Artificial intelligence: increasing labour productivity in a responsible way[1]

24 November 2022

Georgios Petropoulos[2] and Mamta Kapur[3]

TRANSATLANTIC EXPERT GROUP
ON THE FUTURE OF WORK

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

[1] We are very thankful to the valuable comments and suggestions by the other members of the Transatlantic Expert Group on the Future of Work on earlier drafts. A part of this paper has been used as Georgios Petropoulos' expert contribution to the Global Innovation Index 2022 of the World Intellectual Property Organization.

[2] Researcher at Massachusetts Institute of Technology, Stanford University and Bruegel.
[3] Europe Research lead, Talent and Organisation, Accenture.

Co-funded by the European Union

This publication was produced with the financial support of the European Union. Its contents are the sole responsibility of Bruegel AISBL and The German Marshall Fund of the United States and do not necessarily reflect the views of the European Union

Abstract

In the last decade, labour productivity growth has slowed down despite the fast development of new efficient general-purpose digital technologies, including machine learning and artificial intelligence. This productivity slowdown is a big paradox. Does this mean there were false hopes about the potential of these technologies to transform societies and improve people's lives? This paper has two objectives. First, it provides and evaluates alternative explanations for this paradox and proposes specific policy recommendations in order to resolve it and increase productivity. Second, as these recommendations point towards the larger-scale adoption and diffusion of artificial intelligence technologies, it provides a framework that ensures that the ways AI systems are designed, built and scaled up are ethical and responsible by design.

Contents

1. The productivity paradox

Following the Second World War, the United States, Soviet Union and Western European and East Asian countries witnessed phenomenal productivity growth and high employment rates. This lasted until the 1970s, when economic recession occurred in the Western world, with the high inflation and unemployment. In the 1970s a new industrial revolution took place, based on personal computers and information and communication technologies (ICT). It signalled the beginning of a new era in which the rise and systematic adoption of electronics, telecommunications and computers exerted disruptive forces on business and economic models, providing more efficiency gains in the production process and new possibilities and modes of communications. Many experts call this period the fourth industrial revolution (Schwab, 2017, Perez, 2002). The continuous advances in ICT technologies have led to a digital new reality in which new sectors, products and services have been developed in a process of rapid digitization of the world economy, and in which high-level automation has become a popular industrial and business practice.

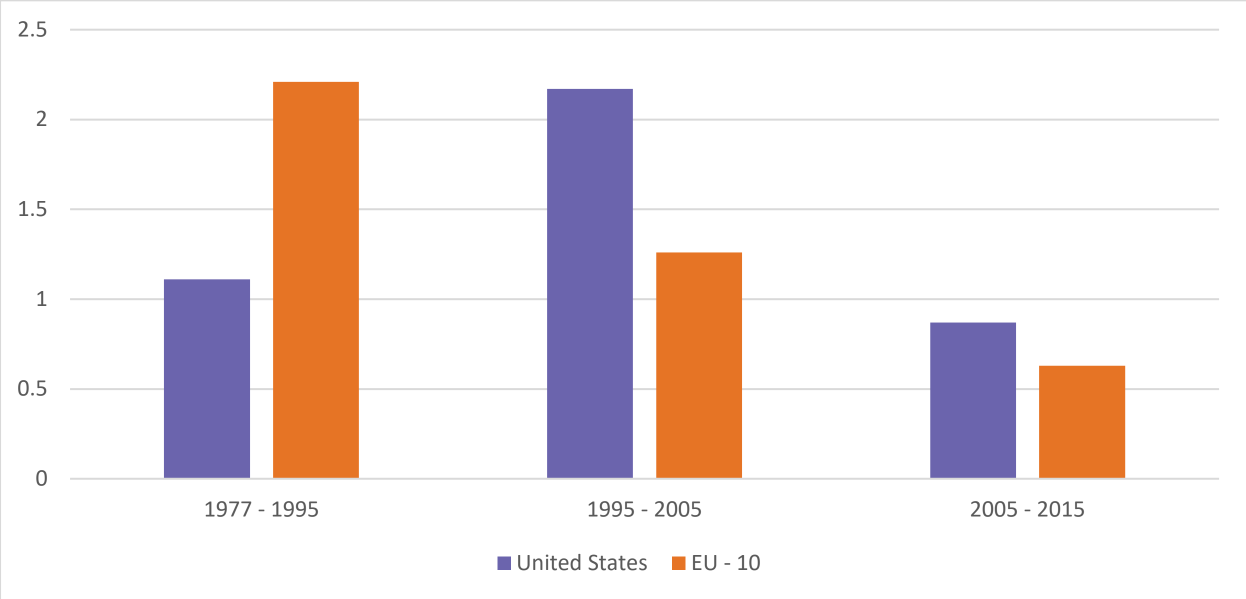
Despite the large scale of ICT deployment in the 1970s and 1980s, substantial benefits in productivity statistics have not been evident, as reflected by the famous quote by Nobel Laurate Robert Solow: *"You can see the computer age everywhere but in the productivity statistics"*[4]. In fact, some productivity growth benefits have been observed, but they took some time to arrive, were US-specific and did not last for long. The US economy experienced a productivity growth revival between 1995 and 2004, which was not observed in other major economies, such as European countries (Figure 1).

