PrometheEUs

THE N(EU) WAY TO ARTIFICIAL INTELLIGENCE

Challenges and Perspectives for Southern Europe

September 2020

Abstract

The idea of an "AI race" has attracted a great deal of public and academic attention in recent times. It is widely understood that China, the US and – lagging slightly behind – the EU are entangled in a competition for the development, the implementation and the regulation of a technology poised to bring about profound social and economic transformations. In the first chapter of this joint paper, we analyze the competing "AI models" of each of the three superpowers, as well as the dangers of framing the competition as a "confrontational race". Afterwards, we justify the need for a global governance framework, and we suggest an approach based on core and extended principles of global governance, where the EU could play a major role.

Europe has to address several crucial challenges when the critical enablers for an AI transition are examined. In the second chapter, we examine the current European AI landscape, which indicates a substantial gap regarding the AI adoption between Northern and Southern Europe for several dimensions ranging from corporate digitisation and AI performance to human capital skills. Next, we thoroughly address the challenges that Europe should confront to achieve inclusive and sustainable growth through AI transition including synergies between academia and the private sector, improved infrastructure and enhancement of ICT skills, focusing on European SMEs' transition to an AI business model by promoting clustering, digitisation and AI culture.

Since 2018, the European Union has placed the need of a common approach to AI at the heart of its digital agenda. In the third chapter of the paper, we recap the DSM strategy and take stock of the EU's AI policies developments, from the 2018 AI Strategy to the 2019 AI HLEG Documents. Afterwards, we analyse the Von der Leyen Commission's Digital Package and particularly the AI White Paper (February 2020), aimed at establishing a regulatory and investment-oriented approach with the twin objective of promoting the uptake of AI and addressing the risks associated with its uses. Finally, we present our remarks on the White Paper, particularly focussing on how to foster EU innovation and a thriving ecosystem based on SMEs.

CONTRIBUTORS

CHAPTER I - EsadeGeo

Tirso Virgós Varela
Óscar Fernández

Capter II – IOBE
Foundation for Economic & Industrial Research

Aggelos Tsakanikas

Chapter III – I-Com
Institute for Competitiveness

Stefano da Empoli
Silvia Compagnucci

**This joint paper on AI has been developed by I-Com, Institute for Competitiveness (Italy), EsadeGeo - Center for Global Economy and Geopolitics (Spain) and IOBE - Foundation for Economic and Industrial Research (Greece).*

PrometheEUs

SUMMARY

PrometheEUs

EXECUTIVE SUMMARY

Introduction

Looking at data, a pattern of competitive disadvantage of Europe compared to the US and China seems to emerge. In terms of patents, according to the European Commission's Joint Research Center (2018), between 2009 and 2018, Chinese AI players filled 57% of requests, versus 13% from the US and 7% from South Korea and the European Union.

In 2018, over 51% of published AI patents were attributed to North America, with the share of Europe and Central Asia declining to 23%, less than two percentage points above East Asia and the Pacific (in 2002 the gap between Europe and East Asia was approximately 20 percentage points) (Artificial Intelligence Index Report 2019, Stanford University).

Out of 41 AI unicorns in March 2019, 18 were based in the US, 17 in China and only 1 in the EU (Global Artificial Intelligence Industry Data Report, China Academy of Information and Communications Technology).

In 2018, a wider look at the start-up environment allowed Roland Berger and Asgard to survey 769 EU startups specialized in AI, much less than in the US (1,393) but significantly more than in China (383). However, close to one-third of the EU startups were based in the UK, now no longer an EU Member State. Moreover, in a world ranking of hosting cities, the first EU hub (Paris) is only in 10th position, due to the extreme fragmentation in Europe.

Fragmentation is definitely one factor at play, reducing EU chances to become an AI world leader, frustrating the high potential to be found in the number of top EU scientists (by far the highest in the world, according to a recent study by Tsinghua University).

However, the gap in the overall amount of investments appears to be the most startling reason for Europe lagging behind the US and China.

In 2018, according to Stanford University estimates, US companies invested $18.7 billion in AI, compared to China's $14.35 billion. The 5 largest EU Member States were not even able to attain together the level of UK investment ($1.255 bn vs. $1.27 bn), only slightly surpassing the much less populated Israel ($ 1.044 bn).

By contrast, also in terms of public investments, the US is allocating large resources, projected to amount to approximately $5 billion ($ 4 bn from the Department of Defense) (Artificial Intelligence Index Report 2019, Stanford University). The same holds true for China.

1. The Global Race for AI: China, the US and the EU

1.1 The three main AI models

China, the US and the EU have shown different approaches to the development, implementation and regulation of AI, as outlined in their respective strategies (New Generation of AI Development Plan, American AI initiative and White Paper on AI). China designed a plan based on three sequential goals: (1) to reach a globally advanced level in AI by 2020, (2) to use AI as the main driver for upgrading its economic and industrial capabilities by 2025, and (3) to become the main innovation hub in the world by 2030. This strategy is mainly focused on the role of the state, which protects and invests in those businesses that have proved successful after an initial phase of fierce domestic competition.

On the other hand, the US gives a much more prominent role to the market and the investments made by big corporations, maximizing space for innovation through lighter public regulation. The US strategy relies to a greater extent on voluntary self-regulation and the protection of values it considers to be "core", such as freedom, human rights and the rule of law. Finally, the EU is trying to develop legislation that paves a third way between the other two models, promoting what is dubbed as "ethical and trustworthy AI". Its strategy aims at fostering cooperation between the public and private sector, as well as ensuring a set of values that are considered to be key (e.g., transparency, accuracy, robustness and non-discrimination).

1.2 The need for a global governance framework

Each of the three models mentioned above is linked to a different set of priorities that could eventually inspire a global regulation. Thus, these differences reinforce the narrative of an "AI race". Certainly, competition between states in the development and governance of strategic technologies – such as AI – is not necessarily negative. However, we must be wary of the risks of this competition taking a confrontational turn. By constructing a network of shared principles, rules and institutions, we believe the balance can be tilted again towards a more desirable combination of competition and cooperation. But what specific incentives do states have to adopt a global governance framework?

