ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Naudé, Wim

Working Paper

The Race against the Robots and the Fallacy of the Giant Cheesecake: Immediate and Imagined Impacts of Artificial Intelligence

IZA Discussion Papers, No. 12218

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Naudé, Wim (2019) : The Race against the Robots and the Fallacy of the Giant Cheesecake: Immediate and Imagined Impacts of Artificial Intelligence, IZA Discussion Papers, No. 12218, Institute of Labor Economics (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/196716

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12218

The Race against the Robots and the Fallacy of the Giant Cheesecake: Immediate and Imagined Impacts of Artificial Intelligence

Wim Naudé

MARCH 2019

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12218

The Race against the Robots and the Fallacy of the Giant Cheesecake: Immediate and Imagined Impacts of Artificial Intelligence

Wim Naudé
Maastricht University, UNU-MERIT, MSM, RWTH Aachen University and IZA

MARCH 2019

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Race against the Robots and the Fallacy of the Giant Cheesecake: Immediate and Imagined Impacts of Artificial Intelligence

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' are imminent. In part, this is because current AI, based on deep learning, is expensive and difficult for (especially small) businesses to adopt, can create new jobs, and is an unlikely route to the invention of a super-intelligence. Even though AI is unlikely to have either utopian or apocalyptic impacts, it will challenge economists in coming years. The challenges include regulation of data and algorithms; the (mis-) measurement of value added; market failures, anti-competitive behaviour and abuse of market power; surveillance, censorship, cybercrime; labor market discrimination, declining job quality; and AI in emerging economies.

JEL Classification: O47, O33, J24, E21, E25

Keywords: technology, articial intelligence, productivity, labor demand, innovation, inequality

Corresponding author:
Wim Naudé
Maastricht School of Management (MSM)
PO Box 1203
6201 BE Maastricht
The Netherlands

E-mail: w.naude@maastrichtuniversity.nl

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] (Auerswald, 2017; Landes, 1999; Mokyr, 2002, 2016; Pinker, 2018; Ridley, 2011). These include inventions and discoveries such as the domestication 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 akin to electricity[2] : Artificial Intelligence (AI) (Agrawal et al., 2019; Brynjolfsson et al., 2017; Trajtenberg, 2018).

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 employment, job quality, inequality, competition and collusive behavior, amongst others (Aghion et al., 2017; Agrawal et al., 2019; Berg et al., 2018; Bloom et al., 2018; Furman and Seamans, 2019; Prettner and Strulik, 2017; Schiller, 2019). According to Acemoglu and Restrepo (2018b, p.2) AI has become the '*most discussed automation technology*'.

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 (Bostrom, 2014; Yudkowsky, 2008) with which economists have not yet engaged in much. For instance, where will continued innovation in AI will ultimately lead to? To 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 shows that AI research, investment and applications are booming after a number of 'AI winters'. Currently most efforts in AI involve computer vision, natural language processing and speech, and the design of algorithms to improve business processes and steer consumer spending. It is argues 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 spiralling 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 tend to be very sensitive to the assumptions used; (ii) automation,[3] being rather domain-specific, affects the tasks that humans perform, and not necessarily their jobs; (iii) net job creation can

[1] Although philosophers of technology and historians of technology have typically differed on the extent of technological determinism (see e.g. Misa (2009)) there is little doubt amongst scholars and policy makers as to the integral role and 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 US$ 21 billion R&D budget is on AI, he would say this.

[3] 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.

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.

Even though the medium- and longer-term impacts of AI are unlikely to be either utopian or apocalyptic, it will have a number of implications that will challenge economists in coming years. These include regulation of data and algorithms; the (mis-) measurement of value added; market failures, anti-competitive behaviour and abuse of market power; surveillance, censorship, cybercrime; labor market discrimination, declining job quality; and AI in emerging economies.

