive Office of the President. Ibid. (p. 34-40).

[140] Though the figures exclude new types of jobs that may be developed in the industry. See *Artificial intelligence, automation, and the economy*, Executive Office of the President. Ibid. (p. 15-17)

[141] Sean Kilcarr, "Defining the American Dream for trucking ... and the nation, too", *Fleetowner*, April 2017

The automation wave calls at least for higher investment and probably the need to reinvent active labor market programs in the coming decades[142]. Such investment should logically be funded by fiscal policies targeting the capital. The 2016 White House report on *Artificial intelligence, Automation, and the Economy* gave an interesting order of magnitude applied to the case of the U.S.: "*increasing funding for job training in the U.S. by six-fold—which would match spending as a percentage of GDP to Germany, but still leave the U.S. far behind other European countries—would enable retraining of an additional 2.5 million people per year*"[143].

AI and other digital technologies offer real potential to innovate new approaches to job-search assistance, placement and hiring processes in the age of personalized services. The efficiency of matching labor supply and demand can be tremendously enhanced by the rise of multi-sided platforms and predictive analytics. The case of platforms such as *LinkedIn* for instance with its 470 million registered users is interesting as an evolution in hiring practices. Tailored counseling and integrated re-training programs also represent promising grounds for innovation.

This, however, won't be enough. A lot will have to be done to create fair and effective life-long skill development/training infrastructure and mechanisms capable of empowering millions of people to viably transition jobs, sectors and potentially geographies – that, too, several times in a lifetime. A lot will also have to be done to address differential geographic impacts which exacerbate income and wealth disparities. Effectively enabling the workforce to be more mobile –both physically, legally and virtually- will be crucial. And this implies of course systemic policy approaches which encompass housing, transportation, licensing, taxes and, crucially in the age of AI, broadband access - especially in rural areas.

To lay solid foundations for this profound transformation, we need more research in at least three complementary areas: first, to devise mechanisms of dynamic mapping of tasks and occupations at risks of automation and associated employment volumes. This mapping of the workforce supply is needed at the macro but also crucially at the micro levels where labor market programs are deployed. Integrated with that, we also need more granular and dynamic mapping of the future jobs/tasks, workplace-structures, associated work-habits, and skill-base spurred by the AI revolution. This mapping of the demand side will be key to innovate, align and synchronize skill development and training programs with future requirements in anticipation, that too on the right timescales. And finally, we need more policy research on the dynamics of professional transitions in different labor market conditions.

To maximize intended impact, create necessary space for trial-and-errors strategies, and to scale up solutions that work, we recommend implementing robust data-driven evidence-based approaches. These approaches should be based on experiments and centered on outcomes in terms of employment but also in terms of earnings. We also recommend exploring new forms of people-public-private partnerships involving civil society as well as new outcome-oriented financial mechanisms such as Social Impact Bonds for instance which could help scale up successful innovations.

A revolution in education?

Understanding components and drivers of AI-labor complementarity and navigating that evolving understanding nimbly to transform primary, secondary and professional education will be capital in the coming decades. Because of the large impact of the rise of AI on economies and societies, this

[142] OECD member countries outside of the U.S. spent, on average, 0.6 percent of GDP on active labor market policies in 2014. The U.S. spent just 0.1 percent of GDP, less than half of what it did 30 years ago. OECD, "Labour market programmes: expenditure and participants", *OECD Employment and Labour Market Statistics* (database), 2016. http://stats.oecd.org/viewhtml.aspx?datasetcode=LMPEXP&lang=en#

[143] This assumes $6,000 per person training/reemployment cost, and an increase in Workforce Innovation and Opportunity Act funding from today's $3B to $18B, to match Germany's spending as a fraction of GDP, with all new funding spent on training. See *Artificial intelligence, automation, and the economy*, Executive Office of the President. Ibid. (p. 33).

implies of course for all countries - almost as a sovereignty imperative, the need to invest in developing AI-related workforce. It is needed to support advances in the field of fundamental research, in the engineering, and of course in the applications, business and socio-political aspects. And the field is by definition interdisciplinary with expanding confines towards biology, cognitive and brain science. Because of the central role of data in developing and training machine learning algorithms, boundaries between fundamental research, applied research, engineering and higher education are likely to blur[144]. We are already seeing a trend whereby fundamental research in AI is shifting away from universities and government laboratories to the biggest technology companies. Academics worry about what they call a "brain drain"[145] which could damage the quality of public research and education down-the-line.

In the *2016 Economic Report of the President*[146], the White House summarized: "*college- and career-ready skills in math, reading, computer science, and critical thinking are likely to be among the factors in helping workers successfully navigate through unpredictable changes in the future labor market*".

Basic literacy and math will more than ever represent the crucial foundation of employability, especially with the accentuation of skill-based job displacement; as will be the quality of early-education since "catching up" will become more difficult; or the need to generalize access to secondary education which should include proven alternatives such as apprenticeship, creative and vocational training[147]. Diversifying and enhancing STEM curriculum beyond computer science to include computational thinking, data science, creativity, innovation and entrepreneurship also appears to be a required evolution.

But beyond that, education will need to change more profoundly and attract the required talent to develop and diffuse innovatively new pedagogies; including centered on emotional intelligence as well as tapping into the power of personalized learning and affective computing. Innovative public-private-partnership should also be explored to favor the emergence of the most effective learning environments and to incentivize good quality investment at scale. But policy-makers will probably retain a key role to ensure innovation diffusion to most, if not all.

