Artificial intelligence: neither Utopian nor apocalyptic impacts soon

Wim Naudé

Published online: 03 Nov 2020.

Submit your article to this journal ⬀

View related articles ⬀

View Crossmark data ⬀

Routledge
Taylor & Francis Group

Artificial intelligence: neither Utopian nor apocalyptic impacts soon

Wim Naudé[a,b]

[a]RWTH Aachen University, Aachen, Germany and IZA Institute of Labor Economics, Bonn, Germany; [b]MSM, Maastricht, The Netherlands and ASC University of Leiden, Leiden, The Netherlands

ABSTRACT

After a number of AI-winters, AI is back with a boom. There are concerns that it will disrupt society. The immediate concern is whether labor can win a 'race against the robots' and the longer-term concern is whether an artificial general intelligence (super-intelligence) can be controlled. This paper describes the nature and context of these concerns, reviews the current state of the empirical and theoretical literature in economics on the impact of AI on jobs and inequality, and discusses the challenge of AI arms races. It is concluded that despite the media hype neither massive jobs losses nor a 'Singularity' is imminent. In part, this is because current AI, based on deep learning, is expensive and difficult for most businesses to adopt, not only displaces but in fact also create jobs, and may not be the route to a super-intelligence. Thus AI is unlikely to have either Utopian or apocalyptic impacts soon. Considering *Amara's Law*, one should however be wary not to underestimate the long-run impacts of AI.

1. Introduction

Technology is the wildcard in humanity's survival, development and future prospects. In the past, various waves of technological innovations and scientific advances facilitated the emergence of *homo sapiens* as the dominant species on the planet[1] (Landes 1999; Mokyr 2002; Ridley 2011; Mokyr 2016; Auerswald 2017; Pinker 2018). These include inventions and discoveries such as the domestication of fire, the wheel, agriculture, cities, the printing press, the microscope, telescopes, the steam engine, the germ theory of disease, electricity, the combustible engine, nuclear science and the personal computer. More recently, a new technology has arrived that has been described as a new general-purpose technology similar to electricity in impact [2]: Artificial Intelligence (AI) (Brynjolfsson, Rock, and Syverson 2017; Trajtenberg 2018; Agrawal, Gans, and Goldfarb 2019).

It is appropriate at the outset to define and describe AI, bearing in mind that there is no generally accepted definition (Van de Gevel and Noussair 2013). Partly this is because intelligence is itself difficult to define. As Tegmark (2017, p. 49) pointed out, 'There's no agreement on what intelligence is even among intelligent intelligence researchers '. The definition proposed by New Scientist (2017, p. 3) is representative and consistent with the earliest views on intelligent machines, defining AI as 'Machines that act intelligently' and can 'make the right decision in uncertain circumstances '. A more descriptive definition[3] is that AI is the 'simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and

the rules for using the information), reasoning (using the rules to reach approximate of definite conclusions) and self-correction'. AI, for the purpose of this paper, is what is termed 'narrow' AI ,[4] which refers to the use of 'machine learning' algorithms in a narrow domain or application to predict and improve predictions by 'learning "from more and more data about a specific domain"' (LeCun, Bengio, and Hinton 2015). It is to be contrasted with the not-yet-available Artificial General Intelligence (AGI) which would be indistinguishable from human intelligence. As with earlier disruptive general-purpose technologies, AI has generated both hopes and fears. Economists for instance have been concerned about the impacts of AI on labor markets, distribution, allocative efficiency, competition and collusive behavior, amongst others (Aghion, Jones, and Jones 2017; Prettner and Strulik 2017; Berg, Buffie, and Zanna 2018; Bloom, McKenna, and Prettner 2018; Agrawal, Gans, and Goldfarb 2019; Furman and Seamans 2019; Schiller 2019; Acemoglu and Restrepo 2020).

These are immediate issues – topics that have also made it onto global and national political agendas. There are however also longer-term concerns, of more existential importance (Yudkowsky 2008; Bostrom 2014) with which economists have not yet engaged with much. For instance, where will continued innovation in AI ultimately lead to? Will narrow AI make way for an AGI? And will this bring about continuing accelerating innovation resulting in a 'Singularity' with exponential economic growth in which a super-intelligence will solve all of humanity's problems (as proposed by Kurzweil (2005)) or one that will intentionally or unintentionally destroy humanity, or perhaps more likely be misused by humanity? There are thus both immediate and long-term issues arising from the recent boom in AI that are of interest to economists. This paper provides a critical survey of the present state of the literature from an economics perspective.

This paper describes how AI research, investment and applications are booming after a number of 'AI winters'. Currently AI is largely applied to enable computers to see, to talk, to process language and to improve business processes and consumer services. It is argued that this 'narrow' AI, which has the 'intelligence of an abacus' is unlikely to lead to a an Artificial General Intelligence (AGI). There has not so far been mass technological unemployment, nor can AI be argued to have been a significant driver of inequality. And it is unlikely to cause mass unemployment or spiraling inequality in the foreseeable future, at least in advanced economies; emerging economies may be different matter – more research is needed in this regard.

The reasons for these conclusions are that (i) the methods used to calculate potential job losses have tended to be very sensitive to the assumptions used and are thus not so reliable; (ii) automation,[5] being rather domain-specific, affects the tasks that humans perform, and not necessarily their jobs; (iii) net job creation can be (and has been historically) positive – because AI stimulates the creation of new jobs or of jobs elsewhere; (iv) the diffusion of AI is much slower than is often assumed: AI is complex and expensive to implement which is reflected in the fact that only a few countries and companies dominate AI investment and research; and (v) the tempo of innovation in AI is slowing down. Some are expecting that another AI winter may be approaching, arguing that decreasing returns has set in in deep learning. Thus at least the short and medium-term impacts of AI are unlikely to be either Utopian or apocalyptic.

The rest of the paper is structured as follows. In Section 2, the current boom in AI is discussed and the nature of AI outlined. In Section 3, the economics literature on the impacts of AI and automation, particularly on jobs and inequality, is reviewed. Section 4 moves on to long term, even imaginary concerns related to the invention of an Artificial General Intelligence (AGI) (a super-intelligence). Here, the dangers of AI-arms races and the need for coherent public policy responses are identified. Section 5 contains a discussion. Section 6 concludes.

2. AI is back

Thou shalt not make a machine in the likeness of a human mind. (Frank Herbert, *Dune,* 1965)

Humans have had an abiding interest in the idea of intelligent machines. A navigational (analogue) computer, the *Antikythera mechanism* ,[6] dating back to 87 BC, has been found sunken off

the coast of Greece. Leonardo Da Vinci designed an automaton (a mechanical knight) in 1495.[7] Interest in AI first seriously arose after the first Industrial Revolution. Two advances were relevant: first, the engineering achievements of the eighteenth and early nineteenth centuries, which include the first use of large industrial machines such as the steam engine; and second, the recognition of evolution by natural selection by Charles Darwin as the process driving the emergence and development of biological species. Inspired by engineering and evolution, Samuel Butler wrote an essay in 1863 entitled *Darwin among the Machines* wherein he predicted that intelligent machines will come to dominate: 'the machines are gaining ground upon us; day by day we are becoming more subservient to them […] that the time will come when the machines will hold the real supremacy over the world and its inhabitants is what no person of a truly philosophic mind can for a moment question '.

The technological and scientific advances of the nineteenth century inspired many works of literary fiction featuring intelligent machines. Much of this is dystopian, somewhat in the spirit of Mary Shelley's *Frankenstein* of 1818. For instance, in 1909 E.M. Forster wrote a novel *The Machine Stops* wherein a post-apocalyptic society is dominated by a super-intelligent machine – and the humans who are left are not too happy about it .[8] In his 1965 novel *Dune* Frank Herbert tells the story of a society in which intelligence machines are so dangerous as to be banned[9] by the commandment 'Thou shalt not make a machine in the likeness of a human mind' (no need for concern: this commandment has not yet been broken). In 1920, Karel Čapek coined the term 'robot' in his 1920 play *Rossums Universal Robots* where intelligent machines (the robots) revolt and wipe out humanity and in 1927 the German movie *Metropolis* featured a similar dystopian future where much of humanity is enslaved in an industrial landscape dominated by machines and robots. The unhappy fictional relationship between human and machine intelligence continues right up to the present: the 2017 HBO TV series *Westworld* likewise dramatizes the uprising of robots against their misuse by humans.

Scientific research into intelligent machines accelerated after the Second World War. Pioneering computer scientist Alan Turing called in an influential paper in 1950 entitled 'Computing machinery and intelligence' for the invention of 'thinking machines '. He argued that the focus ought to be on machines that can play chess and speak English[10]:

> We may hope that machines will eventually compete with men in all purely intellectual fields. But which are the best ones to start with? Many people think that a very abstract activity, like the playing of chess, would be best. It can also be maintained that it is best to provide the machine with the best sense organs that money can buy and teach it to understand and speak English. I think both approaches should be tried. (Turing 1950, p. 460)

The label artificial intelligence (AI) was adopted by the scientific community in 1956 at a conference held at Dartmouth College in the USA (Moor 2006). In 1961 in a paper entitled 'Steps towards artificial intelligence' Marvin Minsky declared that 'we are on the threshold of an era that will be strongly influenced, and quite possibly dominated, by intelligent problem-solving machines' (Minsky 1961). While he was correct in so far as the ICT revolution in general was concerned, as far as AI specifically was concerned, the period between 1956 and 2007 saw research and business interest in AI waxing and waning. AI winters, during which interest and funding in AI declined, occurred in the 1970s and late 1980s up until around 2006[11] (Liu, Shi, and Li 2017). AI winters were the result of the failure of the then mainstream approach, Symbolic Reasoning, to replicate human intelligence (New Scientist 2017). From the late 1980s, cognitive scientists and AI researchers have turned to non-symbolist approaches, inspired by Connectivism, or Artificial Neural Networks (ANN) (Graves, Mohamed, and Hinton 2013). This is reflected, as will be shown below, in the concentration of contemporary AI research in fields like computer vision and speech recognition.

Around 2012[12] a new phase of excitement about AI commenced – AI was back, and with a boom. Venture capital investments in AI-based start-up firms for instance rose from US$ 1.7 billion in 2013 to over US$ 15 billion in 2017, with an acceleration from 2015. Up for grabs is a potential market for AI applications estimated to be worth around $15.7 trillion (PwC, 2017). This boom was the result of a number of confluent developments in ICT that enabled *algorithms*[13] to allow machines to learn from

data, instead of being programmed. These developments included a huge increase in computing power, which has been following exponential growth described by Moore's Law; a significant decline in the price of computing; the emergence of ubiquitous computing through the internet: all with the exponential growth in data (big data) as a result.[14] These confluent ICT developments are discussed by among others Friedman (2016); Brynjolfsson, Rock, and Syverson (2017); McAfee and Brynjolfsson (2017); Tegmark (2017). Engaging historical and futurist perspectives on the rise of humanity and AI are contained in Harari (2011, 2016).