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 more 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*[4], 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[5]. Most recently however, the interest in AI can be traced back to the 1st Industrial Revolution and its aftermath. Two advances were relevant: first, the engineering achievements of the 18th and early 19th 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 19th century inspired many works of literary fiction about intelligent machines. Much of this is dystopian. For instance, in 1909 E.M. Forster wrote a novel *The Machine Stops* wherein a post-apocalyptic society is controlled by a super-intelligent machine - and the humans who are left are not too happy about it[6] . In his 1965 novel *Dune* Frank Herbert tells the story of a society in which intelligence machines are so dangerous as to be banned[7] by the commandment '*Thou shalt not make a machine in the*

[4] See: https://bit.ly/2nNL0Ib
[5] See: http://history-computer.com/Dreamers/LeonardoAutomata.html.
[6] Did E.M. Forster anticipate the internet and smart phone in 1909? See the discussion in https://www.bbc.com/news/entertainment-arts-36289890.
[7] See: https://bit.ly/2NxxjJ1.

likeness of a human mind' (no need for concern: this commandment has not yet been broken). And 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. The unhappy fictional relationship between human and machine intelligence continues right up to the present: the 2017 HBO TV series *Westworld* likewise concerns the uprising of robots against their misuse by humans.

The term 'artificial intelligence' (AI) was adopted by the scientific community in 1956 at a conference held at Dartmouth College in the USA (Moor, 2006). The 1950s heralded the modern agenda to create machine intelligence. Between 1956 and 2007 however, research and business interest in AI waxed and waned. AI winters, during which interest and funding in AI declined, occurred in the 1970s and late 1980s up until around 2006[8] (Liu et al., 2017). Largely these AI winters were due to the failure of the traditional approach to simulate human intelligence, namely Symbolic Reasoning (New Scientist, 2017). From the late 1980s, cognitive scientists and AI researchers have increasingly turned to non-symbolist approaches, inspired by Connectivism, or Artificial Neural Networks (ANN), to simulate aspects of human intelligence, in particular perception, e.g. Graves et al. (2013). This is reflected, as will be shown below, in the concentration of contemporary AI research in fields like computer vision and speech recognition.

At this point the question arises what intelligence and specifically *artificial* intelligence is?

Intelligence is a hard-to-define concept. As Tegmark (2017, p.49) puts it, '*There's no agreement on what intelligence is even among intelligent intelligence researchers*'. There is hence no common definition of AI (Van de Gevel and Noussair, 2013). The definition proposed by New Scientist (2017, p.3) is fairly representative and consistent with the earliest views on intelligent 'machines', stating that AI refer to '*Machines that act intelligently*' and can '*make the right decision in uncertain circumstances*'.

Around 2012[9] 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*[10] 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.[11] These confluent ICT developments are discussed by among others Brynjolfsson et al. (2017); Friedman (2016); McAfee and Brynjolfsson (2017) and Tegmark (2017). Engaging historical and futurist

[8] According to Liu et al. (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) and Hinton et al. (2006).

[9] 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 et al. (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)).

[10] 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).

[11] It is estimated that around 2,5 quintillion gigabytes of data were created daily in 2017 (see `https://bit.ly/2fl2JCT`)

perspectives on the rise of humanity and AI are contained in Harari (2011, 2016).

The simultaneous improvements in tools to handle large datasets, such as machine learning[12], and specifically the sub-set of artificial neural networks, or deep learning[13] 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 that constituted traditional AI in the second half of the 20th 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 'narrow' AI. Whether and when an AGI will be invented are open questions: some see it as unlikely or very far off (e.g. Marcus (2015); Koch (2012); Allen and Greaves (2011)), while a 2017 survey[14] 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; 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 it 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. As Brooks (2017) emphasizes, '*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 fact, in 2017 *The Economist Magazine*[15] 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*[16].

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*

[12] 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).

[13] '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 et al., 2015, p.436). See also **?**.

[14] See `https://blog.aimultiple.com/artificial-general-intelligence-singularity-timing/`

[15] See: `https://econ.st/2Gtfztg`

[16] See : `https://en.wikipedia.org/wiki/FacebookCambridge_Analytica_data_scandal`

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.[17] 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[18]. 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 (e.g. Farboodi et al. (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', 'hear' and communicate. Researchers are also attempting to learn robots to be self-aware, through for instance continuous self-modelling (see e.g. Bongard et al. (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. 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[19] to personal relationships. [20]

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 datasets 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 1,000 petaflop/s[21] a day. This is 300,000 times more compute thrown at deep learning in 2018 than in 2012.[22]

What this reflects is 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 cornered niche.