[4] 'We'd better watch out', *New York Times Book Review*, 12 July 1987, page 36.

Co-funded by the European Union

This publication was produced with the financial support of the European Union. Its contents are the sole responsibility of Bruegel AISBL and The German Marshall Fund of the United States and do not necessarily reflect the views of the European Union

Figure 1: Labour productivity growth rates for US and EU10[5]



Source: Gordon and Sayed (2020).

Gordon and Sayed (2020) found that US productivity revival was driven by intense investments in ICT, which led to additional productivity growth in: i) ICT-intensive service producing industries, and ii) the electric machinery industry that produces computer hardware. In contrast, European countries did not invest heavily in computer hardware, and ICT-producing industries have been of much less importance (O'Mahony *et al*, 2008). As a result, Europe failed to reap the productivity benefits of the ICT revolution.

Within those years of substantial productivity growth, US ICT-intensive industries experienced a productivity growth revival of roughly 2 percentage points, compared to non-ICT intensive industries in which productivity growth remained at the same low levels as before (Stiroth, 2002). In addition to heavy ICT investments, structural market changes, such as the

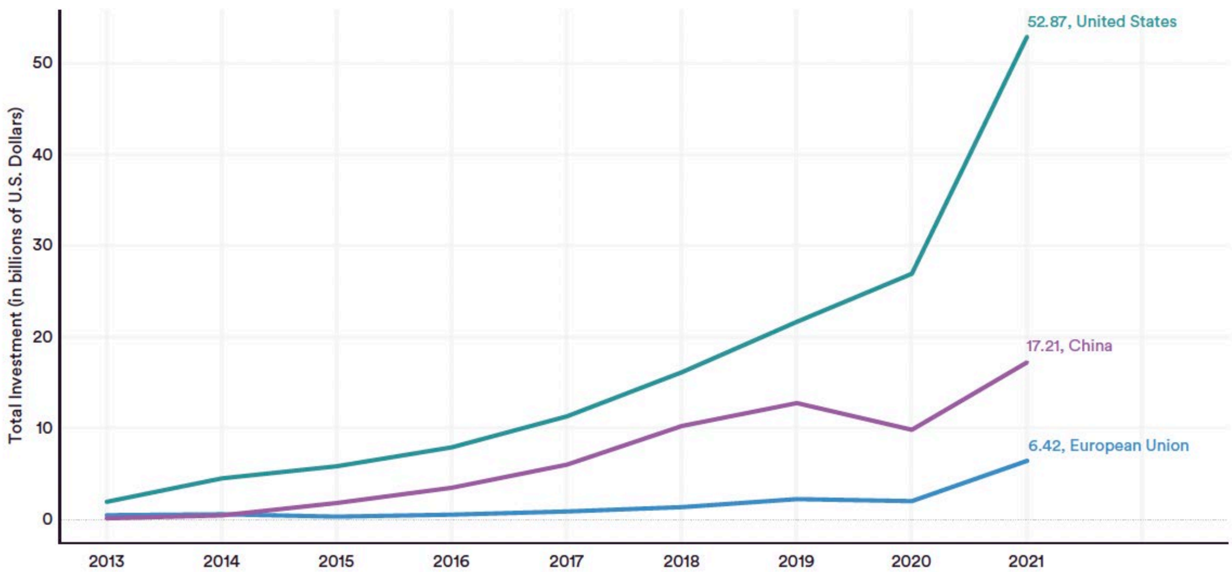
[5] EU10 include: Austria, Belgium, Denmark, France, Germany, Italy, the Netherlands, Spain, Sweden, and the United Kingdom.

emergence of more flexible labour markets, increases in dynamic competition and reallocation effects, also helped the US productivity revival (Jorgenson *et al*, 2008).

Since 2010, the rise of AI has been experienced, which is the next important step in the ICT revolution. AI can be defined as the use of machines and software developed with specific techniques and approaches, for a given set of human-defined objectives, to generate outputs such as content, predictions, recommendations or decisions, influencing the environments they interact with. The most frequently-used techniques and approaches in AI are: machine learning (supervised, reinforcement and unsupervised learning), logic- and knowledge-based approaches (eg inductive logic programming with the use of deductive engines), statistical approaches (like Bayesian estimation), search and optimisation methods.

In 2021, global private investment in AI totalled about $93.5 billion, which is more than double the total private investment in 2020 (Figure 2). This is the greatest year-over-year increase since 2013-2014. The AI investment gap between the US and the EU (without considering the UK) increased in the last few years.

Figure 2: Private investment in AI in the EU, US and China between 2013 and 2021 (billion $)



Source: Zhang *et al* (2022).

A key characteristic of AI systems is that they incorporate a learning-by-doing function that makes them more and more efficient through the execution of their tasks and their experimentation with relevant training data. As a result, AI systems can substantially improve the efficiency of production processes of goods and services if they are 'fed' good quality, relevant (training) data.