The simultaneous improvements in tools to handle large data-sets, such as machine learning,[15] and specifically the sub-set of artificial neural networks, or deep learning[16] analytics, allowed the use of algorithms that could help machines to learn from data and hence exhibit intelligence. In essence, current AI is based on algorithms and data instead of programs and symbolic logic as was the case in the second half of the twentieth century.

AI is thus nothing like human intelligence or an Artificial General Intelligence (AGI) (a strong AI) which would be comparable to human level intelligence. For this reason, current AI in use is described, as was mentioned, as narrow AI. Whether and when an AGI will be invented are open questions: some see it as unlikely or very far off (Marcus 2015; Koch 2012; Allen and Greaves 2011), while a 2017 survey[17] of AI 'experts' found that they tend to believe on average that there is a 50 percent chance that an AGI will be invented by 2060.

The important points to emphasize at this stage is that (i) AI is not yet comparable to human intelligence[18]; and (ii) that it does not really possess intelligence but rather very domain specific competencies (that they develop by learning from very large datasets) (Jankel 2015). This characterization of AI suggests its shortcomings and pitfalls: it works by making predictions within a specific domain, for instance in playing the board game *Go* it would estimate, based on what it has learned through self-play reinforcement learning which move is the best next move (Silver et al. 2017). Similarly, an automated vehicle would, learning from data from many images of road conditions and road behavior, predict the best course to steer (Qu 2018). As such narrow AI depends on the veracity and velocity of data and the data streaming technology and bandwidth to make good predictions and to learn. And here, the word learn should not be equated with how humans learn. Brooks (2017) correctly emphasized that 'machine learning is very brittle, and it requires lots of preparation by human researchers or engineers, special-purpose coding, special-purpose sets of training data, and a custom learning structure for each new problem domain '.

For companies developing AI, the need for data for machine learning turns data into a valuable (intangible) resource. In 2017 *The Economist Magazine*[19] claimed that 'the world's most valuable resource is no longer oil, but data '. In the data-driven economy, larger firms have an advantage, and as the distribution of firm sizes may become more skew over time (Farboodi et al. 2019). These larger firms, operating globally, may also mis-use data through their dominant positions. For example, in 2018 *Facebook* faced an avalanche of criticisms about the way it exploited the data of its hundreds of millions of users and allowed this data to be exploited, including for political purposes in the notorious *Cambridge Analytica Scandal* .[20]

In February 2019, the UK government's Digital, Culture, Media and Sport select committee's report on Disinformation and Fake News found that 'Facebook, in particular, is unwilling to be accountable to regulators around the world. The Government should consider the impact of such monopolies on the political world and on democracy' (House of Commons 2019, p. 91). In 2018, the European Union adopted a data law (GDPR) to regulate how companies deal with data privacy issues, given the increase in such data and the potential for unscrupulous use of such data by companies for training their AI.[21] And in January 2019, the German government announced its intention to put limitations on the ease with which *Facebook* can make use of the data of its users.[22] Regulating large firms with their scale advantages in generating and utilizing data, and supporting smaller firms to access data and compete, is a topic of increasing interest for economists (Farboodi et al. 2019; Chen 2019).

Most current AI investment and research are in visual systems, natural language processing and speech recognition. This allows machines, both embodied (as in robots) or disembodied (as in chatbots) to see, listen and communicate. Researchers are also attempting to learn robots to be self-aware, through for instance continuous self-modeling (see, e.g. Bongard, Zykov, and Lipson (2006)). These fields or competencies which AI acquire (visual, communication) have however seen many applications – from autonomous vehicles to medical diagnostics to online personal assistants and perhaps most pervasively in customized marketing (see, e.g. Dzyabura and Hauser (2011); Chung, Wedel, and Rust (2016)). Much of the start-up excitement in AI is about to which existing product or service or process the recognition and communication abilities of optimizing algorithms can next be added, and/or which of these it can replace or streamline. AI has been touted to improve anything from beer brewing[23] to personal relationships.[24]

An analysis by WIPO (2019) confirms that R&D in AI is not very broad based in terms of functional areas: the bulk is going into developing visual processing for automotive vehicle development. Comparatively little (visible) research seem to go into AGI. Moreover, only a few countries are responsible for virtually the entire current AI-boom: 93 percent of venture capital goes to firms in only three countries/regions: the USA, China and the EU (WIPO 2019). And within these countries, a few very large high-tech companies dominate: in 2017 five high-tech digital-based companies, namely *Amazon, Alphabet* (*Google*), *Microsoft, Apple* and *Facebook* spend US$ 73 billion on research and development (R&D) – almost twice as much as the entire R&D of the United Kingdom. Adequately training AI on large data-sets through deep learning is subject to decreasing returns and getting more and more expensive at the technological edge. Consider for instance that in 2018 *Google* trained *AlphaGo Zero*, its game (Go) playing algorithm, using over 1000 petaflop/s[25] a day. This is 300,000 times more compute used in deep learning in 2018 than in 2012.[26]

This reflects that current AI R&D and business generation is difficult, expensive and subject to decreasing returns. In particular, smaller companies are at a disadvantage facing barriers to entry in the form of gaining access to data and data scientists (Bergstein 2019; Farboodi et al. 2019). AI may be back, but it is back in a rather exclusive niche.

3. Immediate impacts: the race against the robots

AI will Put 10 million Jobs at High Risk – More Than Were Eliminated by the Great Recession . (CB Insights, 2017)[27]

The immediate impact of AI, as an automation technology,[28] is often described as a 'race against the robots '. To remain competitive in this race it is advocated that humans invest in education to re-skill and re-tool. See for instance the book by Brynjolfsson and McAfee (2012), *Race Against the Machine*, or the *Rise of the Robots* by Ford (2016). Warnings have been sounded that the robots may win this race – resulting in job losses and higher inequality. Autor and Salomons (2017) refer to a possible 'Robocolypse' and Brynjolfsson and McAfee (2015) ask 'Will Humans Go the Way of Horses?'

CB Insights certainly seems to expect a Robocolypse, given their sensational headline quoted above, claiming that AI will eliminate more jobs than the Great Recession. Economists tend to suspect that, without compensating re-distributive tax policies and expansion of education, that AI will by raising the skills premium, widening the wage gap, and decreasing the share of labor in GDP, lead to higher income inequality (Autor 2014; Korinek and Stiglitz 2017; Berg, Buffie, and Zanna 2018).

The rest of this section will first (Section 3.1) describe the general links between technological innovation and jobs and inequality, generalizing the standard effects to AI and discussing recent empirical findings. Then, in Section 3.2 an overview is provided of recent advances in the theoretical understanding, within economics, of AI as an automation technology. Section 3.3 evaluates, in light

of these theoretical advances and in light of recent empirical evidence, the likelihood of AI-driven mass unemployment and spiraling inequality.

3.1. Automation, jobs and inequality

3.1.1. Skill-biased technological change

That technological innovations can have an impact on labor markets, for example through automation, has been long recognized in economics. This is elaborated as the theory of skill-biased technological change (SBTC) wherein technological change can either substitute for or complement various types of labor. This can result in changes in demand, and hence wages, for certain types of labor, hence altering the distribution of wages. During the first Industrial Revolution, the new technologies of the time such as the steam engine and power loom complemented low-skilled labour and substituted for more skilled artisans. The result was a higher demand for low skilled labor and a decline in wage inequality, as the wages for lower skilled labor rose. Since the ICT revolution began (around the 1970s) technological change has been biased in favor of higher-skilled labor, thus raising the demand and wages of more highly skilled labor and even replacing lower skilled jobs. See, for example Autor, Katz, and Krueger (1999); Acemoglu (2002); Card and DiNardo (2002); Goldin and Katz (2010).

3.1.2. Impact on jobs

The possibility that skill-biased technological change can lead to higher unemployment and inequality was emphasized in the years following the global financial crisis by Brynjolfsson and McAfee (2012, 2015); Ford (2016) and by the publication in 2013 of very specific predictions of how much job losses automation will cause. In this regard, Frey and Osborne (2013) (also in Frey and Osborne (2017)) estimated that 47 percent of USA jobs could be automated in 10–20 years. Bowles (2017) estimated this at 54 percent for the EU, and the World Bank predicted up to 66 per cent in developing countries (World Bank 2016).

One particular automation technology that has received further attention is robotics, considered to be the 'leading edge' automation technology at present (Autor and Salomons 2018, p. 3). Acemoglu and Restrepo (2017) for instance calculated that one additional robot per 1000 workers will reduce the employment: population ratio in the USA by 0.37 percent. Using a similar method, Chiacchio, Petropoulos, and Pichler (2018) calculated that one additional robot per 1000 workers can reduce EU employment by between 0.16 and 0.20 percent. Carbonero, Ernst, and Weber (2018) estimated that the 24 percent increase in robots worldwide between 2005 and 2014 caused a global decline in employment of 1.3 percent. They estimate that brunt however is born by emerging economies, who saw a 14 percent decrease in employment as result of robots. Around a third of this decline was due to the reduction in offshoring from advanced economies as a result of automation in these countries. In advanced economies, they estimate that the decline over the 10-year period was less than 0.5 percent (Carbonero, Ernst, and Weber 2018).

While many robots use AI, many do not; robotics do not need to rely on AI. And the extent to which AI has penetrated robotics is difficult to measure (Pissarides 2018). Hence, the question is: why has there been such fears about the job destruction potential of AI as for instance reflected in the quote above from CB Insights?

The answer seems that it is believed that AI is an effective automation technology and that it is rapidly diffused. These beliefs are reflected in arguments that AI is a general-purpose technology and hence as disruptive as electricity or fire (Brynjolfsson, Rock, and Syverson 2017). Moreover, it is believed that AI is progressing at an exponential rate, meaning as Chiacchio, Petropoulos, and Pichler (2018, p. 3) claims with reference to a McKinsey study, that AI fuelled disruption could be 10 times faster and 300 times the scale than during the first Industrial Revolution – thus having '3000 times the impact'.

In Section 4, these claims will be critically discussed. For now, however, it should be noted that given predictions of possible mass unemployment and given the speed of the impact, that virtually all global development organizations[29] and many national governments have been highlighting these concerns in recent years, in particular under the banner of *The Future of Work*. The general recommendation from the debate on the future of work is that, if automation is skill-biased, and hence more likely to replace jobs containing routine-tasks, that humans should invest in skills that are less likely to be able to be automated[30] such as creativity, communication, empathy, entrepreneurship, and skills that facilitate humans' working with advanced technology, such as software engineering (Deming 2015; Trajtenberg 2018). Relatedly, there has been concern about the potential negative impact of AI and automation, as well as of advanced in related digital technologies, can have on the quality of jobs (Schiller 2019). Bloom, McKenna, and Prettner (2018, p. 28) note that automation and digitalization have led to a 'profound change in employment relationships; as a result, workers are often not hired on a permanent basis, but on demand. The catchphrase gig economy has been coined for this new type of labor '.