[17] See : https://eugdpr.org
[18] See:https://nyti.ms/2rXdxMJ
[19] Want to know more about the Artificial Neural Beer Network (ANBN)? : https://bit.ly/2GZQyK9
[20] Will your relationship with your better half last? Check : https://bbc.in/2HgFQRk
[21] A petaflop/s-day (pfs-day) entails performing 1015 neural net operations per second for one day.
[22] See: https://blog.openai.com/ai-and-compute/

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[23]

The immediate impact of AI, as an automation technology[24], 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 Brynjolfsson and McAfee (2012) book *Race Against the Machine* or the *Rise of the Robots* by Ford (2016). Warnings have been sounded that the robots may win this race with job losses and rising inequality as result. Autor and Salomons (2017) refers 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 redistributive tax policies and expansion of education, that AI will through raising the skills premium, widening the wage gap, and decreasing the share of labor in GDP, lead to higher income inequality, e.g. Autor (2014); Berg et al. (2018); Korinek and Stiglitz (2017).

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 spiralling 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, and through the different wage implications, lead to changes in the distribution of wages. During the 1st 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 Acemoglu (2002); Autor et al. (1999); Card and DiNardo (2002); Goldin and Katz (2010).

[23] See : https://www.cbinsights.com/research/jobs-automation-artificial-intelligence-risk/
[24] Acemoglu and Restrepo (2018b, p.2) describe AI as 'the currently most discussed automation technology'.

3.1.2 Impact on jobs

The possibility that skill-biased technological change can lead to higher unemployment and inequality was emphasised in the years immediately after the global financial crisis by Brynjolfsson and McAfee (2012), Brynjolfsson and McAfee (2015) and Ford (2016) and by the publication in 2013 of very specific predictions of how much jobs automation will replace. 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 to 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 automaton 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 1,000 workers will reduce the employment: population ratio in the USA by 0,37 percent. Using a similar method Chiacchio et al. (2018) calculated that one additional robot per 1,000 workers can reduce EU employment by between 0,16 to 0,20 percent. Carbonero et al. (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 ten year period was less than 0,5 percent (Carbonero et al., 2018).

While many robots use AI, many do not; robotics is an automation technology that does not need to rely on AI and the extent to which AI has in fact 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 (implicitly) believed that AI is a particularly effective automation technology and that it can be quickly diffused. These beliefs are reflected in the arguments that AI is a general-purpose technology and hence as disruptive as 'electricity' or 'fire' (Brynjolfsson et al., 2017) and that it is progressing at an exponential rate. This means, as Chiacchio et al. (2018, p.3) reports from a McKinsey study, that the belief is that the AI fuelled disruption could be ten times faster and 300 times the scale than during the 1st Industrial Revolution, thus having '*3,000 times the impact*'.

In section 4 below 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[25] and many national governments have been highlighting these concerns in recent years, in particular under the banner of *The Future of Work* (Gries and Naudé, 2018). 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 with more routine-tasks, that human labour invest in skills that are less likely to be able to be automated by AI 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

[25] See for instance the International Labour Organizations (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*.

of jobs (Schiller, 2019). Bloom et al. (2018, p.28) note that automation and digitalization has 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 should be understood to be dependent 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' (Acemoglu and Restrepo, 2018b; Autor and Salomons, 2018). Below reference will be made to the result of overlapping generations (OLG) models by Sachs et al. (2015) and Benzell et al. (2018), and a demand-constrained endogenous growth model proposed by Gries and Naudé (2018) that departs from the 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 below.

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), Acemoglu and Autor (2012), and also Autor and Dorn (2013); Beaudry et al. (2013); Goos and Manning (2007). 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 (2018), 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 et al. (2018) simulates 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 per cent to 8 - 18 percent*' (Berg et al., 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 (Karabarbounis and Neiman, 2014; Rodriguez and Jayadev, 2010; 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) argues 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 et al. (2018), using a general equilibrium firm dynamics model which they calibrate to USA data, finds 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) and 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 do however, 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 et al. (2018) who finds 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 modelling 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 modelled either using overlapping generations (OLGs) frameworks (e.g. Benzell et al. (2018); Sachs et al. (2015)), or R&D -based endogenous growth models. Within the latter, they 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 (2018a), Acemoglu and Restrepo (2018b) or using a variety of intermediate inputs (e.g. Hémous and Olsen (2018) and Gries and Naudé (2018)).