In our view, there are two main justifications for the development of this framework. Firstly, the need to foster stability and a level playing field in international relations. A confrontational race, with countries desperately trying to overtake others in the development of their AI capabilities, could produce adverse outcomes, such as a rise of "AI nationalism", a loss of talent and resources (flowing from less developed nations to the leaders of the race), and unilateral rule-setting by the superpowers and strong international firms. A new form of "data colonialism" could emerge, and the will of citizens from countries on the losing end could be easily subverted, with tensions in the realm of AI translating into more open conflicts, to the detriment of the superpowers

themselves. Moreover, there would be no way to ensure accountability for any actions that threaten international stability.

The second justification for this global framework is related with the negative spill-over effects that AI development may have beyond national borders. For instance, drones and other types of Lethal Autonomous Weapons (LAWs) could become more precise, attain a higher degree of autonomy, and even operate without human oversight once they are sent on a mission. Similarly, cyber-attacks – with their associated economic and possibly political damage – could increase the aforementioned tensions even more.

1.3 A two-tier governance model

In order to avoid the problems of an unregulated AI race and to tame the potential for escalation, we advocate the adoption of a governance model that goes beyond "informal", trust-based arrangements, while avoiding overly intricate frameworks that entail the creation of new institutions and "global oversight bodies".

We would make a distinction between core and extended principles, depending on their substance and on who the adherents would be. On the one hand, the core principles are broad enough to be accepted by all nations, while they have enough depth to regulate some of the most acute problems presented by the development of AI. On the other hand, the extended principles would expand the rules established in the core principles, and initially they would be adopted by a "coalition of the willing", which would gradually become larger if these extended principles prove effective. The ultimate goal of this two-tier, multi-speed governance model is to strike a balance between ambition and realism. We have based our principles on previous work by the OECD and the G20, which made some progress on possible frameworks for future regulations.

Our first core principle would be a common definition on what is meant by AI, as this would lay the foundation for all further discussions. Secondly, human oversight would be required in all operations involving LAWs. Thirdly, transparency in the functioning of AI would be ensured, in order to enable citizens to understand both the process and the outcome of the decisions taken by the machines. The final core principle focuses on the need for robust AI, so that no malfunctioning poses any risk to a human being. The extended principles would build upon these foundations, adding several elements to them. Firstly, legal instruments would be developed in order to ensure the accountability of human operators, programmers or those giving the orders in the actions of LAWs, including a system of reparations. Secondly, a broad agreement would be reached on the protection of the private sphere of citizens, especially with regard to their data. Finally, non-discrimination would be enshrined as a core principle in the development of AI, and this principle would be enforced at both a national and an international level.

We believe that the EU is especially well positioned to spearhead the creation of this global governance framework, given its vast experience in multilateral negotiations and in brokering deals between distant partners. At the same time, the EU is strongly associated with its soft power and its regulatory power, which has given rise to the so-called "Brussels effect" (Bradford,

2012). By becoming a referee, capable of setting the standards for the global regulation of this new technology, the EU could eventually emerge as a clear winner of the AI race.

2. The European Response: Adoption of AI in Europe and challenges ahead

2.1 The European AI landscape

Based on the trends outlined previously, the situation in Europe signals the need for further action. The Digital Economy and Society Index (DESI) for 2019 signals an asymmetry in EUs digital advance. Scandinavian countries outperformed among European members, while most of the Southern – Eastern European countries ranked below the EU average. The United Kingdom, Sweden, Finland, and Ireland seem more ready to welcome an AI transition since they ranked in the top 25% in AI Readiness Index for at least the half categories of the index. On the other hand, Spain, Portugal, Italy, and Greece did not achieve this percentile in any of the examined categories (Bughin et al., 2019). The European presence in several rankings about the AI technologies adoption worldwide is predominated by Germany (WIPO, 2019).

According to McKinsey's 2018 Digital Survey, European companies are less mature not only in their state of diffusion of digital technologies but also in the use of those technologies for new services and business models. Only two European firms are in the worldwide top 30 digitized companies, and at the end of 2017, none of the ten largest internet companies worldwide was based in Europe. Also, less than 50% of the European firms have implemented one AI technology, and most of them are in the pilot stage (Bughin et al., 2019). In terms of investment levels, startups in the United Kingdom received 55% of the EU total investment from 2011 to mid-2018, followed by Germany (14%) and France (13%), when the remaining 25 countries appropriated less than 20% of all private AI equity (OECD, 2019a).

Taking on AI requires a high level of business digitization. According to the Digital Intensity Index (DII), Finland and Denmark possessing more than 10% of highly digitized firms (percentage to total enterprises) followed by Sweden with 8%. In Bulgaria, Greece, Spain, and Italy, the majority of firms (over 55%) disposed of low investments in digital technologies. Likewise, the DESI index suggested a vast gap among European countries regarding the Integration of Digital Technology. Ireland, Netherlands, Belgium, and Scandinavian countries achieve the highest positions, whereas Cyprus, Greece, and Italy rank below the EU average.

The top five contributors to ICT sector value added in the EU for 2016 were the EU's five largest economies (Germany, the United Kingdom, France, Italy, and Spain) accounting for 69% of the total EU ICT sector value-added. As a proportion of GDP, Spain, Italy, Greece, and Portugal ranked below the EU average, and as a share of total employment, Southern European Countries are lagging as well. Bughin et al. 2019 state that Europe is already in supply shortage of the advanced skills needed for AI transformation. The lack of skills is also suggested by the Human capital dimension of DESI 2019. Apart from Malta, the Southern European countries listed below the EU average, indicating a severe insufficiency in digital skills progress.

2.2 European challenges for successful AI adoption

Ensuring inclusive and sustainable growth requires that the geographical disparities regarding AI adoption inside the continent are confronted. An apparent gap in AI readiness inside Europe is evident, with Southern and Eastern Europe being critically behind than the Northern states. Furthermore, it is vital to promote collaboration among research institutions and encourage synergies with the private sector.

Human skills are imperative for AI adoption since both high technological capabilities and emotional intelligence abilities, which cannot be developed by machines, are equally important. Human capital skills are already in shortage of more developed countries concerning AI applications. Vital regional differences are also present in this dimension. Southern EU member states demonstrate a severe insufficiency in digital skills progress. Considerable risk of a potential digital exclusion exists. For instance, in Portugal and Greece, approximately half of their citizens do not possess at least the necessary digital skills. The vast disparity among EU states is present as regards the share of the labour force that has no digital skills. What is more, ICT specialists are not equally shared among EU member states as well. The South experiences a significantly less proportion of them relating to the total level of employment. Slightly more encouraging is the landscape pertaining to the ratio of ICT graduates to the total number of graduates, but still, there is potential to improve. The main problem seems to be not on the availability of ICT professionals but on how they are employed in the business sector. A closer collaboration between the ICT sector and non ICT-firms is necessary in order to highlight the gaps that need to be bridged. ICT expenditures is not a simple trade transaction of ICT goods. For ICTs to be fully integrated in business processes, additional training needs to be implemented and day to day support at least at the first phases of ICTs diffusion.