AI is considered to be a general-purpose technology (like the steam engine, electricity and computers) with a large variety of applications in many industries and sectors[6]. The fastest growing type of AI technologies is related to deep-learning applications. AI-related patent filings have particularly increased in the field of computer vision, including character recognition, biometrics, scene understanding, image and video segmentation, object tracking and augmented reality (WIPO, 2019).

There have been tremendous improvements in the ability of AI systems to perform given tasks. For example, AI systems managed to outperform humans in image recognition in the frame of the ImageNet Large Scale Visual Recognition Challenge.[7] This challenge evaluates algorithms for their capabilities in object detection and image classification at large scale. For any given word, ImageNet contains several hundred images. In the annual ImageNet contest, several research groups compete to get their AI computers to recognise and label images automatically. Humans on average label an image correctly 95 percent of the time. The respective number for the winning AI system in 2010 was 72 percent, but over the next couple of years the error rate fell sharply. In 2015, machines managed to achieve 96 percent accuracy, reducing the error rate below the human average level for the first time.

Another indicative example is the General Language Understanding Evaluation Benchmark (GLUE):[8] GLUE tests single AI systems on nine distinct tasks in an attempt to measure the general natural language understanding of AI systems, and compares it with the respective

[6] Due to the wide variety of applications, general purpose technologies have a large aggregate impact on the economy (Jovanovic and Rousseau, 2005).
[7] https://image-net.org/challenges/LSVRC/.
[8] https://gluebenchmark.com/.

understanding of humans. Tremendous progress has been made in the accuracy of these systems. Though the benchmark was only released in May 2018, the performance of these AI systems surpassed non-expert human performance in June 2019, and continues to improve further.

AI systems have also improved a lot in other tasks including speech recognition, (visual and verbal) question answering, translation from one language to another, language understanding and inference, summarising texts, sentiment analysis and so on.

Given the increasing efficiency of AI systems, it is not surprising that they improve productivity at the firm level. Damioli *et al* (2021) found that more AI patent applications generate a positive effect on companies' labour productivity, especially for SMEs and in service sectors. In a similar spirit, Babina *et al* (2020) used data from online job postings and employment profiles to find that firms that invest in AI experience faster growth in both sales and employment. However, this effect is most significant for the largest firms in each industry. Alekseeva *et al* (2020) illustrated the importance of managers' AI skills in connection with AI adoption increasing firms' market capitalisation, while Brynjolfsson *et al* (2021) focused on AI prediction systems and showed that the firms that adopt them have higher productivity.

Despite i) the increasing investment in AI technologies in the EU and the US; ii) the tremendous improvements in AI machines (which reached efficiency levels equal to or even greater than humans in specific cognitive tasks) and; iii) the AI firm-level positive productivity effects, AI's contribution to the production process has not been captured by aggregate productivity statistics. Since 2005, US labour productivity has grown at an average annual rate of just 1.3 percent. The slow growth observed since 2010 has been even more striking: labour productivity grew just 0.8 percent from 2010 to 2018. Figure 3 presents labour productivity growth rates up to 2019 for France, Germany and the US. In the first decade since the systematic introduction of AI, no striking increases have been seen in labour productivity growth.

Figure 3: Labour productivity annual growth rate for France, Germany and US between 1971 and 2019

Source: OECD Statistics.

The rest of this paper explains this AI productivity paradox and provides some policy recommendations on how to make labour in US and EU economies more productive based on these new efficient technologies.

2. Potential explanations for the AI productivity paradox

Several explanations have been put forward to answer this question (Brynjolfsson *et al*, 2019). The most pessimistic is that high average productivity gains from AI will not be observed in the long-run. As with the ICT revolution that started in the late 1970s but only paid off temporarily in terms of productivity statistics between 1995 and 2005, AI is not expected to have any major impact on the aggregate economy. A main proponent of this explanation, Robert Gordon (Gordon, 2015, 2018) believes that AI is nothing new and will not change people's lives fundamentally[9].

The second explanation has to do with the fact that AI is in many cases associated with intangible capital that is not so easily measured by official statistics. In order to capture its contribution to productivity new methodologies are needed that will help better measure intangible capital and assess its contribution to the production process and online users' well-being (Brynjolfsson *et al*, 2019).

The chief economist of Google illustrated why intangible capital is difficult to measure with the following example[10]:

> *"In 2000, there were 80 billion photos produced. We know that because there were only three companies that produced film. And fast-forward to 2015, there are about 1.6 trillion photos produced. Back in 2000, photos cost about 50 cents apiece. Now they cost zero a piece essentially. So, any ordinary person would say, wow, what a fantastic increase in productivity, because we've got a huge amount of more output and we've got a much, much lower cost. But if we go look at that from the GDP lens, it doesn't show up in GDP for the most part because those photos are typically traded among friends and put in albums and things like that. They're not sold on the market. GDP is the market value of transactions out there, and anything that's not sold or has a zero price isn't going to show up in GDP."*

In addition, the implications of AI for the transformation of work can introduce additional measurement issues. Bachmann *et al* (2022) showed that people working in occupations that are more exposed to AI are more likely to move to self-employment. AI technologies facilitate some new forms of service work for self-employed individuals, such as platform work, which have not yet been captured by official labour statistics.