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, 2018a). Cords and Prettner (2019) introduces 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 modelled 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 et al., 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 (2018) who proposes 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 that technological innovations lead to rising automation, a rise in the skills premium (SBTC), a declining labor share, higher unemployment, and also higher economic growth. Their models joint outcome of higher economic growth and inequality is however at odds with empirical observations and with Piketty (2013) who links slower growth to rising inequality.

Acemoglu and Restrepo (2018b) similarly provides a task-based model wherein automation is an expansion of the set of tasks that can be produced with capital. This leads to some tasks traditionally performed by labor to be done by capital and leads 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) and Sachs et al. (2015) use an overlapping generations (OLG) model and Gries and Naudé (2018) an endogenous growth model to explore how the distributive impacts of automation may lead via the demand side to growth constraints. Sachs et al. (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 increases in economic growth and inequality may be occurring simultaneously, 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 et al. (2015) and Benzell et al. (2018) calls 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 (2018) provide an endogenous growth model that is complimentary to the central concerns in Benzell et al. (2018) and Sachs et al. (2015). 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. Their results suggest, in line with the recommendations of Benzell et al. (2018) and Sachs et al. (2015) 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 summarising five reasons why, from theoretical and empirical considerations, AI will not lead to mass unemployment and spiralling inequality.

The first reason is that the predictions of mass unemployment are based on 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) and Frey and Osborne (2017) are sensitive to these assumptions. Arntz et al. (2016) and Arntz et al. (2017) for instance using different assumptions finds that actually only 9 percent of OECD jobs are possibly subject to automation in the next decade.

Secondly, automation may affect tasks, rather than 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 refers to the effect of new technology raising labour productivity and wages and thus aggregate demand. Autor and Salomons (2018, pp.12-13) illustrates 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 declining prices for basic goods which then increases available income to be spend on other goods; and a *Costco* effect whereby some processing sectors increase their efficiency due to automation and this is passed on to downstream sectors using this as an input.

Recent 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, Cords and Prettner (2019) and Dauth et al. (2017) for Germany, and Gregory et al. (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 concludes as well 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, diffusion of AI may be much slower than thought previously. It is costly and complex for business firms to utilize AI. This may explain the results discussed in point (iii). Studies by Andrews et al. (2016), Bergstein (2019), Brynjolfsson et al. (2017) and OECD (2015) 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 et al. (2017) argue that implementation lags mean that the real impact of AI on productivity (and wages) is yet to be seen. More generally, authors 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 much AI 'innovations' have been characterised 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 rising 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 the 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.[26]

The ways of humans and AI may split in future. A superintelligence 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 superintelligence 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).

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*: we may want giant cheesecakes, and the super-intelligence may be able to give it to us, but we 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[27] that are inspired by new technologies, such as *Transhumanism,*[28] the *Way of the Future* [29] and the *Singularity*[30] movement.

[26] 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`.

[27] Harari (2016, p.314) points out that '*New technologies kill old gods and give birth to new gods*'.

[28] see e.g. O'Connell (2017).

[29] This movement wants to create an AGI to rule as a new God over humanity (see `https://bit.ly/2ybNqqQ`)

[30] 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*

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, p5). 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. a 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: '*Its 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[31] 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 Bostrom (2017) and Yudkowsky (2008) 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 et al. (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 and competing for investing 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*'.

- *can have it all*'.

[31] Pinker does however differ significantly from Kurzweil about the possibility of a Singularity, stating 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 Pinker interviewed on this in 2008 at `https://bit.ly/2U65JVW`).

One problem with resulting 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 et al., 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 et al., 2016; Naudé and Dimitri, 2018). The *AI Now 2018 Report*[32] noted that AI arms races creates 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 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*[33] .