Challenges can be traced to the notably lower business digitalization and e-commerce development in the Southern countries in contrast to the Northern ones where it is blooming. Investment in AI development is significantly lower in the Southern firms compared to their Northern counterparts. This is also evident in the small share of R&D investment, where current statistics point out a high concentration in the four largest economies in the EU. In general, firms located in the South of Europe appear to be less capable of developing, adopting and transferring innovation and technology breakthroughs. Furthermore, small business size, which is a structural characteristic of the business demography in the South, is a significant barrier and affects any efforts to improve the technological advancement and invest heavily in AI.

Specifically, SMEs have to deal with a series of issues for being able to follow the current trends in AI adoption. Cultural barriers and a fear of change and transition to a new business model, are affecting negatively such decisions. The transition to an AI model, in most cases, is not included in the strategic business objectives. Evidence might suggest that firms in Southern Europe cannot fully appropriate the benefits that AI could offer to their operation and services. The shortage of technology experts and the difficulty for SMEs to attract them is another critical issue that delays the AI escalation. From the infrastructure perspective, digitization of the firm and its services is

key for AI enabling. However, SMEs have been slower than bigger players in advancing towards said digitalisation.

On another note, the lack of prominent technology companies in Europe affects the development of novel firms. A new entrepreneur would be more willing to undertake AI investments if larger firms that could buy or invest in this venture were available. This possibility mitigates the undertaken risk substantially, encouraging SMEs' growth. Furthermore, startups and SMEs might be in an adverse position compared to larger firms regarding their ability to absorb public funding, since the bureaucratic and technical procedures are not negligible.

These differences matter. Countries that are AI leaders could have faster growth and higher productivity than the rest, making the lag even more substantial and difficult to catch up. Primarily, if the potential digital gap among the EU and the other digital leaders (China and the US) is taken into account, the in-house challenges should be addressed holistically and instantaneously.

3. The European Commission's Digital Strategy and the AI White paper

The first Von der Leyen Commission digital proposals were published on 19 February 2020, including two Communications (Shaping Europe's digital future & A European Strategy for data), a white paper (Artificial Intelligence: a European Approach to excellence and trust) and two reports (B2G Expert Group Report: Towards a European Strategy on business-to-government datasharing for the public interest and the Commission Report on Safety and liability implications of AI, the Internet of Things and Robotics).

The current EC updated and upgraded the DSM strategy, with its priorities and proposals. In the **Communication "Shaping Europe's digital future"**, the Commission establishes three key objectives to ensure digital transformation complies with European values: 1) a technology that works for people; 2) a fair and competitive economy and 3) an open, democratic and sustainable society.

As data is the essential enabler for AI, the **European Data Strategy** aims at Europe emerging as a leader in the data economy, providing for a single market for data and a larger role for European companies.
The Commission starts from acknowledging that the EU has the potential to be successful in the data-agile economy, thanks to its technology, its know-how and its highly-skilled workforce. However, several issues are holding the EU back from realising its potential in the data economy, mainly due to the fragmentation between Member States (compared to the small number of US and China-based Big Tech firms). Among the most important issues, the strategy lists: 1) availability of data; 2) imbalances of market power; 3) data interoperability and quality; 4) data governance; 5) data infrastructures and technologies; 6) empowering individuals to exercise their rights; 7) skills and data literacy; 8) cybersecurity.

Included in the actions envisaged by the strategy, the Commission aims at supporting business-to-business data sharing, investing in a High Impact Project on European data spaces and

federated cloud infrastructures, by the establishment of EU-wide common, interoperable data spaces (in manufacturing, environment, mobility, health, finance, energy, agriculture, public administration and skills) and the setting up of a cloud service marketplace, empowering individuals regarding their data and investing in skills and general data literacy.

3.1 AI white paper: the striving for excellence and the need for regulation

The AI White Paper aims at setting a framework for trustworthy Artificial Intelligence, based on excellence and trust.

In the so called "**ecosystem of excellence**", among several planned actions, the Commission aims at proposing to the Member States a revision of the 2018 Coordination Plan, facilitating the creation of excellence and testing centers that can combine European, national and private investments. This involves working with Member States to ensure that at least one digital innovation hub per MS has a high degree of specialization in AI, setting up a new public-private partnership in AI, data and robotics in the context of the Horizon Europe Programme.

For the other ecosystem ("**ecosystem of trust**"), the Commission assesses the main risks associated with AI in order to ensure a European regulatory framework for a trustworthy AI.
The risk-based approach allows for a proportionate regulatory intervention, heavier for high-risk AI applications than for other lower-risk applications.
According to the white paper, an AI application should be considered high-risk when it meets the following two cumulative criteria: 1) it is employed in a sector where, given the characteristics of the activities typically undertaken, significant risks can be expected to occur (for instance, healthcare, transport, energy and parts of the public sector); 2) the AI application in the sensitive sector is used in such a manner that significant risks are likely to arise (based on the kind of impact on presumably affected parties). Moreover, the use of AI applications for employment processes, biometric identification and other intrusive surveillance purposes would always be considered as high-risk.

Mandatory requirements for high-risk applications would cover the following areas: 1) training data; 2) data and record-keeping; 3) information to be provided 4) robustness and accuracy; 5) human oversight; 6) specific requirements for certain specific applications, such as biometric identification.
These requirements would be at least in part verified under prior conformity assessments, in line with already existing mechanisms for a large number of products being placed on the EU's internal market.
Of course, ex post controls could be still enforced by competent national authorities.
For non-high-risk applications, the Commission envisages a voluntary labelling scheme, allowing the economic operators to signal the trustworthiness of their products or services.

3.2 Some remarks: how to foster EU innovation and a thriving ecosystem based on SMEs

An ecosystem of excellence

While the EU should strive to improve its current standing in research and innovation, increasing public and private investments but also better coordinating existing initiatives, most companies, especially SMEs, would be either only or mainly AI users, buying technologies developed by other companies. Therefore, for a competitive economy, the EU regulatory framework should lead the vast majority of companies to adopt AI easily and at a cost to be competitive.