---

[9] https://www.irishtimes.com/business/technology/third-industrial-revolution-will-not-transform-our-lives-1.3545650.

[10] https://www.aei.org/economics/googlenomics-a-long-read-qa-with-chief-economist-hal-varian/.

The third explanation is based on the observation that AI and intangible capital investments are concentrated in a few firms (Kaus *et al*, 2020; Altomonte *et al*, 2020). As a result, only a small portion of firms capture most of the benefits from AI technologies and advance their positions in the markets in which they operate. With only few winners from AI, average productivity growth remains low, even if AI technologies are highly productive. De Loecker *et al* (2020) found that market power has significantly increased in the last 15 years in all major economies. In fact, the market power empirical trends are particularly prominent for digital markets, according to Calligaris *et al* (2018), who followed a similar methodology in computing them. In particular, they assigned an index of digital intensity to each sector based on sectoral tangible and intangible ICT investment, purchases of intermediate ICT goods and services, and use of robots. They found that the increase in markups from 2001–2003 to 2013–2014 was greater for the average firm in a digital-intensive sector than for the average firm in the pool of non-digital-intensive sectors[11].

The fourth explanation has to do with the time that needs to pass so that a breakthrough technology to contribute to productivity, and is based on the J-curve (Brynjolfsson *et al*, 2019). The productivity J-curve describes the historical pattern of initially slow productivity growth after the introduction of a breakthrough technology, followed years later by a sharp take-off. For instance, after the introduction of electricity to American factories, productivity was stagnant for over two decades. It was only after managers re-invented their production lines using distributed machinery, a technique made possible by electricity, that productivity belatedly surged. And it's not just electricity that requires such a reinvention of work.

[11] However, such findings should be viewed with some caution. First, how to measure markups is a topic of current debate. For example, Philippon (2019) did not find an increase in markups and concentration in the EU. He only pointed out a sharp increase of concentration in US markets. At the same time, Traina (2018) criticised the way that markups are measured in the literature. Hall (2018) found no evidence that mega-firm-intensive sectors have higher price/marginal-cost markups, but reported some evidence that markups grew in sectors with rising mega-firm intensity. The implications of increasing markups are also debated. One implication is that this trend captures the increase in market concentration. But it may instead refer to higher production efficiency: namely, declining marginal costs, especially in technology-related or information-intensive markets, which lead to increasing markups without necessarily any increase in prices. Nevertheless, the increasing concentration has also been reported by numerous other studies, which implies, even if methodologies are under debate, their qualitative conclusion of rising market concentration depicts reality.

Brynjolfsson *et al* (2021) found that complementary investments in intangible capital are virtually always needed before big technology breakthroughs as diverse as the steam engine or computers ultimately boost productivity. Firms need to rethink their business models, managers need to develop expertise for the digital age, workers need to be retrained to interact with these new technologies, complementary web applications and software need to be designed. Without these complementary innovations it is hard for AI to pay off in terms of aggregate productivity statistics. As a result, over time, there are two distinct phases in the impact of new general-purpose technologies on growth: an initial phase when intangible capital is created and accumulated, followed by a productivity boom.

Consulting the members of the expert group, it became evident that the vast majority believe that a combination of the above reasons explain the AI productivity paradox. First, we can expect the productivity J-curve to be a valid explanation of the paradox. AI requires complementary innovations to take place in order to pay off in aggregate statistics: managers need to get to know what the practical implications of AI are for their business, hiring and training of highly specialised AI talent needs to happen, adopting a more collaborative model of production with the active involvement of both humans and machines can help in arriving to the productivity boom phase.

There are also the measurement issues, but they are not expected to be so important in order to explain the AI productivity paradox alone (Ahmad *et al*, 2017). They may affect the exact shape of the J-curve, namely, they can contribute to the delay until the AI impact is captured in productivity statistics. At times of intangible capital accumulation, because of mismeasurement, it may seem that productivity is smaller than it actually is. But the arrival of the productivity boom phase, with the maturity of such capital investments, will eventually be captured in productivity statistics.

Market power may not have an impact only on the shape, but can also delay the arrival of the productivity boom phase at the J-curve because it limits AI adoption to the top firms in the economy.

The COVID-19 shock, on the other hand, has accelerated the accumulation of intangible capital in the economy (Brynjolfsson and Petropoulos, 2021). The emergence of remote work revealed new efficient ways to produce output even if inputs are restricted (Marcus *et al*, 2022). For example, restrictions on business travel led firms and university researchers to develop new communication and collaboration models to keep output production as high as possible. That effectively can lead to a significant increase in total factor productivity growth in the short-run, at least in sectors in which remote work is possible. However, the biggest impact of the pandemic is expected to be realised in the longer-run. The social distancing restrictions led to a new reality in which investment in digital technologies and digital literacy became necessary. Work and production have been rapidly reorganised through the digital channel. This fundamental shift has had two effects. First, it has allowed the accumulation of intangible capital that is important for arriving at the productivity boom point of the J-curve. Second, it has helped firms and workers understand where the benefits and costs of digital technologies can be found. As the learning curve for these technologies progresses, the COVID-19 shock is more likely to leave a permanent footprint in the organisation of economic relationships and productivity.