Armstrong et al. (2016) and Naudé and Dimitri (2018) models 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 et al. (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.[34]

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

[32] See page 11 of https://ainowinstitute.org/AI_Now_2018_Report.pdf
[33] According to *The Economist Magazine*, *Google* is leading the race: https://econ.st/2RcCIdA
[34] For instance, Googles *Deepmind* AI software filed on average one patent per week between September and December 2016 (https://bit.ly/2ATh5or).

the 'animal spirits' that characterizes 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) (remember, introduced at an over-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[35] 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[36] 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.[37]

5.2 The singularity can be cancelled

A second point (which is related to the first) is that an AGI still seems 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*

[35] See https://bit.ly/2tEQesD

[36] Consistent with this expectation Jia et al. (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.

[37] See: https://www.ft.com/content/e7bc0fd2-3149-11e9-8744-e7016697f225

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'.

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 (2018) 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 is non-trivial. To be precise, current AI R&D and business generation is difficult and expensive (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 performs the vast bulk of research, patenting, as well as receives the bulk (more than 90 percent) of venture capital funding for AI (WIPO, 2019).

This indicates that 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).[38] 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.

5.4 The limits of deep learning

A fourth point that is relevant for the present discussion, and which is somewhat related to the previous point, as well as 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[39]) is that possibility that the current AI-boom may not last very long. Perhaps another AI winter is looming. Deep learning is facing decreasing returns, amongst others AI may need innovations in hardware (including in edge and quantum computing) and the same may apply for the growth in applications based on it (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 is getting harder as a result of the burden of knowledge (Jones, 2009). Sequeira et al. (2018) has shown that entropy will limit the extent of technologically-driven economic growth and provides estimates of when economic growth will converge to a zero rate. Similarly, Moore's Law is not expected to hold for much longer

[38] One perennial problem is of course, finding skilled labor, e.g. data scientists, 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).

[39] As WIPO (2019) data shows, almost 50 percent of all AI patents are in the field of computer vision.

(Waldrop, 2016).

5.5 Message from the empirical evidence...so far

A fifth point is that the impact of AI, over the short and medium term, on job losses, inequality and productivity has been overestimated. Early predictions were that AI would lead to massive job losses, large increases in labour productivity, and spiralling inequality. In the meantime, the experience of much advanced economies has been the opposite. Unemployment is at historical lows and productivity growth is stagnating;[40] 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 labour 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 (Gries and Naudé, 2018; Sequeira et al., 2018; Schiller, 2019).

The empirical evidence is however subject to two shortcomings. First, most studies have focused on advanced economies. They few studies from emerging economies (e.g. World Bank (2016); Carbonero et al. (2018) tend to suggest that the effects of automation may be worse for emerging economies. 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 and existing automation technologies, and therefore one may see AI fas synonymous for automation. 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 machine learning analytics outside the giant tech companies, then the exaggeration or hype in the media about AI becomes even more accentuated.

[40] For example, the UK's ten-year average labor productivity growth since 2007 was the lowest since the 1761 (Lewis, 2018).

Schwartz (2018) describes the current media discourse about AI as '*unhinged*' and quote Lipton that the world is suffering an '*AI misinformation epidemic'*.[41] Despite the rigorous economic studies discussed in section 3 of this paper finding little evidence, and prospects, for mass replacement of jobs by automation, a 2018 Poll by the Pew Research Center[42] found that 65 percent of Americans still believe 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 occurs. Perhaps the two most relevant for the present discussion is 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 fact that different computer hardware may be 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 will need to take place in edge computing. As Batra et al. (2018) point out, '*inference at the edge will become increasingly common for applications where latency in the order of microseconds is mission critical*'.

6 Concluding Remarks

'*Hollywoods vision of AI is often entertaining, generally pessimistic and rarely realistic*' - Greenemeier (2017).

While humans' fascination with intelligent machines dates back centuries, the most recent infatuation can be traced back to the introduction of the term artificial intelligence (AI) in 1956 and 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. In particular, machine (deep) learning lead to the development of computer vision, natural language processing, speech and business intelligence software. These four areas of machine 'talents' have since permeated through many sectors and resulted in myriads of applications: most notably in experimental autonomous vehicles, but also in medical diagnostics, marketing and advertisement, entertainment and games, financial technologies, and others.