There has also been increasing interest in the use of better and new forms of social safety nets for the possible large numbers of technologically unemployed, such as a Universal Basic Income (UBI) grant Pulkka (2017); Cords and Prettner (2019). It can be noted for present purposes that what has generally been absent from the debate on the future of work is a discussion of the role of government to stimulate aggregate demand as a way of creating and protecting jobs: most Western governments have been following fairly conservative fiscal policies since the global financial crisis. In Section 3.2 that follows it will be argued that this is a significant shortcoming, as the potential impact of AI on jobs and inequality also depends on aggregate demand. One of the reasons that this may have been neglected so far is that theoretical growth models tend to be supply-driven, assuming that the rise in productive capacity created by technological progress will automatically create its own demand. In recent theoretical and empirical studies, this shows up as 're-instatement effects' (Autor and Salomons 2018; Acemoglu and Restrepo 2020). Below reference will be made to the result of overlapping generations (OLG) models by Sachs, Benzell, and LaGardia (2015); Benzell et al. (2018), and a demand-constrained endogenous growth model proposed by Gries and Naudé (2020) that breaks with the standard growth-model assumption that supply will automatically create its own demand.

Before showing that economic growth theory can offer differing interpretations and solutions to the possible unemployment impact of AI it is necessary to mention that the initial pessimistic predictions on the impact of AI has made place for less pessimistic predictions in recent years that may necessitate a re-assessment of doomsday scenarios. These will be discussed in Section 3.3.

3.1.3. Impact on inequality

Income inequality can rise if skill-biased technological change result in middle-skilled, middle wage jobs to be replaced by an automation technology (such as AI). This has been termed the 'barbell' effect or the 'hollowing out' of the labor market. See for example the contributions by Acemoglu and Autor (2011, 2012), and also Goos and Manning (2007); Autor and Dorn (2013); Beaudry, Green, and Sand (2013). In the case of a potential general-purpose technology (GPT) such as AI, inequality can also increase as a result of the share of labor in total value added declining because more of the rents from the new technology accrues to the owners thereof. Thus innovation rents from AI can drive higher income inequality (Korinek and Stiglitz 2017). Gries and Naudé (2020), incorporating a venture capital type financial sector and vesting ownership of the AI technology not with labor, illustrates how AI innovation rents can arise and drive inequality, as measured by a decline in the share of value-added accruing to labor. Berg, Buffie, and Zanna (2018) simulate the potential impact of various AI automation scenarios and find that in all of these output increases significantly, but that inequality rises rapidly. Their worst case is when robots substitute only for low skilled labor; as they put it 'the magnitude of the worsening in inequality is horrific. In our base case calibration, the skilled wage increases 56–157 percent in the long run while the wage paid

to low-skill labor drops 26–56 percent and the group's share in national income decreases from 31 percent to 8 – 18 percent' (Berg, Buffie, and Zanna 2018, p. 10).

While the potential of technological progress, including in AI, to lead to inequality *a priori* is recognized, it is however not straightforward to ascribe current or past increases in inequality to AI. This is because the labor share in value added and within-country inequality already started to worsen during the 1980s (Rodriguez and Jayadev 2010; Karabarbounis and Neiman 2014; Van Zanden et al. 2014) – when in fact there was an AI winter.

This is not to say that technological change, in particular automation, did not contribute towards some of the rise in inequality since the 1980s. Automation via other technologies than AI, such as the internet, the personal computer and mobile phone, has been found to have played a role, especially in the USA, although the magnitude and significance is a point of disagreement amongst economists, with other factors such as globalization and de-unionization also playing a role (Autor 2014). In the case of Germany, Naudé and Nagler (2018) argue that too little innovation, rather than too fast or too much innovation has contributed to growing income inequality, in the context of globalization and erosion of union power.

Autor et al. (2017) explain the declining labor share not as a result of AI or automation, but with reference to the rise of 'superstar firms' [that is a number of small firms that gain a large size of the market due to fiercer competition] as a result of globalization and digital technologies which enhances consumer power. Hopenhayn, Neira, and Singhania (2018), using a general equilibrium firm dynamics model which they calibrate to USA data, find that the decline in the labor share can be explained by the declining population growth rate. They show that a declining population growth rate leads to a shift in the firm-age distribution towards older firms. And older firms employ relatively fewer 'overhead' labor to 'production' labor. Hence over time, and in aggregate, the labor share will decline as the firm-age distribution shifts.

Even if automation did play a role in the historical decline of the labor share and even in job losses, then this should not be taken to imply that AI so far has had a significant impact, or even that AI is a particularly effective automation technology. So far, hard empirical evidence on AI specifically is lacking. Some of the most interesting recent studies, such as for instance (Autor and Salomons 2018; Bessen et al. 2019) measures the impact of automation either by proxies for technological change such as total factor productivity (TFP), robot density, patents and ICT investments or 'automation costs'; in other words the current state of the art is to use fairly broad measures for automation and not for AI specifically. As Bessen et al. (2019, p. 6) admit they 'do not know the exact automation tech '.

The research cited so far has tended to deal with within-country inequality and with inequality and automation in advanced economies. While it is well established that an important reason for between-country inequality is due to technology gaps (see e.g. Comin and Mestieri (2013)) there is a relative lack of research on the potential impact on automation on global inequality and on emerging economies. The existing research done is, however, pessimistic. For instance, the World Bank (2016) estimated that the jobs susceptible to automation is larger in emerging economies that in developed countries (reflecting the greater labor intensity in these countries). Similarly worrisome is the econometric estimates of Carbonero, Ernst, and Weber (2018) who find that while global employment declined on average by 1.3 percent between 2005 and 2014 due to a 24 percent increase in the number of robots over this period, the decline in employment in emerging economies was up to 14 percent. They conclude [p. 2] that 'these results demonstrate that if there are concerns about automation and robots in particular, these should first and foremost be addressed to emerging economies '.

3.2. Recent theoretical advances

A number of recent additions to theoretical modeling in economics have provided deeper insights into understanding the range of impacts that automation (and AI) may have on key variables such as

the number of jobs, productivity and inequality. Broadly, automation and AI have been modeled either using overlapping generations (OLGs) frameworks (e.g. Sachs, Benzell, and LaGardia (2015); Benzell et al. (2018)), endogenous growth models or labor market models. Within the latter, the main approaches have been to define total output to be produced using either a variety of tasks, in so-called *task-based models* (e.g. Acemoglu and Restrepo (2018b, 2020) while in endogenous growth models, a characteristic feature is to model total output being produced by combining human labor services with a variety of intermediate inputs (e.g. Hémous and Olsen (2018); Gries and Naudé (2020)).

In many of these models, a distinction is made between high-skilled labor and low-skilled labor (e.g. Prettner and Strulik (2017)) or high-tech and low-tech workers (Benzell et al. 2018) and between routine-tasks and non-routine tasks (Acemoglu and Restrepo 2020). Cords and Prettner (2019) introduce a third production factor in addition to labor and capital, which they call 'automation capital'. An automation technology can then be simulated as either complementing or substituting for labor. Most often automation is modeled to substitute for low-skilled labor or routine-tasks and complement high-skilled labor and non-routine tasks. *Routine-replacing technical change* (RRTC) is when automation reduces the demand for middle-skilled labor, as this labor may be performing more routine tasks (Autor and Salomons 2018; Gregory, Salomons, and Zierahn 2019). Exceptions to the idea that automation is skill-biased or only of concern for middle-skilled routine tasks, include Susskind (2017) who proposes a *task encroachment model* wherein output is produced by labor and two varieties of capital: 'traditional' capital and 'advanced' capital, with the latter eventually driving labor out; and Gries and Naudé (2020) who propose a human service input wherein *all* types of labor can be potentially substituted.

To illustrate these theoretical models of automation, consider for example Prettner and Strulik (2017) who uses a R&D-based growth model (see Jones (1995)) with high-skilled and low-skilled labor. A critical assumption they make is that high-skilled labor compliments machines or robots (the automation technology) and that machines and robots substitute for low-skilled labor. They then show how technological innovation drives economic growth in a process accompanied by more automation, an increase in the skills premium, a decline in the labor share and higher unemployment.

Acemoglu and Restrepo (2018a, 2018b) similarly provide a task-based model wherein automation is an expansion of the set of tasks that can be produced with capital. In Acemoglu and Restrepo (2020), they embed the task-based model within a endogenous growth model. These models have as a result that some tasks, traditionally performed by labor, can be taken over by capital and would lead to a reduction (a displacement) of the demand for labor. However, they allow for the adoption of automation to create new tasks and hence raise the demand for labor as a countervailing effect, so that the net impact of automation on jobs will depend on the relative strengths of the 'displacement' effect and the 'countervailing' effects – see also the discussion above. They postulate that AI-driven innovation will be different from other types on innovation in being more likely to generate countervailing effects. In their model, the simultaneous occurrence of growth with higher income inequality, as in Prettner and Strulik (2017), is recognized as an area for further study (as an extension of their model) as they recognize that the rise in income inequality could mean that 'the rise in real incomes resulting from automation ends up in the hands of a narrow segment of the population with much lower marginal propensity to consume than those losing incomes and their jobs' (Acemoglu and Restrepo 2018b, p. 33).

To allow for and model the joint occurrence of low growth and higher inequality due to automation, Benzell et al. (2018); Sachs, Benzell, and LaGardia (2015) use an overlapping generations (OLG) model and Gries and Naudé (2020) an endogenous growth model to explore how the distributive impacts of automation may lead via the demand side to growth constraints. Sachs, Benzell, and LaGardia (2015) model automation through robotics, which they define as machines that 'allow for output without labor' (Ibid, p. 3). They find that an increase in automation can raise output over the short term but will also lead to lower wages and consumption over the longer run, due to the lower demand for labor. Thus over the short-term higher economic growth and

inequality may be simultaneous, but this will turn out to be immiserating growth. This will particularly be likely in their model if the goods produced by robots are close substitutes for good produced by labor, and when traditional capital is a more important factor of production in non-robotic production. Benzell et al. (2018) build on this by focusing not so much on how automation is incorporated into the production function, but rather on what it implies for saving and consumption patterns. Their model predicts a decline in the labor share, tech booms and busts, and 'growing dependency of current output on past software investment' (Ibid, p. 21). Both Sachs, Benzell, and LaGardia (2015) and Benzell et al. (2018) call for more attention to the distributional aspects of automation; in the words of the latter their 'central massage is disturbing. Absent appropriate fiscal policy that redistributes from winners to losers, smart machines can mean long-term misery for all' (Benzell et al. 2018, p.22).

Gries and Naudé (2020) provide an endogenous growth model, with the task-approach embedded, that addresses the central concerns in Sachs, Benzell, and LaGardia (2015); Benzell et al. (2018). They add to the latter imperfect markets and information, allow for the gradual diffusion of AI, and show that progress in AI can be consistent with high employment, low productivity and low growth, and not high productivity and high unemployment as per most of the available endogenous models. This is due to their model allowing non-representative household to have different consumption and savings preferences and hence for aggregate demand to act as a brake on growth – a major point of difference from the Acemoglu and Restrepo (2020) task-based endogenous growth model. Their results suggest, in line with the recommendations of Sachs, Benzell, and LaGardia (2015); Benzell et al. (2018) that further consideration be given to Keynesian-type policies and active labor market policies in order to stimulate aggregate demand as a response to AI-induced declining aggregate demand.