3. Policy recommendations for increasing labour productivity

Looking ahead, it is not sufficient to only explain the AI productivity paradox and then assume a passive role, waiting for the productivity boom phase to arrive. Specific policies should be prioritised to maximise knowledge spillovers without impeding innovators' incentives, and to adopt new frameworks that are more suitable for measuring the contribution of AI to productivity.

Knowledge spillovers have traditionally been a central objective of government policy interventions. Under a strong intellectual property regime that keeps the value of innovation high, policies that aim at the better and wider diffusion of AI technologies can be beneficial in building the intangible capital needed to arrive at the productivity boom phase. Becker (2015) and Bloom *et al* (2019) illustrated how R&D tax credits on AI investments can work well towards this goal. In fact, many countries provide additional fiscal incentives for R&D, such as allowing an additional deduction to be made against tax liabilities. However, country measures differ in terms of their generosity. An overall estimation by Bloom *et al* (2019) on these policy measures concluded that a 10 percent fall in the tax price of R&D leads to at least a 10 percent increase in R&D in the long run. Hence, AI tax credits can prompt the better diffusion of these technologies and could allow for a significant accumulation of intangible capital in order to reach a critical mass, putting the economy on the productivity boom path.

Policies are also needed that focus on the supply of human capital and especially the supply of AI talent. One of the major obstacles for the diffusion of AI technologies is the lack of AI talent. Adjusting educational and training policies in order to facilitate a greater supply and variety of AI talent can be very beneficial for improving the technology frontier in industrial production. Managers should become more familiar with the practical implications of AI in order to contribute to the reorganisation of work, towards a model in which AI machines and labour act as complements.

As Elon Musk put it: *"Yes, excessive automation at Tesla was a mistake. To be precise, my mistake. Humans are underrated"*[12]. There is a risk that managers choose to overinvest in automated technologies that do not add much in terms of productivity, while neglecting the productivity boost of combining labour and AI capital in a harmonious way. Parallel to investment in AI, managers should fundamentally change their perspectives on how their firms should adjust their work environments so that workers can become more efficient by using AI machines. A human-centric approach is needed in industrial production and the provision of tasks, in order to grasp the full benefits of technology.

Acquiring AI talent is an important part of this. In fact, AI talent is very concentrated in few superstar firms. Jin *et al* (2021) used US online job posting data from Burning Glass Technologies from January 2010 to June 2020, and found that the top employers account for a large percentage of the total demand for frontier technology skills, including AI, machine

[12] https://twitter.com/elonmusk/status/984882630947753984?lang=en.

learning, natural language processing, cloud computing and big data. More than 26 percent of all job vacancies in the last decade that required AI skills were posted by the top 10 firms that employed people with AI skills. The respective concentration percentage of more 'traditional' information technology skills is only 6.9 percent. Wide adoption of AI in order to maximise its knowledge spillovers, and therefore the social benefits from AI, would require smaller firms to be able to hire AI experts, which will help them make complementary investments in intangible capital in order to grasp a fair share of these benefits.

More AI-related productivity would also require the market-power failure to be addressed. Market power can explain why AI and intangible capital investments are concentrated in only a few firms. As a result, only a small portion of firms capture most of the benefits from AI technologies, and advance their positions in the markets in which they operate.

Addressing the market-power failure in AI-related markets would require a combination of market regulation, competition policies and labour-market policies (Parker *et al*, 2022). Market regulation should set the basic principles of operation so that specific firms do not have an unfair competitive advantage that allows them to grow at the expense of their competitors, even if they are not more efficient in terms of production costs and quality of products and services. Competition policy should ensure that these regulatory principles are adequately enforced, giving to antitrust authorities the ability to intervene in a timely manner and have access to relevant information, in order to evaluate cases of market misconduct. Labour-market policies should embrace flexibility, allowing the AI talent to flow across different firms, but policies should also give workers adequate social protecti]on.

It is therefore a combination of tax, education, labour and competition policies that could speed up the arrival of the productivity boom (revival) phase of the J-curve. Parallel to that, new ways need to be found to better measure productivity in the digital age and with respect to AI (Brynjolfsson and Petropoulos, 2022). Current measurements, such as GDP, are insufficient when they only factor in tangible goods and services that are offered at positive prices. In the digital economy, many intangible goods and services are provided at no financial cost to consumers. These still increase consumer welfare, create jobs and generate profit. Moreover, advancements in AI decision-making and prediction could generate new opportunities for economic growth that have never previously been realised.

4. A framework for building a responsible AI approach by design

Ensuring greater AI adoption and diffusion can bring major opportunities, from increasing efficiencies and improving outcomes, to reimagining industries altogether. AI can transform the relationship between people and technology. However, as AI decisions increasingly influence and impact people's lives at scale, there has also been a rising level of discussion and questions have been raised around AI ethics, data governance, trust, legality and responsibility issues. In particular, bias, discrimination and fairness have emerged as areas of paramount concern, alongside explainability of decisions taken based on AI.

It would therefore be a mistake when implementing policies that seek to more widely diffuse AI technologies, not to also build a framework of ethical rules that guarantees AI's responsible deployment by all shareholders. Rules should cover both the creation of AI technologies and their use.