As the "online-to-in-person" learning continuum grows more mature, the contours of teachers' role are also very likely to evolve: from that of content providers towards that of content curators, educators, coaches and mentors able to guide learners along personalized path adapted to labor market needs. Crucially, civic education will also need to evolve to equip future citizens with data and AI literacy as well as adequate understanding of trends and stakes, including related to the governance of AI and other emerging technosciences. As we have seen in this study, the serious ethical and political choices abound regarding how societies will decide to collectively embrace the rise of AI. Forging consensus will not be easy, especially considering how the rise of income, wealth, geographic and opportunity disparities may unravel the social fabric both in developed and in emerging countries.

---

[144] The case of Yann LeCun is emblematic. A pioneer in machine learning, computer vision, mobile robotics and computational neuroscience with a long career in academia in France and in the U.S., he joined Facebook as Director of AI Research in 2013 while retaining his position of Professorship at New York University, and simultaneously starting a research partnership between Facebook and New York University's Center for Data Science. He also convinced Mark Zuckerberg to let him run Facebook AI Research operations from New York City creating a dedicated lab there a few blocks from NYU in addition to the laboratories in Menlo Park CA and London. See https://www.facebook.com/yann.lecun/posts/10151728212367143

[145] Richard Waters, "AI academic warns on brain drain to tech groups", *Financial Times,* November 2016. https://www.ft.com/content/298e2ac0-b010-11e6-a37c-f4a01f1b0fa1

[146] The White House, Economic Report of the President 2016, Chapter 4.

[147] Research in the U.S. suggests that apprenticeship fetches a significant premium at a given skill level—as much as $300,000 over a lifetime. Debbie Reed, et al. An Effectiveness Assessment and Cost-Benefit Analysis of Registered Apprenticeship in 10 States. Mathematica Policy Research, 2012.

4. Recommendations

Recommendation 1: Private and public institutions should be encouraged to examine whether and how they can responsibly leverage AI and machine learning in ways that will benefit society. Government organizations should consider partnerships with AI researchers and practitioners that can help apply AI tactics to the broad social problems these institutions already address in other ways.

Recommendation 2: Countries and international organizations should work on developing global coordination mechanisms to collectively monitor the state of AI globally. These should include:

1. Transparency standards over research and development activities;
2. Confidence-building measures regarding the use of machine learning for defense application;
3. Competition policies to address the potential adverse effect of industrial concentration.

Recommendation 3: Countries and international organizations should support more research and foresight on the dynamics (velocity and magnitude) of the rise of artificial intelligence to enable a more detailed mapping and monitoring of potential outcomes to inform adaptation and risk-mitigation strategies. This should include, a focus on:

1. A better understanding of the notion of "Artificial Intelligence" itself, its components and its intersection with other related fields (robotics, big data, computer science) to identify the most appropriate analytical boundaries;
2. A better understanding of the "Artificial General Intelligence" vs. "Artificial Narrow Intelligence" boundary to inform monitoring strategies;
3. A better understanding of the existing and possible paths of convergence between computer science and neuroscience - especially brain science; and their consequences;
4. A better understanding of the technoscientific paths of supercomputing (e.g. quantum computing) to understand future dynamics of Moore's law.

Recommendation 4: Countries and international organizations should support the development of predictive models analyzing the impact of autonomous machines on labor markets, the economy and welfare mechanisms and future jobs/tasks, venture structures and associated skill base. This should include:

1. Research on the interplays between minimum wages and the dynamics of automation;
2. Research on professional transitions in different labor market conditions;
3. Research, development and experiment of dynamic and predictive mapping tools for tasks, occupations and associated skill base with associated employment volumes;
4. Research, development and experiments (data driven and evidenced based) on active labor markets programs, universal basic income, unemployment insurance and other hybrid policy-options;
5. Research on the evolution of power dynamics between capital owners and labor (aggregate and micro levels) and collective bargaining power.

Recommendation 5: Countries and international organizations should support more research in social science on the interaction between science, technology and society in the age of AI.

Recommendation 6: Countries and international organizations should support more research on the need for enhanced statutory and regulatory protections against algorithmic biases (e.g. hiring practice, access to credit or insurance, criminal justice, policing, surveillance)

Recommendation 7: Countries and international organizations should work to create a global policy framework and harmonized norms regulating the flows and stocks of data and especially personal data, and the use of algorithms. This should include:

1. Research on the interrelations between data regulation and machine learning algorithms regulation (e.g. data portability; right to explanation);
2. Research on solutions to make advanced machine learning algorithms (e.g. deep neural networks) controllable, accountable and transparent.

Recommendation 8: Private and public institutions as well as professional organizations should be encouraged to proactively lead public debates engaging AI experts, practitioners, stakeholders (e.g. trade unions) and society at large on the governance of AI. This should include:

1. Ethically aligned design principles of AI and autonomous machines. Actors should for instance actively get involved in the *Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems* led by IEEE or other similar endeavor that exist today and may in the future;
2. Actors should for instance actively get involved in the effort led by *The Future Society at Harvard Kennedy School, AI Initiative* to organize a revolving global conference on the governance of AI; and get actively involved in the ongoing online global civic consultation.