EU objectives to increase R&D and productive excellence should not jeopardize the possibility for EU citizens and companies to have access to the best available AI technologies at a competitive price. A balanced approach should be used taking into full account the interests of all the concerned parties, including the vast majority of citizens and companies that would be adopters rather than R&D and/or commercial producers in the AI ecosystem.

Both training and advice to SMEs should be key activities for AI specialized digital innovation hubs (DIHs). For this reason, foreseeing only one DIH per Member State may involve a sizeable geographical barrier for SMEs, especially in larger countries. A more distributed network of DIHs providing expertise to SMEs in different regions should be pursued, possibly involving trade associations and larger AI technology players.

An ecosystem of trust

To start with, a EU-wide regulatory approach is preferable in order to avoid major risks of internal market fragmentation. Therefore, Member States should refrain from unilateral moves and look for agreements and alliances at EU level.

Although some new legislation is certainly required and a EU-wide regulatory framework is surely preferable to national, current legislation should apply whenever possible in order to avoid excessive market fragmentation and uncertainty and increase compliance costs for companies, especially SMEs.

Whenever possible, a clearer interpretation of current legislation to be applied to all products, including those embedding AI, should be chosen instead of new legislation reserved to AI products.

Although many mentioned concerns deserve a high level of scrutiny and sometimes need to be addressed by ad hoc regulation, it would be fairer to compare AI applications with a human-based benchmark. It would not be realistic to expect AI achieving an error-free perfection where, in the same field, the same standard is not currently applied. This requirement could significantly stifle innovation, especially from SMEs and new entrants.

If the two proposed cumulative criteria to determine “high-risk” AI applications seem quite logical and could help provide legal certainty, exceptional additional instances should be better defined and limited to specific cases in order to avoid any ambiguity, where the aim of the risk-assessment approach is exactly the opposite.

Concerning enforcement, the regulatory framework should mostly rely on ex-ante self-assessment, instead of an external procedure, in order to speed up the innovation process and ensuring a thriving European AI ecosystem, setting low compliance costs for SMEs, and ex-post enforcement, paramount to guaranteeing full compliance by AI developers and deployers.

PrometheEUs

CHAPTER I
The Global Race for AI: the US, China, and the EU

1.1 The state of the race

A recurring topic in today's public debate is that of an "AI race" between superpowers: the United States (US), China and the European Union (EU). There is ongoing competition for the development and implementation of a "General Purpose Technology" (Gill, 2020) which **could have an impact similar to that of the steam engine or electricity**. If a country were to win the race, it could become the new economic leader of the world, hosting "the next Google or Amazon" (Moore, in LeVine, 2018). Consequently, the outcome of this struggle will have a lasting impact on the future of our societies.

While the US is widely considered to be the country with the most advantages for the development and implementation of AI, **several experts (Lee, 2019; Castro et al., 2019) believe that China is already gaining ground and even surpassing the US in certain areas**. Lee, for instance, argues that China realized its need to invest in AI after the defeat of the world's top Go player, Ke Jie, in a match against Google's AlphaGo Master computer. It was a highly symbolic victory for AI in a game that was considered impossible to master by a non-human, due to the enormous number of possible moves. Ke's defeat opened the eyes of the leaders of the Communist Party of China (CPC) to the potential of this new technology[1] and created a sense of urgency to win the oncoming race.

The EU, on the other hand, lags behind both China and the US. Even European Commission reports have recognized this situation (Craglia et al., 2019). Some observers argue that this is due to a lack of this "sense of urgency" (Schenker, 2019) or "Sputnik moment", in the words of Lee. Others point to the institutional framework of the EU and its difficulties to amass enough data, computational power and top AI researchers to match the other two competitors (Stix, 2020).

Currently, of the top 30 AI patent applicants worldwide, only 4 are European (WIPO, 2019). Similarly, according to CB Insights (2020), only 6 of the 100 most promising start-ups in the world come from the EU. The weight of the Information and Communications Technology (ICT) sector

[1] Lee calls this "The Sputnik moment", echoing the realization by the Americans of the progress that the Soviets had made in the Space Race during the Cold War after the launching of the Sputnik satellite. Similarly, China saw how far ahead the US and its big companies were in terms of AI development.

in the GDP is 1.66% in the EU, compared to 3.33% in the US and 2.16% in China (Bughin et al., 2019). And of the worldwide investments in AI made in 2017, 38% came from the US, 48% from China and 13% from the rest of the world combined (CB Insights, 2017). Nevertheless, the EU fares better in terms of publications in academic journals and has the largest number of AI researchers, surpassing even the US (Castro et al., 2019). Yet it lacks in ability to retain this talent, which usually moves to the US and its academic institutions and big companies. **Consequently, some of the EU's best assets end up benefitting its competitors.**

To put an end to this drain of talent, and to recover the lost ground, the European Commission has recently released its strategies both on AI and data. Their declared objective is to "empower people with a new generation of technologies" (EU Commission, 2020), making the transformation work for both people and business. **The EU is not alone in the development of strategies to build a model of AI governance.** Along with the US and China, countries such as Israel, Russia and Singapore have developed their own frameworks for harnessing the potential of this technology. Israel, for instance, is considered to be one of the nations that would benefit most from widespread adoption of AI, given its small population and high-skilled workforce (Horowitz et al., 2018). Singapore sits at the top of the "AI readiness Index" (Oxford Insights, 2019) and has developed the idea of a "smart nation", while Russia has focused especially on the military applications of AI. However, given that the US, China and the EU are widely considered as the three main blocs competing in the "AI race", the scope of this chapter will be limited to them. **In the next section, we will examine their different models, as well as the pros and cons of these.**

1.2 State-controlled development, free markets and a third way

The Chinese approach to AI is outlined in the "New Generation of Artificial Intelligence Development Plan" (State Council, 2017), which establishes a three-step strategy for the Asian country. Firstly, by 2020, the aim was to be a globally advanced AI player. By 2025, China's plan is to become the world leader in AI in certain sectors, using this technology as the main driving force for upgrading its economic and industrial capabilities. Finally, by 2030, China aspires to lead the world in terms of AI capabilities, becoming the main "innovation hub" and, consequently, the winner of the race.