3.3. Why the Robocalypse can be cancelled

This section can now be concluded by summarizing five reasons why, from theoretical and empirical considerations, fears of AI leading to mass unemployment and spiraling inequality may be exaggerated.

The first reason is that the much publicized predictions of mass unemployment were based on doubtful assumptions. Moreover, it has been shown that the methods used to calculate potential job losses in initial reports, e.g. by Frey and Osborne (2013, 2017) are very sensitive to these assumptions. Arntz, Gregory, and Zierahn (2016, 2017) for instance using different – perhaps more tenable – assumptions find that actually only 9 percent of OECD jobs are possibly subject to automation in the next decade.

Secondly, automation tends to affect tasks, rather than entire jobs, e.g. Autor (2015). Some empirical evidence that may be consistent with this, is that job churning in the USA is at historically low levels, despite the spread of AI and robotics through the economy. If AI destroyed jobs (and even created new jobs) job churning would increase, not decrease, as is observed (Atkinson and Wu 2017).

Third, net job creation may be positive as automation (also by AI) lead to creation of new jobs or jobs in other sectors than those negatively affected. There are a number of mechanisms through which such a supply shock can create its own demand and hence reduce the net impact on unemployment. Acemoglu and Restrepo (2018b) identify a 'reinstatement effect' which is the effect of new technology raising labour productivity and wages and thus aggregate demand. Autor and Salomons (2018, pp. 12–13) illustrate the reinstatement effect to be due to an *Uber effect* when consumption of a good increases because its price drops and its quality improves as result of automation; or a *Walmart effect* when automation leads to lower prices of basic goods which then increases the income that consumers can spend on other goods; and a *Costco* effect whereby some upstream processing industries increase their efficiency due to automation and this is passed on to downstream sectors using this as an input.

Empirical studies indeed find that automation has not led to net job losses: see for instance Autor and Salomons (2018) for the OECD countries, Berriman and Hawksworth (2017) for the UK, Dauth et al. (2017); Cords and Prettner (2019) for Germany, and Gregory, Salomons, and Zierahn (2019) for 27 European countries. The latter study found that automation created 1.5 million net new jobs between 1999 and 2010 in Europe. Bessen et al. (2019) study the impact of automation on individual workers in the Netherlands using a dataset covering more than 36,000 workers between 2000 and 2016. Although they find that automation did displace some workers, they did not consider the aggregate employment impact, and conclude that the risk of displacement through automation is 'an order of magnitude smaller than the risk of losing work due to adverse economic conditions' (Bessen et al. 2019, p. 41). AI thus have not yet been so effective as even to show up in individual level data since 2000.

Fourth, the diffusion of AI may be much slower than is commonly thought. It is costly and complex for business firms to utilize AI. This may explain the results discussed in point (iii). Studies by OECD (2015); Andrews, Criscuolo, and Gal (2016); Brynjolfsson, Rock, and Syverson (2017); Bergstein (2019) among others note lagging and leading firms in technology adoption and the existence of implementation lags. Not all companies are staffed by hundreds of PhDs in data science, as the internet giants are (Bergstein 2019). Brynjolfsson, Rock, and Syverson (2017) argue that implementation lags mean that the real impact of AI on productivity (and wages) is yet to be seen. More generally, author such as Gordon (2018) has argued that automation is having a evolutionary rather than a revolutionary impact on jobs in the USA, replacing workers 'slowly' and 'only in a minority of sectors' [IBID, p. 1].

Fifth, the tempo of innovation in AI may be slowing down. There have been AI winters in the 1970s and 1980s. Some are predicting another AI winter, predicting that deep learning is soon to be exhausted (Hao 2019). The innovation in AI may not only be slowing down, but the costs thereof may be underestimated. Based on the arguments of Komlos (2014) there may be a case to make that many AI 'innovations' have been characterized more by Schumpeterian 'destruction' rather than 'creation', that is to say that the net benefits may be overestimated because there is a high degree of substitutability between new AI-driven services and products and the older products and services that they displace. As was noted in Section 2, the bulk of AI applications is of advances in machine vision and speech to make existing products, services and processes somewhat better, not so much yet to provide completely new products, capital and processes. As Cowen (2016, p. 43) remarked 'most new technologies today generate only marginal improvements in well-being '.

In conclusion, the theoretical arguments and empirical evidence surveyed in this section is neither on the side of the tech optimists nor of the tech pessimists: it is unlikely that AI will cause massive job losses in advanced economies, or be the major cause of higher inequality within countries. On the other hand, AI may not create much new jobs either and perhaps even exacerbate global inequality (if its impact on emerging economies are more negative). If AI research and investment are subject to decreasing returns and running up against increasing complexity, the AI-boom may be relatively short-lived: an anti-climax, given current high expectations. Dealing with lack of aggregate demand and stagnating productivity growth, as a result of demographic changes and firm dynamics in hyper-globalized digital markets, may perhaps be the more pertinent policy challenge, rather than the race against the robots.

4. Existential impacts: the fallacy of the giant cheesecake

Now it's clear, that if you're baking a cheesecake, how large a cheesecake you can bake depends on your intelligence. A superintelligence could build enormous cheesecakes – cheesecakes the size of cities. And Moore's Law keeps dropping the cost of computing power. By golly, the future will be full of giant cheesecakes! Yudkowsky (2008:314)

In Arthur C. Clarke's 1968 novel *2001: A Space Odyssey* an intelligent monolith – an AI – from another civilization teaches ancient hominids on Earth to use tools, i.e. develop technology. They develop over time to create their own AI: a supercomputer named **Hal 9000**. In 2001, a human

crew takes to their spaceship to travel to Jupiter to find the makers of the monolith. On its way, Hal exterminates almost all the humans before the captain, Dave Bowman, manages to shut him down with a screwdriver. Dave subsequently finds no trace of biologically based or human-like makers of the monoliths. The AI creators of the monoliths seems to have been only interested in meeting Hal, a fellow AI. They had no interest in the humans.[31]

The ways of humans and AI may split in future. A super-intelligence may have little interest in humanity. What would be the implications? Does this mean that the world is 'hastening toward some sort of digital apocalypse' (Harris 2015) or will a super-intelligence cause a moment of 'Singularity' followed by a tech utopia? The rest of this section will discuss the alignment problem (Section 4.1) and the challenge posed by AI arms-races (Section 4.2). It should be emphasized at the outset that, even if the invention of a superintelligence may not be realized quite soon, or a Singularity be reached, that an understanding of the potential issues at stake, such as the quality of AI, its ethical basis and the dangers that any arms races for AI superiority may result in, is necessary to inform current discussions around the regulation and governance of AI.

4.1. The alignment problem

The moral of *2001: A Space Odyssey* as interpreted by Robert J. Sawyer and recounted above is that 'the ultimate fate of biological life forms is to be replaced by their AIs '. A super-intelligence AGI will have its own agenda, and not necessary those of its human makers. This is basically what Yudkowsky (2008) conveys through the *Fallacy of the Giant Cheesecake*: humans may want giant cheesecakes, and the super-intelligence may be able to deliver, but humanity cannot assume this. More likely the AGI will pursue its own agenda, which may not be aligned with human concerns. Perhaps most guilty of the Fallacy are some members (though not all) of the quasi-religious and religious movements[32] that are inspired by new technologies, such as *Transhumanism*,[33] the *Way of the Future* [34] and the *Singularity*[35] movement.

More formally, the challenge of creating an AGI that is aligned with human interest, the alignment problem, has been discussed under the headings of the control and political problems. According to Bostrom (2017, p. v) the control problem is 'the problem of how to control what the superintelligence would do', in other words the challenge to 'design AI systems such that they do what their designers intend '. The political problem is 'how to achieve a situation in which individuals or institutions empowered by such AI use it in ways that promote the common good' (Ibid, p. 5). Promoting the common good would include steering the development of the AGI such that it prevents the monopolization of the AGI to benefit only a few (Bostrom 2017) and it would also be requiring to minimize the threats emanating from what has been termed Moore's Law of Mad Science: 'every 18 months, the minimum IQ necessary to destroy the world drops by one point' (Yudkowsky 2008, p. 338).

For Kurzweil, proponent of *The Singularity*, there is however no alignment problem and hence no Fallacy. Kurzweil (2005) considers a Singularity as inevitable because technology is developing along an exponential curve. He expects the Singularity, characterized by an 'intelligence explosion' (i.e. an AGI) and exponential economic growth, to occur by 2045. In an interview with Ford (2018, pp. 243–244), Kurzweil explains that an alignment problem is unlikely because of the way the technology evolves and is integrated into human society and even on the individual level with humans: 'It's not us versus the AIs, which has been the theme of many AI futurist dystopian movies. We are going to merge with it. We already have' (Ford 2018, p. 243).

Pinker (2018), no transhumanist nor a futurist, has a fairly similar view[36] on the alignment problem declaring that 'artificial intelligence is like any other technology. It is developed incrementally, designed to satisfy multiple conditions, tested before it is implemented, and constantly tweaked for efficacy and safety' (Pinker 2018, p. 300).

Others, however, such as Yudkowsky (2008); Bostrom (2017) have made more of the alignment problem and have raised warnings that AI may indeed pose an existential risk. And in 2015 more

than 150 scientists and engineers, including luminaries such as Stephen Hawking, Max Tegmark and Peter Norvig, signed an 'open letter on artificial intelligence' in which they stressed the existential threats of AI and called on society to ensure the development of a more robust and safer AI to reduce various risks (such as posed by autonomous lethal weapons), including existential risks (see also Russel, Dewey, and Tegmark (2015)).

4.2. Going beserk: AI arms races

While is there is no agreement on whether or not the alignment problem constitutes a real or imagined problem, the possibility that an AGI can be invented would spur on arms races for AI – not only arms races in the sense of investing in and competing for inventing a superintelligence but also literal races for AI arms, including autonomous lethal weapons. As Harris (2015) describe the conundrum: 'what would the Russians or the Chinese do if they learned that some company in Silicon Valley was about to develop a superintelligent AGI? This machine would, by definition, be capable of waging war – terrestrial and cyber – with unprecedented power. How would our adversaries behave on the brink of such a winner-take-all scenario? Mere rumors of an AGI might cause our species to go berserk '.

A problem with AI arms races is that the participants may cut corners on the design and end up with an inappropriate, second-best AI and even a second best ('unfriendly') AGI (Armstrong, Bostrom, and Schulman 2016). Worse even, the arms race could be won by malevolent agents, intent on misusing the AGI for their own narrow purposes (if they can) (Armstrong, Bostrom, and Schulman 2016; Naudé and Dimitri 2020). The *AI Now 2018 Report*[37] noted that AI arms races create an accountability gap: 'The "move fast and break things" culture provides little incentive for ensuring meaningful public accountability '.