[41] See: https://bit.ly/2uPZk6x
[42] See : https://pewrsr.ch/2xmr8Qn/

The progress in AI can be traced through the sharp increase, first in scientific publications on the topic (starting around 2002), secondly in patent publications, especially from 2012, and thirdly in venture capital to AI-start-up companies, especially from 2016.

Also, around 2013, economists started to ask what the implications of this AI-boom could be in terms of its impact on labor markets (productivity, jobs, inequality) and competitiveness of markets (market concentration and collusion). At first, there were much concern that AI will cause huge job losses, giving rise to fears that human labor is in a race against the robots. Some however, falling into a *Fallacy of the Giant Cheesecake*, considered the present AI-boom as the harbinger of a tech-utopia, the Singularity, which will be achieved when AI advanced to the state that it becomes a super-intelligence (or AGI).

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 the boom in AI is driving significant business process efficiency and generating consumer surpluses, and automating many tasks performed by humans, it has so far not led to mass technological unemployment. AI is also not a main driver of income inequality (although the labor share of GDP may decline as a more general tendency of capital to substitute for labor much more easily over time, and emerging economies may be more negatively affected than more advanced economies).

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 (and complementary investments and innovations are needed for this, such as in edge computing or quantum computing). Some are predicting another AI winter, others are relatedly arguing that deep learning could soon be exhausted. Perhaps the most important feature of AI, although neglected in the media hype, is its slow and limited diffusion outside the big tech companies, internet giants and few leading countries - as a result of requiring expensive investment, large volumes of reliable data, time-intensive commitments, changes in organizational structures, as well as sophisticated data expertise.

As was noted, the experience of much advanced economies has been the opposite to what was at first predicted. Instead of high unemployment and high productivity growth, unemployment is at historical lows and productivity growth is stagnating; global inequality is also down, and where within-country inequality has increased, for example in Germany, this has been found to be due to too slow, rather than too fast technological innovation. More research however is needed on the distributional impacts of AI between countries. Given 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 the potential vulnerabilities of emerging economies to AI.

Theoretical models used by economists have also evolved to deal better with the nature of an AI technology that can both complement and substitute for labor, and these have variously shown that automation can affect tasks within jobs rather than jobs itself, can stimulate an increase in labor demand through for the re-institution of labor due to increases in labor productivity

and wages, and moreover that AI progress can be consistent with low wage and productivity growth, relatively high employment, and economic stagnation.

As far as the long-term impacts of AI is concerned, this paper has concluded that a *Singularity* or AGI is only a remote possibility. The term artificial intelligence is in fact a misnomer and current research into deep learning will probably not lead to invention of an AGI. Both a blissful Singularity and a robocalypse seems at present to belong only to the realms of science fiction. Of course, considering *Amara's Law*, this paper may be grossly underestimating the longer-run impacts of technology.

In conclusion: even though AI is unlikely to have either utopian or apocalyptic impacts, it will have a number of implications that will challenge economists in coming years. These include regulation of data and algorithms; the (mis-) measurement of value added; market failures, anti-competitive behaviour and abuse of market power; surveillance, censorship, cybercrime; labor market discrimination, declining job quality; and AI in emerging economies.

Acknowledgements

I am grateful to participants at conferences, workshops and seminars at RWTH Aachen University, Sussex University (*15th European Network on the Economics of the Firm meeting*) and Utrecht University for their helpful comments and questions on earlier and related versions of this work. In particular I wish to thank Seth Benzell, Niels Bosma, Tommaso Ciarli, Nicola Dimitri, Maarten Goos, Pedro Mazeda Gil, Terry Gregory, Thomas Gries, Thomas Kittsteiner, Marion Ott, Frank Piller, Anna Salomons, Erik Stam and Ed Steinmueller for their interest and suggestions. I take sole responsibility for errors and omissions.

References

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

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment earnings. *In Ashenfelter, O. and Card, D. eds. The Handbook of Labor Economics vol 4B. Amsterdam: Elsevier.*

Acemoglu, D. and Autor, D. (2012). What does Human Capital Do? A Review of Goldin and Katzs The Race between Education and Technology. *Journal of Economic Literature*, 50(2):426–463.