Despite some references to the need for ethical codes, especially in the realm of automated weapons, China's strategy does not mention the word "discrimination". One of the main concerns regarding AI is the possibility of biased algorithms. The risk emerges from the fact that algorithms are developed by humans, who have either conscious or unconscious prejudices. It has been demonstrated that AI machines replicate these prejudices, and consequently they can create situations of unfairness, despite their presumed objectivity. For instance, a medical school in the UK tried to use AI to select which applicants would receive an interview invitation. Women and people from racial minorities had less of a chance of being interviewed, even if they had the same academic record as white male applicants (Hacker, 2018). However, the Chinese strategy dodges this issue and chooses to focus on the development of this technology for a "more

efficient economy". Along with this goal, it also aims to foster civic and military cooperation, increase public safety and build "smart cities".

Although the Chinese strategy mentions the need to collaborate with private industry, as well as the role of the market in the development of AI, **it features a clear "state-oriented" approach.** China aims to "build an international innovation base", recruit top talent and foster the expertise in AI of its current researchers, as well as "establish financial guidance" for capital investments. Consequently, the state would play a huge role in the development of the strategy, which could be beneficial in terms of the amount of money and resources available to invest[2]. **Conversely, there is a risk of increased state surveillance, as well as possible obstacles to innovation**, stemming from a reluctance to contradict the guidelines set forth by the CCP.

Following a very different tradition, the US model is outlined in its American AI initiative (Executive Order 13859, 2019), which mentions four different areas of interest in the use of AI: innovation, industry, American workers and American values. As the names of these areas suggest, the strategy is mostly inward-looking and does not represent a genuine attempt to create global rules. Recently, the US Office of Science and Technology Policy criticized the EU for its "over-regulation". In its statement, it condemned "heavy-handed innovation-killing models", arguing that they would hamper innovation, and encouraging "American allies" to adopt a similar model to that of the US. The remarks on "international leadership in AI" included in the American AI Initiative, as well as the desire of the White House to align its allies with respect to regulation, seem to be steps in the direction of a future global governance model for AI. However, essentially, the US idea of global regulation appears to consist of a worldwide extension of the US model.

The US strategy is focused on the idea of transparency and innovation, and oriented towards turning the US into the global hub of AI, as well as preventing the use and abuse of this technology by authoritarian regimes. **The main goal is to avoid any possible hindrance to innovation, while fostering trust in AI and promoting US economic growth**. This means that the US strategy is more market-oriented, for despite the increase in R&D investment by the Federal government and the call for increased cooperation among agencies, the lion's share of the investment still comes from private enterprises. In 2017, for instance, Amazon spent 16.1 billion dollars on R&D, while the National Science Foundation (NSF) and the Defence Advanced Research Projects Agency (DARPA) spent only 5.3 billion. Amazon, together with the NSF, was also the company in charge of the programme to foster fairness in AI.

[2] However, China's state investment in "national champions" does not imply that there is no competition among businesses. As Saz and Brosch (2020) argue, domestic competition is fierce in China. SMEs compete to develop new applications of AI, as well as more efficient forms of implementation, while trying to avoid being overtaken by their rivals (Lee, 2018). This produces a competitive process from the bottom up, before the state nominates the winner to become the national champion. Consequently, despite its state-oriented approach, the Chinese model does possess some clear market characteristics.

Consequently, the American model is skewed towards a market-oriented set of rules[3]**.** By arguing that over-regulation can hamper both developments in AI and the ability of business to grow and implement that technology, innovation and sparse regulation have become the trademarks of the Trump administration's strategy. Its American AI initiative argues that the US has always been a champion of freedom, human rights and**, especially, "the opportunity to all to pursue their dreams".** Consequently, the US regulation focuses on the market, businesses and individuals, as opposed to the more interventionist practices of other countries. However, a global governance framework which relies on voluntary ethical codes and little intervention by regulators will encounter several hurdles when seeking to fulfil all these ideals at the same time.

Finally, the EU has presented a model that aims to navigate between the waters of the free-market approach of the US and the more state-centred vision of China. The European Commission has repeatedly advocated an "ethical and trustworthy AI", focusing on the need to foster accountability, transparency and diversity, and to avoid biases in the implementation of this technology. These principles are in line with the Ethics Guidelines presented by the High-Level Expert Group on AI (European Commission, 2019), which recommended a lawful, transparent, robust, accountable and non-discriminatory AI, controlled at all times by humans. In this way, they argued, it will be "trustworthy" in the eyes of citizens and capable of spreading its benefits throughout society.

Thus, the EU seems to be trying to pave a third way, along the lines of the development of the welfare state in the 20th century, which offered an alternative to the economic systems of both the US and the USSR. The EU highlights the advantages of cooperation between the public and private sector, the need for human supervision, and the importance of unbiased data input and the democratic accountability of AI. These principles can provide the basis for a global governance framework that does not "over-regulate" and which also protects the fundamental values that have been enshrined in benchmark legislation such as the EU's General Data Protection Regulation (GDPR).

The White Paper on Artificial Intelligence (European Commission, 2020) is the first step towards the development of this model. It focuses on aspects that are only tangential to other strategies, such as gender discrimination in algorithms or stricter regulations in sectors deemed "high-risk". The EU claims that it wants to build an ecosystem of trust and excellence that would be regulated under a "complete" framework, but not an overly restrictive one, which could stifle growth and innovation. To sum up, the EU has taken "a regulatory and investment-oriented approach", capable of both harnessing the potentiality of AI and addressing the risks associated with it**. With this strategy, the Commission aims to produce legislation that can be used as a reference for future international treaties** and to build trust with allies, so that a multilateral dialogue on a future global framework may be conducted. In the following section, we will address the

[3] This is not to say that, in the US, the state is completely absent from the process of development and implementation of AI. For instance, the NSF has several grant programmes as well as partnerships with industry and academia to invest in technology and the training of future experts in this field. This follows Mariana Mazzucato's concept of "the entrepreneurial state" (2013).

problems of this multilateral perspective in a "confrontational race" and we will argue in favour of creating a set of global norms.

1.3. The concept of an "AI race" and the need for a global governance framework

There are a number of problems surrounding the reference to an "AI race". **The concept of a "race" usually evokes the notions of "weapons" or "arms"** embedded in the competition between the US and the USSR during the Cold War. However, the development of AI is poised to bring enormous benefits to our societies. It has been estimated that it could add $13 trillion to the global economic output by 2030 (Wladawsky-Berger, 2018). It will also help in the development of autonomous vehicles, smart houses and smart cities, and its applications include more efficient farming, better adaptation and mitigation measures against climate change, and improved healthcare**. Like many other disruptive technologies before it, such as electricity, AI has the potential to make a lasting and positive impact on our living conditions**.