Because of the huge potential reward and winner-takes-all effects that characterize the possession of powerful AIs and ultimately an AGI (that can be controlled…) arms races are a reality, amongst countries such as the USA, China and Russia (see, e.g. Apps (2019)) and among the dominant internet and smart phone giants such as *Alphabet, Amazon, Facebook, Google, Huawei* .[38] Armstrong, Bostrom, and Schulman (2016); Naudé and Dimitri (2020) model the dynamics of AI arms races. The latter use an *All Pay Contest* model to explore how a government could try to limit arms races for an AGI. They note that an AGI has the characteristics of a 'single best effort public good', which is the kind of global public good that requires a deliberate effort of one country or a coalition of countries to be delivered but will benefit all countries in the world once it is available (Barrett 2007).

They call for either public procurement of an AGI innovation by one country or a coalition, in order to incorporate into the design of an AGI certain limitations and ethics; and moreover, to increase transparency and reduce the number of competing teams. Their model also suggest that the taxation of AI can help to reduce some of the incentive to develop an AGI. Their call for public policies such as taxation and procurement to steer the quality of AI and reduce the incentives of high-speed arms races (the 'move fast and break things' culture) is in line with views that the best way to tackle the control problem is through the incremental improvement of the technology; the 'constant tweaking' as Pinker (2018) describes it. Public procurement of a trustworthy AI is also consistent with the call by Cockburn, Henderson, and Stern (2017) call for government support for AI-innovation based on the argument that the private sector may under-invest in AI due to the fact that an AI can innovate by itself, generating positive externalities.[39]

In conclusion, if it is assumed that an AGI is possible, then the question is whether there will be an alignment problem. While this is a topic of disagreement, it is certainly the case that the lure of an AGI will fuel arms-races. These may have sub-optimal outcomes for the quality of even current AI. Misuse of data and market position, collusion, cybercrime, discrimination, poor working conditions, illegal surveillance, the spectre of autonomous lethal weapons, and many other current issues of concern are partly due to lack of transparency, winner-take-all-effects, competition for scarce skills – all issues that are made worse by arms-race conditions.

5. Discussion

The purpose of this paper has been to survey the recent literature in economics on both immediate and long-term issues – benefits and threats – of the recent boom in artificial intelligence (AI). Combinations of beliefs about technological determinism, tech-optimism and the 'animal spirits' that characterize bubbles in investment markets, make the public discourse and public policy formulation in this area very complex – and it is an area that is already very complex due to the inherent nature of AI.

It was useful therefore to survey what scholarly research has so far shown or concluded to be more or less realistic as to what and how the impacts of AI on the economy will unfold – over both the more immediate term as well as over the longer term. This survey has raised six points for further discussion that will be elaborated in this section.

5.1. An Abacus

A first point is whether or not the term artificial intelligence (AI) (introduced during an optimistic period in 1956) is perhaps a misnomer and should be ditched. As was explained, this *narrow* or domain-specific AI consists of deep learning techniques, using large volumes of data. This is not intelligence, even though the answers it can provide can be pretty impressive. It is rather the case that 'current and foreseeable smart technologies have the intelligence of an abacus: that is, zero' (Floridi 2018, p. 157). Some have more disparagingly described narrow AI as 'glorified statistics'. A joke about AI making the rounds[40] goes like this: 'When you're fundraising, it's AI. When you're hiring, it's ML. When you're implementing, it's logistic regression '.

AI is however more than just glorified regression or optimization because of the ability of the software to get better by learning from data. Moreover, precisely how deep learning takes place and how predictions come about is an unknown. This black box quality of the results from deep learning could very well lead to the even slower diffusion of AI over time, for at least two reasons: one is due to given the tightening of regulations on data use, such as the GDPR in the EU[41] and the European Commission's ethical guidelines for trustworthy AI, which includes the requirement of traceability and auditability of AI decisions (EC 2018). Another reason is the growing concern about the use of deep learning in science because results often cannot be adequately explained or replicated.[42]

5.2. The singularity can be cancelled

A second point (which is related to the first) is that an AGI is remote, placing hopes and speculations about a super-intelligence and Singularity in the realm of science fiction rather than of fact. The core problem is that scientists cannot replicate the human brain or human intelligence and consciousness because they do not fully understand it (Meese 2018). Penrose (1989) has (controversially) argued that quantum physics may be required to explain human consciousness. Koch (2012) provides a rigorous criticism from the point of biology of those claiming the imminence of a singularity or super-intelligence, stating that they do not appreciate the complexity of living systems. Dyson (2019) believes the future of computing is *analogue* (the human nervous system operates in analogue) and not digital. Allen and Greaves (2011) describe a 'complexity brake' applying to the invention of a super-intelligence, which refers to the fact that 'As we go deeper and deeper in our understanding of natural systems, we typically find that we require more and more specialized knowledge […] although developments in AI might ultimately end up being the route to the singularity, again the complexity brake slows our rate of progress, and pushes the singularity considerably into the future '.

Even if the invention of a superintelligence may not be realized quite soon, or a Singularity be reached, that an understanding of the potential issues at stake, such as the quality of AI, its ethical basis and the dangers that any arms races for AI superiority may result in, is necessary to inform current discussions around the regulation and governance of AI. Because of the *Fallacy of*

the Giant Cheesecake pursuit of an AGI can lead to an AGI arms race. Certainly, if key players (governments and big corporations) believe that through increased compute narrow AI will eventually lead to an AGI, then such arms races will spill over into the present arms races in AI. These can lead to sub-optimal outcomes which, as Naudé and Dimitri (2020) shows, will need government regulation, taxation and incentivizing to minimize.

5.3. More difficult than it seems

A third point for discussion is the extent to which narrow AI has been taken up and implemented – and what it will take to realize the general-purpose technology potential of AI. Section 3 illustrated that the requirements for businesses to utilize or implement AI are non-trivial. To be precise, current AI R&D and business generation is difficult and expensive[43] (Bergstein 2019; Farboodi et al. 2019). Most AI research and implementation is done by a minority of companies in a minority of countries: in fact, around 30 companies in three regions, North-America, the EU and China perform the vast bulk of research, patenting, as well as receives the bulk (more than 90 percent) of venture capital funding for AI (WIPO 2019).

Smaller companies, which makes up the vast majority of business firms, are at a disadvantage, facing huge barriers to entry in the form of gaining access to and analyzing of data (Farboodi et al. 2019).[44] There seems little prospect of excessive automation, which Acemoglu and Restrepo (2018b, p. 3) speculate is a potential cause of the slowdown in productivity in the West.

While AI is a technique that is more difficult than it seems to develop and adopt, especially in small businesses, it also not clear to many businesses that AI can add value to their business. For example, during the 2020 COVID-19 pandemic, it was soon clear that despite potential, AI yet had little to offer in practical terms to fight the pandemic, and moreover required large and relevant data-sets that were not yet available (Naudé 2020).

Finally, developing AI models at scale will require solving a difficult problem, namely developing 'green AI'. Machine Learning (ML) at the time of writing implied significant environmental cost. Schwartz et al. (2019) determined for instance that 'training a large, deep-learning model can generate the same carbon footprint as the lifetime of five American cars, including gas'.

5.4. The limits of deep learning

A fourth point, that is related to the previous point as well as to the facts presented on the rather narrow concerns of AI (most AI abilities are limited to machine vision, natural language processing, speech, and business intelligence[45]) is that possibility that the current AI-boom may not last very long. Perhaps another AI winter is looming. Deep learning appears to be facing decreasing returns, and further progress in AI may require innovations in hardware (e.g. in edge and quantum computing) (Hao 2019).

On whether another AI winter is looming the jury is still out, although there are signs that the low hanging fruits of deep learning may have been reaped (Hao 2019) and that alternative approaches to AI may be needed in order to better simulate human intelligence (Bellmund et al. 2018; Dyson 2019). Fundamentally, innovation may be getting harder as a result of the burden of knowledge (Jones 2009). Sequeira, Gil, and Afonso (2018) have shown that entropy will limit the extent of technologically driven economic growth and provide estimates of when economic growth will converge to a zero rate. Similarly, Moore's Law is not expected to hold for much longer (Waldrop 2016).

5.5. Message from the empirical evidence…so far

A fifth point is that the impacts of AI, over the short and medium term, on job losses, inequality and productivity have been overestimated. Early predictions were that AI would lead to massive job

losses, large increases in labour productivity, and spiraling inequality. In the meantime, the experience of much advanced economies has been the opposite. Unemployment is at historical lows and productivity growth is stagnating [46]; global inequality is down, and where within-country income inequality has increased, for example in Germany, this has been found to be due to too slow, rather than too fast technological innovation, with technology broadly measured by Total Factor Productivity (TFP) (Naudé and Nagler 2018).

Empirical evidence, still relatively scant but accumulating, has found little evidence of net job destruction as a result of automation in advanced economies. Theoretical economic models have also developed to reflect better with the nature of an automation technology that can both complement and substitute for human labor: these have variously shown that automation can affect tasks within jobs rather than jobs itself, can increase labor demand through (for instance 're-institution' of labor due to labor productivity and wages rising) and moreover that AI progress can be consistent with low wage and productivity growth, relatively high employment, but with economic stagnation, entropy, and a decline in the quality of jobs (Sequeira, Gil, and Afonso 2018; Schiller 2019; Gries and Naudé 2020).

The empirical evidence is however subject to two shortcomings. First, most studies have focused on advanced economies. The few studies from emerging economies (e.g. World Bank (2016); Carbonero, Ernst, and Weber (2018)) tend to suggest that the effects of automation may be worse. A second shortcoming is that most studies do not measure the impact of AI directly, but rather via various measures of 'automation' of tasks, using various broader indicators such as industrial robots and 'automation costs' which may or may not be appropriate proxies. It seems the assumption, still to be rigorously tested, is that AI is an effective automation technology, or similar to previous automation technologies. Given however the comments above on the slow diffusion of AI, and the relatively slow penetration of AI to many emerging economies outside of China, more research is needed to better understand AI as an *automation* technology, and to understand the potential vulnerabilities of emerging economies to AI.

5.6. A misinformation epidemic

A final point is that if one considers the nature of machine learning (and deep learning), the relatively few firms and countries driving AI innovation, and the somewhat limited basic core functionalities being developed (mostly computer vision and language), as well as perhaps most importantly the slow and expensive diffusion of ML outside the giant tech companies, then the exaggeration or hype in the media about AI becomes even more accentuated.

Schwartz (2018) describes the current media discourse about AI as '*unhinged*' and quote Lipton that the world is suffering an 'AI misinformation epidemic '.[47] Despite the lack of evidence, and prospects, for mass replacement of jobs by automation, a 2018 Poll by the Pew Research Center[48] found that 65 percent of Americans still fear that their jobs will be taken over by robots and computers within 50 years.

According to Brooks (2017), there are a number of reasons why this misinformation and hype about the impacts of AI occur. Perhaps the two most relevant for the present discussion are that (i) people have a tendency to overestimate the impact of a new technology over the short-term but to underestimate it over the longer-term (*Amara's Law*) and (ii) the slow rate of diffusion of a new, especially new digital technology, is much slower and evolutionary, rather than revolutionary, than most people think.