If unaddressed, the ethical concerns surrounding AI could lead to poor AI performance, resulting in limited to no value from the investment made; regulatory implications, resulting in an inability to use the existing AI solutions; employee resistance to AI, affecting adoption rates; reputational risks, affecting company bransd and putting the survival of companies at risk; unintentional infringements of the law, legal actions, fines and settlements and lack of trust among stakeholders. The question then arises: how can ethical and responsible AI systems be designed, built and scaled up?

Trust in AI is key to realising value from this technology. With trust comes great responsibility, but also a great opportunity, because trust triggers the loyalty and engagement that will drive business innovation and growth. To achieve trust, as organisations start scaling up their use of AI, adhering to laws, regulations and ethical norms is critical to building a sound data and AI foundation. This can be done by implementing **Responsible AI**.

Responsible AI is the practice of designing, building and deploying AI in a manner that empowers employees and businesses and impacts customers and society fairly. At its core, Responsible AI is about considering the impact the use of AI will have on people. Responsible AI builds trust and lays the foundation for successful scaling by taking a 'human first' approach – using technology to help people make better decisions, while keeping them accountable through the right governance processes and technical steps.

Regulation – both imposed by governments and self-imposed by organisations – will be a key part of this equation. Organisations need to be transparent in their use of AI to maintain trust and avoid bias (both in the algorithms created and the datasets used to train them). Organisations should also offer a right of appeal against decisions taken by algorithm, accounting for security concerns, and considering how humans will take back control from an algorithm when necessary. Governments and regulators are considering how to supervise and set standards for the responsible development and use of AI. Countries including the United Kingdom, Brazil and China are already taking action, either by developing existing requirements related to AI (for example, in regulation such as the EU's general data protection regulation), or through the development of new regulatory policy. The EU's draft AI Act, proposed in April 2021, is the best-known example: once ratified, anyone who wants to use, build or sell AI products and services within the EU will have to comply with the requirements of the legislation, depending on the category under which their AI systems fall under i.e. unacceptable-risk AI systems, high-risk AI systems, and limited- and minimal-risk AI systems (European Commission, 2022). The draft EU AI regulations should serve as a reality check for the organisations to ensure they have robust processes and governance systems in place to manage to AI risks and comply with the regulations. Instead of getting discouraged from pursuing the development and use of AI systems, future-ready organizations will take this opportunity to create a robust governance system and risk management framework that will allow their organisation to innovate and deploy AI in a responsible manner.

A 2021 Accenture study of 850 C-suite executives across 17 geographies and 20 industries sought to understand organisations' attitudes toward AI regulation and assess their readiness to embrace it. Nearly all (97 percent) respondents believed that regulation will impact them to some extent, and 77 percent indicated that compliance is a company-wide priority. More than 80 percent said they would commit 10 percent or more of their total AI budget to meeting regulatory requirements by 2024. However, most organisations have yet to turn these attitudes and intentions into action. Only 6 percent of organisations have built their Responsible AI foundation and put their principles into practice, while the majority (94 percent) are struggling to operationalise across all key elements of Responsible AI (Accenture, 2021).

With Responsible AI, organisations can shape key objectives and establish their governance strategies, creating systems that enable AI. Important benefits from responsible AI are:

- **Minimise unintended bias:** Build responsibility into AI to ensure that the algorithms – and underlying data – are as unbiased and representative as possible. This entails ensuring the policies, processes, governance systems and the organizational culture are built with ethics and fairness principles at the core.

- **Ensure AI transparency:** To build trust among employees and customers, develop explainable AI that is transparent across processes and functions.

- **Create opportunities for employees:** Empower employees to raise doubts or concerns with AI systems and effectively govern technology, without stifling innovation.

- **Protect the privacy and security of data:** Leverage a privacy and security-first approach to ensure personal and/or sensitive data is never used unethically.

- **Benefit all stakeholders:** By creating an ethical underpinning for AI, one can mitigate risk and establish systems that benefit all stakeholders, including shareholders, employees and society at large.

There is an urgent need to establish flexible guiding principles to govern AI that are general enough to evolve with a rapidly changing technological environment, but are also specific enough to be useful for applications. Clearly articulated and established guiding principles can help enable a cohesive stance on ethics and AI that remains consistent with long-term technological advancements. The guiding principles will also help create a culture within an organisation that allows for ethically responsible attitudes and behaviours from top to bottom.

A strong, existing compliance structure within an organisation is the greatest starting point for implementing an ethical governance process around AI – employees who feel that they can confidently raise personal problems at a company without actions taken against them will similarly be emboldened to raise ethical issues they identify in technology. By leveraging existing structures within the organisation, a stronger ethical culture can be created around responsible AI.

For organisations, the upside of being responsible by design is an improved ability to meet future requirements, better mitigate risks and create sustainable value for themselves and their stakeholders. Being responsible by design will become more beneficial over time, especially as governments and regulators consider new standards for the development and use of AI.