But the risks must also be taken into account. For example, there are obvious military applications of AI (Barnes and Chin, 2018), such as recognition of enemies in crowds, more efficient fully automated weapons, or planes equipped with more sophisticated systems. **AI also provides tools for increasing the surveillance power of the state**, with enhanced techniques of facial recognition and internet control. Consequently, some authors have argued that the very concept of a "race" may be dangerous for the purposes of creating a sustainable global governance framework (Cave & OhÉigeartaigh, 2018). **Framing AI development as a competition can lead to a "winner takes it all" mentality**, which could damage the prospects of peaceful relations between the world superpowers in the future. More specifically, there is a risk of falling into a "Thucydides trap"[4] (Allison, 2017), with the current US-China tensions leading to an armed confrontation.

In a similar vein, some statements made by political leaders have fed this narrative of a zero-sum race where only one power can remain standing at the end. Russian President Vladimir Putin said that "**whoever becomes the leader of the AI sphere will become the ruler of the world**" (Gigova, 2017), and his French counterpart Emmanuel Macron argued that "as always, the winner takes it all in the AI field" (Thompson, 2018). As Cave and OhÉigeartaigh observe, there is a feeling of "survivalism", with powers panicking at the possibility of their rivals having military superiority and, consequently, a geopolitical advantage in any dispute. Other authors contend that by referring to a "race", the focus rests on military advances and not on the wide range of applications of this new technology (Gady, 2019). If these arguments were true, talk of a race could have a performative effect on international relations, prompting the superpowers to increase their R&D spending on military AI applications in order to gain an edge over their rivals, thereby increasing the likelihood of conflict.

[4] This theory, put forward by Graham Allison, states that the emergence of a new great power which threatens to displace the current superpower has a high probability of leading to war. In 12 out of the 16 historical episodes covered by his study, a conflict ensued.

Although there is merit in these theses, believing in a non-competitive development of AI would be tantamount to ignoring the dynamics of both the development and the implementation of similar technologies in the past. Furthermore, competition does not make conflict inevitable, and in fact there are good reasons to cherish a competitive environment. For example, it can foster innovation in the AI domain and bring about beneficial spill-over effects. **What we should do, instead of discarding the concept altogether, is to reframe the idea of a "race",** moving away from its military connotations towards a combination of competition and cooperation that does not result in conflict.

In this respect, a more interconnected world where superpowers are linked together in a web of economic, institutional and diplomatic relations (Nye, 2019) may be less prone to moving towards confrontation. **However, this stability should not be taken for granted, as these same superpowers can use asymmetric power relations in their favour**, coercing other countries through the use, for example, of economic sanctions. This **"weaponized interdependence"** (Farrell & Newman, 2019) presents a risk to the prospects of a non-conflictual development of AI, and consequently, all the actors in the race should be compelled to engage in a multilateral dialogue to find common ground for a global governance framework. **By constructing a network of principles, rules and institutions, the balance could be tilted towards cooperation and more benign forms of competition.**

Why should we build this set of global norms? In our view, there are **two main reasons**. Firstly, there is the need to foster stability as a key principle of the global order. Secondly, there are certain issues related to the development of AI that have transnational significance, with the potential to produce negative spill-over effects. As these issues cannot be effectively addressed through national action alone, and will require international coordination.

The first reason is related to **multilateral stability in international relations.** The development of AI may lead to the re-establishment of a world divided into blocs. By protecting "national champions" or introducing protectionist measures against foreign innovations, the race for the development and implementation of AI could easily turn into a confrontational one. It is not hard to imagine how a variety of regulatory standards for AI may lead to a greater pressure to overtake rivals in the race and adopt a more aggressive discourse. Ian Hogarth (2019) has called this phenomenon "**AI nationalism**", whereby all countries would avoid giving so much as an inch of advantage to their rivals. This, in turn, would threaten the stability of the international system, and the prospect of a confrontation would not be unthinkable.

Dafoe (2018), on the other hand, raises the possibility of these protectionist measures not being enough to guarantee the survival of the countries in the race. Their talent could still flee, "inevitably gravitating towards existing leaders in AI". If this is the case, there will not be the slightest semblance of a **level playing field in AI across countries**. Even though a global governance framework would not have the capacity to create perfectly fair competition, it must aim to give every country a chance to compete. This "fair competition" requires international standards to allow every country to implement AI at its own pace. An unregulated playing field would allow either power blocs or big companies to set their own "international" standards,

without respect for the will of the citizens of the affected states. As Harari has stated (2019), this could lead to a sort of **"data colonialism"**, with big companies or states obtaining access to the personal information of citizens from weaker countries. In turn, this would also foster tensions between competing powers, which could result in a more confrontational approach to international relations. Quoting Dafoe again, "norms involve a rough mutual understanding about what actions are unacceptable and what sanctions will be imposed in response". **In the absence of these rules, it would be impossible to hold the misconduct of others accountable**, threatening the stability of the system.

The second reason is related to the **need to counter some negative spill-over effects of AI development and preserve some basic public goods**. There are a series of issues that have a transnational impact, and if left unregulated, they could represent a grave threat to all states[5]. Firstly, and fundamentally, we have the development of lethal autonomous weapons. Unmanned Aerial Vehicles (UAVs), or "drones", have been widely adopted for civil and military purposes. According to the Bureau of Investigative Journalism (BIJ, 2019), there have been 430 strikes in Pakistan since January 2004. Between 2,515 and 4,026 people have been killed, including an estimated 500 to 1,000 civilians, many of whom were children. The European Parliament released a briefing (Latici, 2019) outlining the risks of unregulated military drones and discussing the main dangers associated with this new technology, such as breaches in citizens' private sphere and the possible invasion of a sovereign state's territory. However, there is no international regulation in place to hold drones "accountable", which forces states to **"apply international law written for the Second World War to Star Trek technology"** (Bakerjian, 2019). Horowitz, Kreps and Fuhrmann (2016) have discussed different aspects relating to the proliferation of drones in military arsenals. On the one hand, drones could have a significant impact on anti-terrorist operations (US President Barack Obama even referred to them as a "cure-all for terrorism")[6]. On the other hand, they serve a dual purpose in autocratic regimes, where there are fewer constraints on the actions of the executive**. Drones could be used for better surveillance of the population**, as well as to prevent a possible pro-democratic coup from a strong military (Greitens, 2017)[7]. Boyle (2013) believes that the use of drones may also have

[5] Dani Rodrik (2019) has convincingly argued that the idea that "multilateral agreements" always work better is more normative than empirical, and that excessive global regulation can hamper regional innovation, for example. Still, we believe that the current scenario fulfils Rodrik's criteria for supporting said regulation (Rodrik, 2020), namely, the existence of "beggar-thy-neighbour", zero-sum tendencies and/or the need to preserve global public goods (e.g., international security). Any global arrangements, however, need to be flexible enough to allow for regional experimentation on more comprehensive ways of regulating this technology.