Brooks (2017) explains this last point referring to the slow pace of replacement of hardware and physical infrastructure, which limits the extent to which a digital technology such as AI can spread, stating as an example that 'many of the cars we are buying today, which are not self-driving, and mostly are not software-enabled, will probably still be on the road in the year 2040. This puts an inherent limit on how soon all our cars will be self-driving '. It is not only sunk investments in hardware and physical capital that slows down the adoption and impact of the newest digital technologies, but also the likelihood that different computer hardware is needed to continue to make

progress in AI (see related to this the discussion on the slowing down of progress in AI in Section 3). For instance, at the time of writing machine learning use large amounts of parallel processing power, which is mostly delivered through servers in the cloud, as opposed to capacity on a local smart phone or hardware. This creates what is called the *von Neumann Bottleneck* which is the problem that 'data need to be transferred back and forth from one place to the other overtime a calculation takes place' (Zarkadis 2018). To overcome the von Neumann Bottleneck, more investment and research may be needed in the field of edge computing. According to Batra, Queirolo, and Santhanam (2018) 'inference at the edge will become increasingly common for applications where latency in the order of microseconds is mission critical '.

6. Concluding remarks

Hollywood's vision of AI is often entertaining, generally pessimistic and rarely realistic. (Greenemeier 2017)

While humans' fascination with intelligent machines dates back centuries, and the term itself was coined back in 1956, the most recent infatuation can be traced back to around 2012 when the confluence of big data, cloud computing, the internet, and exponential changes in computing power and computing cost, allowed machine learning algorithms to perform exceedingly complex tasks, many associated with human intelligence. It can even be argued that the world is seeing a AI-boom, reflected in a sharp increase in scientific publications, patents publications, and venture capital investment in AI-start-ups.

The interest in AI has generated hopes and fears. It has generated hopes that AI will improve productivity and growth prospects, and even accelerate progress towards a tech-utopia, the Singularity, which may be achieved when AI advances to the extent that it becomes a super-intelligence (AGI). But it has also generated fears that AI would negatively impact on jobs and inequality or even over the longer-run pose an existential threat to humanity.

This paper has argued that the economics literature has been converging on a view that is, in contrast to that of the typical Hollywood film, less pessimistic and more realistic (even if perhaps marginally less entertaining). In particular, the economic view is that while AI is improving business process efficiency and generating consumer surpluses, and automating many tasks traditionally performed by humans, it has *so far* not led to mass technological unemployment. AI is also not (yet) a main driver of income inequality.

AI is unlikely to cause huge job losses (or job creation) in the foreseeable future, at least in advanced economies (there seem to be valid concerns that emerging economies may face more significant impacts: but more research is needed). The main reasons for this conclusion are based on (i) the fact that the methods used to calculate potential job losses are sensitive to assumptions used; (ii) automation may affect tasks more significantly, rather than the jobs within which they are performed; (iii) net job creation can be positive because automation stimulates the creation of new jobs or jobs elsewhere; (iv) diffusion of AI may be much slower than is thought or assumed; and (v) the tempo of innovation in AI is slowing down – some are even predicting another AI winter.

As far as the ability of AI to cause a Singularity or pose an existential threat to humanity are concerned, this paper has concluded that a *Singularity* is out of reach and moreover unlikely in near future. Current research into deep learning, which is dominating research on AI, is probably not the method for inventing an AGI. A blissful Singularity and a Robocalypse thus presently belong to the realms of science fiction. Of course, considering *Amara's Law*, this paper may be grossly underestimating the longer-run impacts of AI.

Finally, it was concluded that concern over AI arms races may be justified nevertheless. This may be particularly the case in light of the 'misinformation epidemic' concerning AI. The hype and hysteria about AI has led to an 'unhinged' debate about AI, which may have as consequence stifling regulations as well as detrimental AI arms races. These consequences could hasten a premature AI-winter through

inappropriate controls and a loss of public trust, unfortunately at a time when the world needs more, not less, technological innovation, and for this technology to diffuse much faster.

Notes

1. Philosophers of technology and historians of technology have typically differed on the extent of technological determinism (see, e.g. Misa (2009)). There is, however, no doubt amongst scholars and policy makers as to the importance of technology in human society.
2. According to the CEO of Google, '*AI is probably the most important thing humanity has ever worked on. I think of it as something more profound than electricity or fire*' (see https://bit.ly/2vY3n3Z). As a sizeable chunk of Google's billion-dollar R&D budget is on AI, he would say this.
3. https://searchenterpriseai.techtarget.com/definition/AI-Artificial-Intelligence
4. Ford (2016, p. 225) stresses that 'Even IBM's Watson, perhaps the most impressive demonstration of machine intelligence to date, doesn't come close to anything that might reasonably be compared to general, human-like intelligence. Indeed, outside the realm of science fiction, all functional artificial intelligence technology is, in fact, narrow AI'.
5. Much of the existing economic literature on the topic is concerned with automation. Automation is not new of course – it has always characterized production and can be achieved through application of many kinds of technology. For example, the power loom of 1784 automated the task of weaving, and was a key enabling technology of the 1st Industrial Revolution. AI is only one specific technology that can automate tasks.
6. See https://bit.ly/2nNL0Ib
7. See http://history-computer.com/Dreamers/LeonardoAutomata.html.
8. Did E.M. Forster anticipate the internet and smart phone in 1909? See the discussion in https://www.bbc.com/news/entertainment-arts-36289890.
9. See https://bit.ly/2NxxjJ1.
10. At the time of writing (2019) both approaches have indeed been tried and implemented. E.g. in chess there has been IBM's Deep Blue algorithm and Google's *AlphaZero*. The latter is reported to have taught itself in 2017 to play chess in four hours and then to defeat other leading chess computer programs. And as far as talking is concerned, in 2018 Google introduced *Google Duplex*, an AI telephone assistant that can conduct a complex conversation.
11. According to Liu, Shi, and Li (2017), the most recent AI winter came to an end after 2006 with breakthroughs in the technique of deep learning by Hinton and Salakhutdinov (2006); Hinton and Osindero (2006).
12. The year 2012 can be chosen as a convenient starting date for the current boom in AI research and applications, based on recent data from WIPO (2019) showing that the sharp increase in patenting started around this date. Venture capital investments in AI similarly surged after 2012. Moreover, 2012 comes suitably and soon after 2006 and 2007, the dates that has been identified first for breakthroughs in machine learning (specifically deep learning) (see Liu, Shi, and Li (2017) and the confluence of a number of key technologies that enabled AI (such as the smart phone and cloud computing) responsible for the exponential increase in data and connected things (see Friedman (2016)).
13. An algorithm is 'a sequence of rules that should be performed in an exact order to carry out a certain task' OECD (2017, p. 8).
14. It is estimated that around 2.5 quintillion gigabytes of data were created daily in 2017 (see https://bit.ly/2fl2JCT).
15. Machine learning can take place through decision-tree learning, association rule learning, artificial neural networks, deep learning, inductive logic programming, support vector machines, clustering, Bayesian networks, reinforcement learning, representation learning, similarity and metric learning, genetic algorithms and rule-based machine learning (Tien 2017).
16. 'Deep-learning methods are representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level […]higher layers of representation amplify aspects of the input that are important for discrimination and suppress irrelevant variations' (LeCun, Bengio, and Hinton 2015, p. 436).
17. See https://blog.aimultiple.com/artificial-general-intelligence-singularity-timing/.
18. Although it is the case that AI is not yet comparable to human intelligence this does not mean that it will not in future develop into an AGI. The only point that I wish to make for now is that at currently AI is not much more than deep learning and that more skepticism is warranted in the public discourse, which as I will deal with in more detail in Section 5.6, tend towards hype and hysteria when it comes to AI.
19. See https://econ.st/2Gtfztg.
20. See https://en.wikipedia.org/wiki/Facebook–Cambridge_Analytica_data_scandal.
21. See https://eugdpr.org.
22. See https://nyti.ms/2rXdxMJ.

23. Want to know more about the Artificial Neural Beer Network (ANBN)? : https://bit.ly/2GZQyK9.
24. Will your relationship with your better half last? Check https://bbc.in/2HgFQRk.
25. A petaflop/s-day (pfs-day) entails performing 1015 neural net operations per second for 1 day.
26. See https://blog.openai.com/ai-and-compute/.
27. See https://www.cbinsights.com/research/jobs-automation-artificial-intelligence-risk/.
28. Acemoglu and Restrepo (2018b, p. 2) describe AI as 'the currently most discussed automation technology'.
29. See for instance the International Labour Organization's (ILO) *Global Commission on the Future of Work*; the World Trade Report 2017 on *Trade, Technology and Jobs*, and the World Development Report 2019 on *The Changing Nature of Work*.
30. Huang and Rust (2018) argues that the belief that AI will not be likely or soon to automate jobs or tasks extensively using intuition or empathy, or jobs in the services sector, is mistaken, and that even such tasks will be automated.
31. This interpretation of what the novel 2001: A Space Odyssey is about is indebted to Robert J. Sawyer, see https://lifeboat.com/ex/ai.and.sci-fi. This description is virtually the entire plot of the novel. In 1968, Stanley Kubrick managed to make a 160-minute movie out of it that has been puzzling audiences ever since, see https://bit.ly/2GXVjFe.
32. Harari (2016, p. 314) points out that '*New technologies kill old gods and give birth to new gods*'.
33. See, e.g. O'Connell (2017).
34. This movement wants to create an AGI to rule as a new God over humanity (see https://bit.ly/2ybNqqQ)
35. As described by Vance (2010) the Singularity movement '*offers a modern-day, quasi-religious answer to the Fountain of Youth by affirming the notion that, yes indeed, humans – or at least something derived from them – can have it all*'.
36. Pinker does however differ significantly from Kurzweil about the possibility of a Singularity, declaring that it will 'never, ever' occur and more memorably that 'Sheer processing power is not a pixie dust that magically solves all your problems' (see Steven Pinker interviewed on this in 2008 at https://bit.ly/2U65JVW).
37. See page 11 of https://ainowinstitute.org/AI_Now_2018_Report.pdf.
38. According to *The Economist Magazine*, *Google* is leading the race: https://econ.st/2RcCldA
39. For instance, Google's *Deepmind* AI software filed on average one patent per week between September and December 2016 (https://bit.ly/2ATh5or).
40. See https://bit.ly/2tEQesD.
41. Consistent with this expectation Jia, Jin, and Wagman (2018) finds that the GDPR legislation in the EU were quick to impact negatively on the relative ability of EU firms to raise venture capital for technology investment.
42. See https://www.ft.com/content/e7bc0fd2-3149-11e9-8744-e7016697f225.
43. Interestingly, however, computer technology in general and AI specifically may improve the efficiency of R&D through facilitation of open innovation, in effect driving innovation in innovation, see, e.g Piller and Walcher (2006); Antonelli (2017).
44. One potential problem is of course, finding suitably skilled labor, such as data scientists and business analysts, to implement AI. A 2018 EY survey of executives in the USA found that 80 percent indicated a shortage of skilled labor as factor that constrains the implementation of AI (McKendrick 2018).
45. As WIPO (2019) data shows almost 50 percent of all AI patents are in the field of computer vision.
46. For example, the UK's 10-year average labor productivity growth since 2007 was the lowest since the 1761 (Lewis 2018).
47. See https://bit.ly/2uPZk6x.
48. See https://pewrsr.ch/2xmr8Qn/.