Organisations can build a Responsible AI foundation supported by these four key pillars:

- **Organisational:** Strong leadership is pivotal to empower employees and elevate Responsible AI as a business imperative. To democratise this way of working, successful organisations recognise the need for new roles, and actively upskill, re-skill or hire. Organisations should nurture cultures that empower individuals to raise concerns over AI systems, without stifling innovation. Clear success criteria, incentives and training are all critical requirements.

- **Operational:** Organisations need to ensure considerations for AI are built into their core values and robust compliance processes. They need to establish transparent, cross-domain, governance structures, including systems, measures and controls that enable AI to flourish. These build internal confidence and trust in AI technologies by identifying roles, expectations and accountabilities.

- **Technical:** Leading organisations deploy AI models, systems and platforms that are trustworthy and explainable by design. Technology tools should support fairness, explainability, robustness, accountability and privacy. Organisations can leverage proven qualitative and quantitative techniques for assessing potential risks to reach cross-domain consensus on mitigation.

- **Reputational:** Leading organisations clearly articulate their responsible business mission, anchored in their values. Ongoing measurement and monitoring of key Responsible AI metrics ensures they're managing risk and communicating with transparency.

Being responsible by design means that organisations understand the importance of incorporating Responsible AI into their data and AI strategies from the start. They operate a responsible data and AI approach across the complete lifecycle of all of their models, enabling the organisation to engender trust and scale AI with confidence. With the foundations in place to support the responsible use of AI across the enterprise, it becomes easier to adapt as new regulations emerge. That way, businesses can focus more on performance and competitive advantage.

Organisations can become responsible by design by taking the following steps:

1. **Articulate clear principles and governance structures for AI**. Organisations need to review their existing business values and evaluate how Responsible AI fits into their overall mission. They need to define and communicate their Responsible AI mission and principles, as well as key objectives and key performance indicators (KPIs), across

the organisation. Establish an AI-specific governance framework that includes roles/responsibilities required to support key initiatives. Use C-Suite to drive broad Responsible AI awareness across the organisation, positioning Responsible AI as critical to the business strategy and risk-management decision making. Operationalise the Responsible AI governance model across the organisation and include incentives to accelerate adoption.

2. **Develop a risk management framework that monitors and operationalises current and future policies.** Organisations need to review their existing risk management models and conduct a gap analysis of all risk processes against all potential AI risks (data protection, human rights, ethical risks, accuracy, legal, etc). They need to document their Responsible AI risk management strategy and update the company-wide risk management framework to incorporate the new considerations for AI. Operationalise the updated risk management framework across the enterprise, leveraging new procedures and checkpoints throughout the data and AI lifecycle. Establish traceability/auditability processes to monitor decisions and changes, and measure key KPIs across the data and AI lifecycle.

3. **Invest in technology tools that support fairness, explainability, robustness, accountability and privacy, and build these into AI systems and platforms.** Organisations need to review existing tools and techniques that support responsible data and AI (ie tools that monitor bias, fairness, explainability). Review the model development and lifecycle management process to understand how AI tools/systems are developed and maintained. Define measurable performance metrics and establish techniques for continuous monitoring, control and re-assessment of data and AI systems. Establish and communicate clear roles and responsibilities for those managing every stage of the AI development lifecycle.

4. **Building a culture of Responsible AI.** It is important that every employee has a sufficient understanding of, and confidence in, the approach the organisation is taking

to ensure the responsible use of AI. With AI talent already scarce, organisations must consider how to attract or develop the specialist skills required for Responsible AI roles – keeping in mind that teams responsible for AI systems should also reflect a diversity of geography, backgrounds and 'lived experience'. The different perspectives that they bring are essential to spot potential bias and unfairness, and to minimise unconscious bias in product design, build and testing. When it comes to AI, culture is essential to uniting the whole organisation around responsible principles and practices. It is critical for organisations to establish a system for continuous review of roles, skills and training to match advances in AI and address new risks. Consider where roles need to be augmented or added to support the responsible development and use of AI at scale. Specify new roles, skills and learning agendas required to support Responsible AI across the organisation.

All of these elements are part of an innovation-friendly blueprint for Responsible AI that can be applied across functions and projects, allowing the ethical implications of AI to be understood and managed and ensuring the organisation has the foundations in place to adapt as new regulations and guidance emerge. That way, organisations can focus more on performance and competitive advantage.

References

Accenture (2021), "From AI Compliance to Competitive Advantage". Available at https://www.accenture.com/_acnmedia/PDF-179/Accenture-Responsible-by-Design-Report.pdf#zoom=40

Ahmad, N., J. Ribarsky and M. Reinsdorf (2017) 'Can potential mismeasurement of the digital economy explain the post-crisis slowdown in GDP and productivity growth?' OECD Statistics Working Papers, No. 2017/09, OECD Publishing, Paris

Alekseeva, L., Gine, M., Samila, S., & Taska, B. (2020) AI Adoption and Firm Performance: Management versus IT. Available at SSRN 3677237.

Altomonte, C., Favoino, D., Morlacco, M., & Sonno, T. (2021) Markups, intangible capital and heterogeneous financial frictions. Centre for Economic Performance, London School of Economics and Political Science.

Babina, T., Fedyk, A., He, A. X., & Hodson, J. (2020) Artificial intelligence, firm growth, and industry concentration. Firm Growth, and Industry Concentration. Mimeo.