[6] Most of the literature, however, does not support the hypothesis that targeted strikes are a "cure-all for terrorism". Firstly, because after the killing of a terrorist leader, another member of the organization, eager for blood and revenge, could take his place (Abrams & Mierau, 2017). Secondly, because the killing can create a "martyr", thereby unifying the terrorist group (Bob & Nepstad, 2007). Finally, because dangerous terrorist organizations usually rely on sophisticated internal structures and campaigns for popular support, so they can easily survive for a period of time, even without their leader (Jordan, 2014).

[7] This, in fact, would solve the autocrat's dilemma described by Sheena Greitens in her book *Dictators and their secret police: coercive institutions and State Violence* (2017). She argues that dictators have to choose between a strong and unified military or a more fragmented one. The former would be useful against contentious actions by

the secondary effect of "cheapening" the cost of war, thereby increasing the probability of armed conflicts, due to lower economic losses and fewer lives being placed at risk. Even if these risks were non-existent and drones were exclusively in the hands of democracies, the mere possibility of armed drones falling into the hands of terrorist groups should motivate us to try and find a common set of norms. **The goal would be, if not to restrict their use, at the very least to contain the most harmful effects of their proliferation**.

As a response to the abovementioned concerns, Buchanan and Keohane (2015) tried to develop what they called "a drone accountability regime". Given the mechanical character of UAVs, the high probability of collateral casualties in their attacks, and their ability to enter the sovereign territory of other states more easily than conventional armed forces, the authors consider it necessary to develop a set of rules with a clear idea of who would be accountable for their actions. Today, AI is mainly used to improve the capabilities of piloted drones, either by enhancing the resolution of their cameras, improving their targeting systems or raising their speed. However, there is a common concern in the literature about AI based on the use of this technology to develop fully autonomous drones. This would mean the lack of any human presence in the vehicle's actions, be this in piloting or supervisory tasks. Buchanan and Keohane argue that it is necessary to have a human either remotely piloting the drone or in charge of designing its action protocols before its mission. A fully automated drone, with no oversight, could kill a citizen, and without international instruments to hold somebody to account (the government or group giving the order, the designer, the programmer…), diplomatic conflicts could ensue. Therefore, in view of capability improvements through narrower forms of AI and the development of fully autonomous vehicles, the risks posed by drones will require an international framework and the establishment of appropriate legal procedures. Without them, it would be extremely difficult to find a peaceful solution to a targeted killing, and an escalation of hostilities would become more likely.

Secondly, the development of AI has also increased the risks and damage caused by cyber-attacks (Brundage et al., 2018). By enhancing the capabilities of aggressors and helping to create deep fakes, this technology could have a great impact on the capabilities of several countries. Those with limited defences would be especially vulnerable, and, without a global set of norms in place, institutionally unprepared to face and respond to these attacks. This issue also correlates with the drone problem, as cyber-attacks could result in remote control of military or civilian drones (Thornhill, 2020). The effects of these attacks are also felt in economies worldwide, and this represents a problem not only for the big powers, but also for countries in dire financial situations. **For example, in 2019, the US economy suffered losses of more than $1 trillion due to cyber-attacks**. In an international normative vacuum, there is no shared principle of defence

the population, but also dangerous for the regime, as a well-organized coup would be more plausible. The latter would be less effective when seeking to contain protests, but it would also be less prepared for a coup. The use of drones could provide a cheap source of surveillance for civilian leaders, as well as a tool to reduce the size of the armed forces and the possibility of a coup, while keeping military capabilities intact.

of cybernetic stability and security, nor are there **any bodies that a country can turn to in order to obtain reparations** (Nye, 2019).

Finally, the privacy of citizens, as well as possible discrimination by means of AI, may also be addressed by a global normative framework. Data constitutes the input that allows AI to "learn" and perform its tasks efficiently. The gathering and use of this information can lead to breaches in the rights of citizens by invading their personal sphere or using their personal data in ways that they have not consciously consented to. This is especially true for citizens of less advantaged countries, which could find themselves on the losing end of the phenomenon of "data colonialism" described earlier. Similarly, by reproducing the biases of their programmers, AI machines could replicate patterns of discrimination, which would be harmful for the rights of citizens of all countries, both at a national level (e.g., in recruitment processes) and in their stays in other nations (e.g., if accused of a crime).

As we have seen here, if we want to preserve the stability of the international system and protect citizens worldwide against the detrimental spill-over effects that AI could have on several fronts, it is necessary to establish a set of global rules. In the following section, we will take stock of what has been done at an international level to develop this global governance framework. We will also trace a roadmap towards a future set of norms that aspires to be realistic enough to be accepted by all states. **The overarching goal is to build a future in which AI works not only to the advantage of businesses and states, but also to the benefit of humankind.**

1.4 Building a global governance framework: a cooperative approach to AI

As in the realm of cybersecurity (Nye, 2011; 2017), both academia and the private sector have called for the adoption of a certain framework of norms at a global level. However, political developments have not been particularly encouraging, although some steps have been taken. At the 2018 meeting of the signatories of the Convention on Certain Conventional Weapons, Amnesty International reported that 26 states supported a total ban on fully automated lethal weapons, with many also calling for the development of new instruments of international law. In 2017, the office of the High Commissioner of Human Rights produced a report on the possible discriminatory effects of algorithms. In 2017 and 2018, the International Telecommunication Union (ITU), a specialized agency of the UN, organized the "AI for Good" Global Summits. Their goal was to put stakeholders in contact and foster the development of safe and inclusive technologies which would work to the benefit of all societies.