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

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

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

Aghion, P., Jones, B., and Jones, C. (2017). Artificial Intelligence and Economic Growth. *Working Paper no. 23928. National Bureau of Economic Research.*

Agrawal, A., Gans, J., and Goldfarb, A. (2019). Economic Policy for Artificial Intelligence. *(In Lerner, J. and Stern, S. eds. Innovation Policy and the Economy, Vol. 19. NBER. Pp. 139-159).*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Berg, A., Buffie, E., and Zanna, L. (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 Hawksworth, J. (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, Salomons, A., and den Berge, W. V. (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., McKenna, M., 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., Zykov, V., and Lipson, H. (2006). Resilient Machines Through Continuous Self-Modeling. *Science*, 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*, pages 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 McAfee, A. (2012). Thriving in the Automated Economy. *The Futurist*, March-April:27–31.

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

Brynjolfsson, E., Rock, D., and Syverson, C. (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., Ernst, E., and Weber, E. (2018). Robots Worldwide: The Impact of Automation on Employment and Trade. *Working Paper no. 36, ILO Research Department.*

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

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

Cockburn, I., Henderson, R., and Stern, S. (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 Mestieri, M. (2013). If technology has arrived everywhere, why has income diverged? *NBER Working Paper No. 19010. National Bureau for Economic Research.*

Cords, D. and Prettner, K. (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*, March/April:42–46.

Dauth, W., S.Findeisen, J.Suedekum, and Woessner, N. (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.*

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

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

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

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 Osborne, M. (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 Osborne, M. (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 Optimists Guide to Thriving in the Age of Accelerations. *New York: Picador.*

Furman, J. and Seamans, R. (2019). AI and the Economy. *In Lerner, J. and Stern, S. eds. Innovation Policy and the Economy, Vol. 19. NBER. Pp. 161-191.*

Goldin, C. and Katz, L. (2010). The Race between Education and Technology. *Harvard University Press.*

Goos, M. and Manning, A. (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., Mohamed, A., and Hinton, S. (2013). Speech Recognition with Deep Recurrent Neural Networks. *IEEE International Conference on Acoustics, Speech and Signal Processing, 26 31 May.*

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

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

Gries, T. and Naudé, W. (2018). Artificial Intelligence, Jobs, Inequality and Productivity: Does Aggregate Demand Matter? *IZA Discussion Paper no. 12005. IZA Institute of Labor Economics, Bonn.*

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

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

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

Hopenhayn, H., J.Neira, and Singhania, R. (2018). From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share. *Working Paper no. 25382, National Bureau of Economic Research.*

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

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

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

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.

Jones, C. (1995). R&d-based models of Economic Growth. *Journal of Political Economy*, 103(4):759–783.

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

Koch, C. (2012). Modular Biological Complexity. *Science*, 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 Stiglitz, J. (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., Bengio, Y., and Hinton, G. (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., Shi, Y., and Li, P. (2017). Analysis of the Relation Between Artificial Intelligence and the Internet from the Perspective of Brain Science. *Procedia Computer Science*, pages 377–383.

Marcus, G. (2015). Machines wont be Thinking Anytime Soon. *Edge.*

McAfee, A. and Brynjolfsson, E. (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*, Volume 9, article 2173.

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

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. and Dimitri, N. (2018). The Race for an Artificial General Intelligence: Implications for Public Policy. *IZA Discussion Paper no. 11737. IZA Institute of Labor Economics, Bonn.*

Naudé, W. and Nagler, P. (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). Algorithms and Collusion: Competition Policy in the Digital Age. *Paris: Organization for Economic Cooperation and Development.*

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

Piketty, T. (2013). Capital in the 21st Century. *Harvard University Press.*

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 Strulik, H. (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.-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 Oct.*

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

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

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

Sachs, J., Benzell, S., and LaGardia, G. (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.*

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

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., and Hassabis, D. (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 Artificial 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.*

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

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

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

Waldrop, M. (2016). The Chips are Down for Moore's Law. *Nature*, 530:144147.

WIPO (2019). Technology Trends 2019: Artificial Intelligence. *Geneva: World Intellectual Property Organization.*

World Bank (2016). World Development Report 2016.

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

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