Bachmann, R., Gonshor, M., Milasi, S., & Mitra, A. (2022) Technological progress and the dynamics of self-employment: worker-level evidence for Europe.

Becker, B. (2015) Public R&D Policies and Private R&D Investment: A Survey of the Empirical Evidence. Journal of Economic Surveys 29(5): 917–42.

Bloom, N., Van Reenen, J., & Williams, H. (2019) A toolkit of policies to promote innovation. Journal of Economic Perspectives, 33(3), 163-84.

Brynjolfsson, E., Jin, W., & McElheran, K. (2021) The power of prediction: predictive analytics, workplace complements, and business performance. Business Economics, 56(4), 217-239.

Brynjolfsson, E., & Petropoulos, G. (2021) The coming productivity boom. MIT Technology Review.

Brynjolfsson, E., & Petropoulos, G. (2022). Advancing a More Productive Tech Economy. A White Paper for the Stanford Institute for Human – Centered Artificial Intelligence and the Stanford Digital Economy Lab.

Brynjolfsson, E., D. Rock, C. Syverson (2021) The Productivity J-Curve: How Intangibles Complement General Purpose Technologies, American Economic Journal: Macroeconomics.

Calligaris, S., C. Criscuolo, and L. Marcolin (2018) Mark-Ups in the Digital Era. OECD Science, Technology, and Industry Working Paper.

Damioli, G., Van Roy, V. & Vertesy, D. (2021) The impact of artificial intelligence on labor productivity. Eurasian Bus Rev 11, 1–25

De Loecker, J., J. Eeckhout, and G. Unger (2020) The Rise of Market Power and the Macroeconomic Implications. Quarterly Journal of Economics 135, no. 2: 561–644.

European Commission (2022), "Shaping Europe's digital future - Regulatory framework proposal on artificial intelligence" Available at: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

Gordon, R (2015) "Secular stagnation: A supply-side view", American Economic Review, Papers & Proceedings 105(5): 54-59.

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

Gordon, R. J., and Sayed, H. (2020) Transatlantic technologies: The role of ICT in the evolution of US and european productivity growth (No. w27425). National Bureau of Economic Research.

Hall, R. (2018) Using Empirical Marginal Cost to Measure Market Power in the US Economy. NBER Working Paper 25251.

Jin, W., G. Petropoulos, and S. Steffen (2021) Concentration of artificial intelligence and other frontier IT skills, Bruegel Blog, 21 October.

Jovanovic, B., Rousseau, P. L. (2005) Chapter 18 - General Purpose Technologies. Handbook of Economic Growth, Editor(s): Philippe Aghion, Steven N. Durlauf, Elsevier, Volume 1, Part B, Pages 1181-1224.

Jorgenson, Dale W., Ho, Mun S., and Stiroh, Kevin J. (2008) A Retrospective Look at the U.S. Productivity Growth Resurgence. The Journal of Economic Perspectives, 22(1), 3-24.

Kaus, W., Slavtchev, V. and Zimmermann, M. (2020) 'Intangible Capital and Productivity. Firm-Level Evidence from German Manufacturing', IWH Discussion Paper No 1/2020.

Marcus, J. S., G. Petropoulos and A. Aloisi (2022) COVID-19 and the Accelerated Shift to Technology-Enabled Work from Home. Available at SSRN: https://ssrn.com/abstract=4025515.

O'Mahony, Mary, Timmer, Marcel P., and van Ark, Bart (2008) "The productivity gap between Europe and the United States: trends and causes," Journal of economic perspectives 22 (Winter, no. 1), 25-44.

Parker, G., G. Petropoulos, M.W. Van Alstyne (2022) Digital platforms and antitrust. The Oxford Handbook of Institutions of International Economic Governance and Market Regulation. Editors: Eric Brousseau, Jean-Michel Glachant and Jérôme Sgard.

Perez, Carlota (2002) *Technological Revolutions and Financial Capital. The Dynamics of Bubbles and Golden Ages*, (Edward Elgar, Cheltenham).

Philippon, T. (2019) The Great Reversal: How America Gave Up on Free Markets. Cambridge, MA: Belknap Press of Harvard University Press.

Schwab, K. (2017) The fourth industrial revolution. Currency.

Stiroh, Kevin J. (2002) 'Information Technology and the U.S. Productivity Revival: What Do the Industry Data Say?' American Economic Review 92 (December, no. 5), 1559-76.

Traina, J. (2018) Is Aggregate Market Power Increasing? Production Trends Using Financial Statements. Available at SSRN: https://doi.org/10.2139/ssrn.3120849.

WIPO Technology Trends 2019 - Artificial Intelligence. DOI: 10.34667/tind.29084.

Daniel Zhang, Nestor Maslej, Erik Brynjolfsson, John Etchemendy, Terah Lyons, James Manyika, Helen Ngo, Juan Carlos Niebles, Michael Sellitto, Ellie Sakhaee, Yoav Shoham, Jack Clark, and Raymond Perrault, "The AI Index 2022 Annual Report," AI Index Steering Committee, Stanford Institute for Human-Centered AI, Stanford University, March 2022.