Acknowledgments

Disclosure statement

References

Acemoglu, D. 2002. "Technical Change, Inequality, and the Labor Market." *Journal of Economic Literature* 40 (1): 7–72.

Acemoglu, D., and D. Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, eds. D. Card and O. Ashenfelter, 1043–1171. Amsterdam: Elsevier.

Acemoglu, D., and D. Autor. 2012. "What Does Human Capital Do? A Review of Goldin and Katz's The Race Between Education and Technology." *Journal of Economic Literature* 50 (2): 426–463.

Acemoglu, D., and P. Restrepo. 2017. *Robots and Jobs: Evidence from US Labor Markets*. NBER Working Paper no. 23285. National Bureau for Economic Research.

Acemoglu, D., and P. Restrepo. 2018a. *Artificial Intelligence, Automation and Work*. Working Paper no. 24196. National Bureau of Economic Research.

Acemoglu, D., and P. Restrepo. 2018b. "The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares and Employment." *American Economic Review* 108 (6): 1488–1542.

Acemoglu, D., and P. Restrepo. 2020. "Robots and Jobs: Evidence From US Labor Markets." *Journal of Political Economy* 128 (6).

Aghion, P., B. Jones, and C. Jones. 2017. *Artificial Intelligence and Economic Growth*. NBER Working Paper 23928, National Burea for Economic Research.

Agrawal, A., J. Gans, and A. Goldfarb. 2019. "Economic Policy for Artificial Intelligence." In *Innovation Policy and the Economy*, Vol. 19, eds. J. Lerner and S. Stern, 139–159. Chicago: NBER.

Allen, P., and M. Greaves. 2011. *The Singularity isn't Near*. MIT Technology Review, October 12.

Andrews, D., C. Criscuolo, and P. Gal. 2016. *The Global Productivity Slowdown, Technology Divergence and Public Policy: A Firm Level Perspective*. OECD Productivity Working Papers, No. 5. Paris, OECD.

Antonelli, C. 2017. "Digital Knowledge Generation and the Appropriability Trade Off." *Telecommunications Policy* 41: 991–1002.

Apps, P. 2019. *Commentary: Are China, Russia Winning the AI Arms Race?* Reuters, 15 Jan. https://www.reuters.com/article/us-apps-ai-commentary/commentary-are-china-russia-winning-the-ai-arms-race-idUSKCN1P91NM).

Armstrong, S., N. Bostrom, and C. Schulman. 2016. "Racing to the Precipice: A Model of Artificial Intelligence Development." *AI & Society* 31: 201–206.

Arntz, M., T. Gregory, and U. Zierahn. 2016. *The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis*. OECD Social, Employment and Migration Working Paper no. 189. Paris: OECD.

Arntz, M., I. Gregory, and U. Zierahn. 2017. "Revisiting the Risk of Automation." *Economic Letters* 159: 157–160.

Atkinson, R., and J. Wu. 2017. "False Alarmism: Technological Disruption and the U.S. Labor Market, 1850–2015." *Information Technology and Innovation Foundation*, 1–28. May.

Auerswald, P. E. 2017. *The Code Economy: A Forty-Thousand-Year History*. Oxford: Oxford University Press.

Autor, D.. 2014. "Skills, Education, and the Rise of Earnings Inequality Among the Other 99 Percent." *Science (New York, N.Y.)* 344 (6186): 843–851.

Autor, D.. 2015. "Why are There Still so Many Jobs? The History and Future of Workplace Automation." *Journal of Economic Perspectives* 29 (3): 3–30.

Autor, D., and D. Dorn. 2013. "The Growth of Low Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103 (5): 1553–1597.

Autor, D., D. Dorn, L. Katz, C. Patterson, and J. V. Reenen. 2017. *The Fall of the Labour Share and the Rise of Superstar Firms*. Working Paper no. 23396. National Bureau of Economic Research.

Autor, D., L. Katz, and A. Krueger. 1999. "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics* 113: 1169–1214.

Autor, D., and A. Salomons. 2017. "Robocalypse Now: Does Productivity Growth Threaten Employment?" In *Proceedings of the ECB Forum on Central Banking: Investment and Growth in Advanced Economies*, 45–118. European Central Bank.

Autor, D., and A. Salomons. 2018. "Is Automation Labor-Displacing? Productivity Growth, Employment and the Labor Share." *Brookings Papers on Economic Activity, BPEA Conference*, Washington DC, 8-9 March.

Barrett, S. 2007. *Why Cooperate? The Incentive to Supply Global Public Goods*. Oxford: Oxford University Press.

Batra, G., A. Queirolo, and N. Santhanam. 2018. *Artificial Intelligence: The Time to Act is Now*. McKinsey, January.

Beaudry, P., D. Green, and B. Sand. 2013. *The Great Reversal in the Demand for Skill and Cognitive Tasks*. Working Paper No.18901. National Bureau of Economic Research.

Bellmund, J., P. Gaerdenfors, E. Moser, and C. Doeller. 2018. "Navigating Cognition: Spatial Codes for Human Thinking." *Science (New York, N.Y.)* 362 (6415): eaat6766.

Benzell, S., L. Kotlikoff, G. LaGardia, and J. Sachs. 2018. *Robots are Us: Some Economics of Human Replacement*. Working Paper 20941. National Bureau of Economic Research.

Berg, A., E. Buffie, and L. Zanna. 2018. *Robots, Growth, and Inequality: Should we Fear the Robot Revolution? (The Correct Answer is Yes)*. IMF Working Paper no. 18/116.

Bergstein, B. 2019. *This is Why AI has Yet to Reshape Most Businesses*. MIT Technology Review, 13 February.

Berriman, R., and J. Hawksworth. 2017. *Will Robots Steal our Jobs? The Potential Impact of Automation on the UK and other Major Economies*. PwC UK Economic Outlook, March.

Bessen, J., M. Goos, A. Salomons, and W. van den Berge. 2019. *Automatic Reaction – What Happens to Workers at Firms that Automate*? Law and Economics Series Paper no. 19-2, Boston University School of Law.

Bloom, D., M. McKenna, and K. Prettner. 2018. *Demography, Unemployment, Automation, and Digitalization: Implications for the Creation of (Decent) Jobs, 2010–2030*. IZA Discussion Paper no. 11739. IZA Institute of Labor Economics, Bonn.

Bongard, J., V. Zykov, and H. Lipson. 2006. "Resilient Machines Through Continuous Self-Modeling." *Science (New York, N.Y.)* 314 (5802): 1118–1121.

Bostrom, N. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.

Bostrom, N. 2017. *Strategic Implications of Openness in AI Development*. Global Policy, 1–14.

Bowles, J. 2017. *The Computerization of European Jobs*. Bruegel.org, 24th July.

Brooks, R. 2017. *The Seven Deadly Sins of AI Predictions*. MIT Technology Review, 6 October.

Brynjolfsson, E., and A. McAfee. 2012. "Thriving in the Automated Economy." *The Futurist* 46 (2): 27–31.

Brynjolfsson, E., and A. McAfee. 2015. "Will Humans Go the Way of Horses?" *Foreign Affairs* 94: 8–14.

Brynjolfsson, E., D. Rock, and C. Syverson. 2017. *Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics*. NBER Working Paper no. 24001. National Bureau for Economic Research.

Carbonero, F., E. Ernst, and E. Weber. 2018. *Robots Worldwide: The Impact of Automation on Employment and Trade*. Working Paper no. 36, ILO Research Department.

Card, D., and J. DiNardo. 2002. "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles." *Journal of Labor Economics* 20 (4): 733–783.

Chen, R. 2019. *Policy and Regulatory Issues with Digital Business*. Policy Research Working Paper no. 8948, Washington DC: The World Bank.

Chiacchio, F., G. Petropoulos, and D. Pichler. 2018. *The Impact of Industrial Robots on EU Employment and Wages: A Local Labor Market Approach*. Working Paper no. 02, Bruegel, Brussels.

Chung, T., M. Wedel, and R. Rust. 2016. "Adaptive Personalization Using Social Networks." *Journal of the Academy of Marketing* 44 (1): 66–87.

Cockburn, I., R. Henderson, and S. Stern. 2017. "The Impact of Artificial Intelligence on Innovation". *Paper presented at the NBER Conference on Research Issues in Artificial Intelligence, Toronto, September.*

Comin, D., and M. Mestieri. 2013. *If Technology has Arrived Everywhere, Why has Income Diverged*? NBER Working Paper No. 19010. National Bureau for Economic Research.

Cords, D., and K. Prettner. 2019. *Technological Unemployment Revisited: Automation in a Search and Matching Framework*. GLO Discussion Paper no. 308, Global Labor Organization.

Cowen, T. 2016. "Is Innovation Over? The Case Against Pessimism." *Foreign Affairs* 95 (2): 42–46.

Dauth, W., S. Findeisen, J. Suedekum, and N. Woessner. 2017. *German Robots – The Impact of Industrial Robots on Workers*. CEPR Discussion Paper No. 12306.

Deming, D. 2015. *The Growing Importance of Social Skills in the Labor Market*. Working Paper no. 21473. National Bureau of Economic Research.

Dyson, G. 2019. *The Third Law: The Future of Computing is Analog*. Medium. 12 February.

Dzyabura, D., and J. Hauser. 2011. "Active Machine Learning for Consideration Heuristics." *Marketing Science* 30 (5): 801–819.

EC. 2018. "Draft Ethics Guidelines for Trustworthy AI." *ECs High-Level Expert Group on Artificial Intelligence*. European Commission. Brussels.

Farboodi, M., R. Mihet, T. Philippon, and L. Veldkamp. 2019. *Big Data and Firm Dynamics*. Working Paper no. 25515, National Bureau of Economic Research.

Floridi, L. 2018. "The Ethics of Artificial Intelligence". In *Megatech: Technology in 2050*, ed. D. Franklin. Chapter 13, 155–163. London: Profile Books.

Ford, M. 2016. *The Rise of the Robots: Technology and the Threat of Mass Unemployment*. London: Oneworld Publications.

Ford, M. 2018. *Architects of Intelligence: The Truth About AI From the People Building It*. Birmingham: Pack Publishing.

Frey, C., and M. Osborne. 2013. "The Future of Employment: How Susceptible are Jobs to Computerization?" *Oxford Martin Programme on the Impacts of Future Technology*, University of Oxford.

Frey, C., and M. Osborne. 2017. "The Future of Employment: How Susceptible are Jobs to Computerization?" *Technological Forecasting and Social Change* 114: 254–280.

Friedman, T. 2016. *Thank You for Being Late: An Optimist's Guide to Thriving in the Age of Accelerations*. New York: Picador.

Furman, J., and R. Seamans. 2019. "AI and the Economy." In *Innovation Policy and the Economy*, Vol. 19, eds. J. Lerner and S. Stern, 161–191. Chicago: NBER.

Goldin, C., and L. Katz. 2010. *The Race Between Education and Technology*. Cambridge MA: Harvard University Press.

Goos, M., and A. Manning. 2007. "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain." *Review of Economics and Statistics* 89 (1): 118–133.

Gordon, R. 2018. *Why Has Economic Growth Slowed When Innovation Appears to be Accelerating*? NBER Working Paper 24554. National Bureau for Economic Research.

Graves, A., A. Mohamed, and S. Hinton. 2013. "Speech Recognition with Deep Recurrent Neural Networks." *EEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver BC, 26–31 May.

Greenemeier, L. 2017. *Ghost in the Sell: Hollywoods Mischievous Vision of AI*. Scientific American, 30 March.

Gregory, T., A. Salomons, and U. Zierahn. 2019. *Racing with or Against the Machine? Evidence from Europe*. IZA Discussion Paper no. 12063. IZA Institute of Labor Economics, Bonn.

Gries, T., and W. Naudé. 2020. *Artificial Intelligence, Income Distribution and Economic Growth*. IZA Discussion Paper no. 13606.

Hao, K. 2019. *We Analyzed 16,625 Papers to Figure Out Where AI is Headed Next*. MIT Technology Review, 25 January.

Harari, Y. 2011. *Sapiens: A Brief History of Humankind*. London: Vintage.

Harari, Y. 2016. *Homo Deus: A Brief History of Tomorrow*. London: Vintage.

Harris, S. 2015. *Can We Avoid a Digital Apocalypse*? SAM Harris Blog, 16 January, at https://samharris.org/can-we-avoid-a-digital-apocalypse/.

Hémous, D., and M. Olsen. 2018. *The Rise of the Machines: Automation, Horizontal Innovation and Income Inequality*. Mimeo, University of Zurich.

Hinton, G., S. Osindero, and T.-W. Teh. 2006. "A Fast Learning Algorithm for Deep Belief Nets." *Neural Computation* 18: 1527–1554.

Hinton, G., and R. Salakhutdinov. 2006. "Reducing the Dimensionality of Data with Neural Networks." *Science (New York, N.Y.)* 313: 504–507.

Hopenhayn, H., J. Neira, and R. Singhania. 2018. *From Population Growth to Firm Demographics: Implications from Concentration, Entrepreneurship and the Labor Share*. NBER Working Paper no. 25382.

House of Commons. 2019. *Disinformation and Fake News: Final Report*. Digital, Culture, Media and Sports Committee. UK Government, 18 February.

Huang, M.-H., and R. Rust. 2018. "Artificial Intelligence in Service." *Journal of Service Research* 21 (2): 155–172.

Jankel, N. 2015. *AI vs Human Intelligence: Why Computers Will Never Create Disruptive Innovations*. Huffington Post, 26 April.

Jia, J., G. Jin, and L. Wagman. 2018. *The Short-Run Effects of GDPR on Technology Venture Investment*. Working Paper no. 25248. National Bureau of Economic Research.

Jones, C. 1995. "R&D – Based Models of Economic Growth." *Journal of Political Economy* 103 (4): 759–783.

Jones, B. 2009. "The Burden of Knowledge and the Death of Renaissance Man: Is Innovation Getting Harder?" *Review of Economic Studies* 76 (1): 283–317.

Karabarbounis, L., and B. Neiman. 2014. "The Global Decline of the Labor Share." *Quarterly Journal of Economics* 129 (1): 61–103.

Koch, C. 2012. "Modular Biological Complexity." *Science (New York, N.Y.)* 337 (6094): 531–532.

Komlos, J. 2014. *Has Creative Destruction Become More Destructive*? Working Paper no. 20379. National Bureau of Economic Research.

Korinek, A., and J. Stiglitz. 2017. *Artificial Intelligence and Its Implications for Income Distribution and Unemployment*. NBER Working Paper no. 24174. National Bureau for Economic Research.

Kurzweil, R. 2005. *The Singularity is Near: When Humans Transcend Biology*. New York: Viking Press.

Landes, D. 1999. *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor*. New York: Norton & Co.

LeCun, Y., Y. Bengio, and G. Hinton. 2015. "Deep Learning." *Nature* 521: 436–444.

Lewis, J. 2018. *Bitesize: The Past Decade's Productivity Growth in Historical Context*. Bank Underground Blog, 25 April.

Liu, F., Y. Shi, and P. Li. 2017. "Analysis of the Relation Between Artical Intelligence and the Internet From the Perspective of Brain Science." *Procedia Computer Science* 122: 377–383.

Marcus, G. 2015. *Machines won't be Thinking Anytime Soon*. Edge.

McAfee, A., and E. Brynjolfsson. 2017. *Machine, Platform, Crowd: Harnessing Our Digital Future*. New York: W.W. Norton & Company.

McKendrick, J. 2018. *Artificial Intelligence Will Relieve Skills Shortages, If We Could Find Enough People to Build It*. Forbes, 16 May.

Meese, T. 2018. "The How and Why of Consciousness?" *Frontiers in Psychology* 9: 2173.

Minsky, M. 1961. *Steps Towards Artical Intelligence*. https://web.media.mit.edu/ minsky/papers/steps.html.

Misa, T. 2009. "History of Technology." In *A Companion to the Philosophy of Technology*, eds. J. K. B. Olsen, S. A. Pedersen and V. F. Hendricks, 7–17. Chichester: Wiley-Blackwell.

Mokyr, J. 2002. *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Princeton: Princeton University Press.

Mokyr, J. 2016. *A Culture of Growth: The Origins of the Modern Economy*. Princeton: Princeton University Press.

Moor, J. 2006. "The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years." *AI Magazine* 27 (4): 87–91.

Naudé, W. 2020 April. "Artificial Intelligence Vs. COVID-19: Limitations, Constraints and Pitfalls." *AI & Society* 35: 761–765. doi:10.1007/s00146-020-00978-0.

Naudé, W., and N. Dimitri. 2020. "The Race for An Artificial General Intelligence: Implications for Public Policy." *AI & Society* 35 (2): 367–379.

Naudé, W., and P. Nagler. 2018. *Technological Innovation, Entrepreneurship and Productivity in Germany, 1871–2015*. SPRU Working Paper Series (SWPS), 2018-02. Science Policy Research Unit, University of Sussex.

New Scientist. 2017. *Machines That Think*. London: John Murray Learning.

O'Connell, M. 2017. *To Be a Machine: Adventures Among Cyborgs, Utopians, Hackers, and the Futurists Solving the Modest Problem of Death*. New York: Doubleday Books.

OECD. 2015. "The Future of Productivity." OECD. Paris.

OECD. 2017. "Business Dynamics and Productivity."

Penrose, R. 1989. *The Emperors New Mind: Concerning Computers, Minds and the Laws of Physics*. Oxford: Oxford University Press.

Piller, F., and D. Walcher. 2006. "Tookits for Idea Competitions: A Novel Method to Integrate Users in New Product Development." *R/&D Management* 36 (3): 307–318.

Pinker, S. 2018. *Enlightenment Now: The Case for Reason, Science, Humanism, and Progress*. New York: Viking.

Pissarides, C. 2018. *Work in the Age of Robots and AI*. IZA 20th Anniversary Celebration, Berlin, 28–29 June.

Prettner, K., and H. Strulik. 2017. *The Lost Race Against the Machine: Automation, Education, and Inequality in an R&D-Based Growth Model*. Hohenheim Discussion Papers in Business, Economics and Social Sciences no. 08-2017.

Pulkka, V. 2017. "A Free Lunch with Robots Can a Basic Income Stabilise the Digital Economy?" *Transfer* 23 (3): 295–311.

Qu, J. 2018. *Training Self-Driving Cars Using Reinforcement Learning*. Medium, 28 October.

Ridley, M. 2011. *The Rational Optimist: How Prosperity Evolves*. London: Fourth Estate.

Rodriguez, F., and A. Jayadev. 2010. *The Declining Labor Share of Income*. Human Development Research Paper No. 2010/36, UNDP.

Russel, S., D. Dewey, and M. Tegmark. 2015. *Research Priorities for Robust and Beneficial Artificial Intelligence*. AI Magazine, Association for the Advancement of Artificial Intelligence, Winter, 105–114.

Sachs, J., S. Benzell, and G. LaGardia. 2015. *Robots: Curse or Blessing? A Basic Framework*. Working Paper 21091. National Bureau of Economic Research.

Schiller, R. 2019. *Narratives about Technology-Induced Job Degradations Then and Now*. Working Paper no. 25536. National Bureau of Economic Research.

Schwartz, R. 2018. *The Discourse is Unhinged: How the Media gets AI Alarmingly Wrong*. The Guardian, 25 July.

Schwartz, R., J. Dodge, N. Smith, and O. Etzioni. 2019. Green AI. Arxiv, https://arxiv.org/abs/1907.10597, 13 August.

Sequeira, T., P. Gil, and P. Afonso. 2018. "Endogenous Growth and Entropy." *Journal of Economic Behavior & Organization* 154 (C): 100–120.

Silver, D., T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis. 2017. "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm." arXix:1712.01815 [cs.AI].

Susskind, D. 2017. *A Model of Technological Unemployment*. Economic Series Working Papers 819. University of Oxford.

Tegmark, M. 2017. *Life 3.0: Being Human in the Age of Artificial Intelligence*. London: Penguin Books.

Tien, J. 2017. "Internet of Things, Real-Time Decision Making, and Articial Intelligence." *Annals of Data Science* 4 (2): 149–178.

Trajtenberg, M. 2018. *AI as the Next GPT: A Political-Economy Perspective*. NBER Working Paper no. 24245. National Bureau for Economic Research.

Turing, A. 1950. "Computing Machinery and Intelligence." *Mind; a Quarterly Review of Psychology and Philosophy* 59: 433–460.

Vance, A. 2010. *Merely Human? That's so Yesterday*. New York Times, 12 June.

Van de Gevel, A., and C. Noussair. 2013. "The Nexus between Artificial Intelligence and Economics". *Springer Briefs in Economics*.

Van Zanden, J., J. Baten, P. Foldvari, and B. van Leeuwen. 2014. "The Changing Shape of Global Inequality, 1820–2000: Exploring a New Dataset." *Review of Income and Wealth* 60 (2): 279–297.

Waldrop, M. 2016. "The Chips are Down for Moore's Law." *Nature* 530: 144–147.

WIPO. 2019. *Technology Trends 2019: Artificial Intelligence*. World Intellectual Property Association, Geneva.

World Bank. 2016. "World Development Report."

Yudkowsky, E. 2008. "Artificial Intelligence as a Positive and Negative Factor in Global Risk." In *Global Catastrophic Risks*, eds. N. Bostrom and M. N. Cirkovic, Chapter 15, 308–345. Oxford: Oxford University Press.

Zarkadis, G. 2018. *Forget Algorithms. The Future of AI is Hardware*! The Huffington Post, 14 January.