In 2019, the OECD put forward a series of "guiding principles for Artificial Intelligence" (OECD, 2019) to "stand together in unleashing AI innovation" (Murgia & Shrikanth, 2019). They admitted the "pervasive and transformative effects of AI on economic sectors and societies" and aimed to build a loose regulatory framework. Firstly, **by defining what AI is** ("a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations and decisions influencing real and virtual environments, with different degrees of autonomy"). Secondly, **by establishing recommendations, such as the avoidance of discriminatory algorithms, a regulation based on human rights and democratic values, the robustness and transparency of AI, and a regime of accountability**. Finally, by calling for cooperation in matters

relating to AI regulation and in the adoption of these principles. The 36 countries of the OECD signed the document, as well as Argentina, Brazil, Peru, Colombia, Costa Rica and Romania. Later that year, Malta and Ukraine also signed the principles. China, however, did not, and it is clear that **a set of global rules that does not include the Asian country will always be defective.**

Nevertheless, later in 2019, the G20 Digital Economy Ministers, as well as representatives from Senegal, Egypt, Singapore, Netherlands, Spain, Estonia, Vietnam and Nigeria, signed a statement on Trade and the Digital Economy (G20, 2019). This statement was focused on the benefits that digitalization and the widespread adoption of AI could bring to human societies, and it proposed a series of principles for a future global regulation, drawn from the recommendations of the OECD, while calling for cooperation between all stakeholders. **Despite its non-binding character, it represented a first step towards international coordination.** Countries with very different views, such as China and the US, agreed to sign a document that aimed to further an inclusive and "human-centred" AI.

These are promising, but incomplete advances. **If we are to tackle the broad challenges that AI presents, we cannot rely on the voluntary self-regulation of big corporations**, on strictly regional and national measures, or on non-binding documents (Floridi et al., 2018). As we have seen before, it is imperative to reframe the idea of a race towards a mix of competition and cooperation that does not drift towards conflict**. In order to achieve this, we must establish a set of international norms to regulate the development and implementation of this General-Purpose Technology.**

As Buchanan and Keohane argue (2015), it is necessary to distinguish between the "goals" and the "standards" of a global governance framework. If our goal is to produce a set of norms that would be accepted by all states, then we must lower the standards of those rules. **Norms, in fact, are interwoven in the social fabric, and if they are not considered to be legitimate, citizens or states are unlikely to comply with them** (Finnemore and Hollis, 2016). Consequently, the procedure of development and adoption of these norms is important, and an excessive focus on fulfilling too many objectives at the same time could reduce the legitimacy of these norms in countries that do not share the principles behind them.

The fact that we can only aspire to develop a relatively modest set of global rules is not necessarily a bad thing in itself, **as long as these rules enshrine certain principles that can keep the most pernicious uses of AI in check**. At the same time, a series of **more ambitious principles** could be developed for certain groups of democratic nations with shared institutional and cultural features, as it would be easier for them to reach a **"coalition of the willing"** (Baum, 2013) with higher standards and no problems of legitimacy. If this regulation were to be effective in helping these countries develop their AI capabilities within a framework of cooperation, it is plausible that other states would try to join the coalition, and so these extended principles would be adopted by more and more nations. We will begin our proposal with the **core principles**, those we believe could be adopted at a global level, and then we will define the more ambitious **extended principles.**

To begin with, a logical first step would be to **agree on a common definition of what AI is** (Walch, 2020). In this respect, the principles drawn up by the OECD provide a definition that could easily be accepted worldwide[8]. Secondly, a principle that could also be accepted with few reservations would be **the demand for human oversight of all operations involving Lethal Autonomous Weapons (LAWs).** Both Russia and the US have blocked any attempts to draft treaties on this matter (Brzozowski, 2019), fearing a possible ban on the production of these weapons. What is at stake here is ensuring that **there is always a human operator**, either in the control, programming, or determination of the targets of the weapon. Such a demand, which ensures a degree of accountability for those using these weapons, could gather unanimous support (Anderson & Waxman, 2017).

Thirdly, there would be a principle based **on transparency in the functioning of AI systems**, so that citizens from all countries would be able to understand both the process and the outcome of the decisions taken by the machine (Calo, 2018). A final principle would be **the need for robust AI,** subject to constant improvement and supervision, to ensure that no malfunctioning poses a risk to any human being. **All these principles would adhere to the guidelines of both the OECD and the G20, which have been accepted by countries with different institutional and cultural backgrounds**. Thus, although they would not be as ambitious as the extended principles, they would provide a framework for a global regulation of AI, with enough flexibility to accommodate different views (Whittlestone et al., 2019).

Moving on to the extended principles, we consider that all countries belonging to the "coalition of the willing" **would aim to preserve the life of their nationals, and hold accountable those who violate their citizens' right to live**. Given these premises, LAWs, as well as any form of AI used for military purposes, should always be under human control (Horowitz & Scharre, 2015). However, the aim is to go beyond the core principles. In these extended principles, either the programmer or the controller of the weapon, depending on each case, should be subject to judgement on the basis of the actions of the machine. Governments may also be held responsible for actions conducted on their behalf, and **a system of reparations could be broadly accepted.**

Secondly, data is widely considered as the "fuel" of AI development, and it has become a disputed resource. At the same time, despite differences in the perceptions of what constitutes the public and the private sphere**, these states would grant their citizens a certain measure of privacy.** Consequently, there could be a broad agreement to protect the privacy of citizens in certain aspects of their lives. **The GDPR and the European Digital Strategy could be the foundations for future regulations in this area**. The coalition supporting these principles should be willing to admit that data management that goes unchecked by governments and big companies could end up posing risks for democracies, and they should act accordingly.

[8] An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy.

Finally, **given the global spread of AI, an agreement could be made that recognizes the need to fight against discrimination in the development of algorithms**. Firstly, at a national level, groups that have already been marginalized, such as women or racial minorities[9], could face even more discrimination on the basis of biased algorithms. Secondly, at an international level, AI could perpetuate discrimination against citizens from different countries, as they could be banned from applying for residency or work permits in certain countries almost automatically. In line with this proposition, it would be accepted that **there needs to be transparency with regard to the premises on which the AI acts, who the designer is, and for what purpose it has been conceived.** These three principles could easily be accepted by all democratic nations, as they are built upon the work that has already been undertaken by the EU, the OECD and the G20. They would represent a step forward in developing "binding" instruments, instead of mere recommendations, on a scale that goes beyond the regional.

We have refrained from crafting a proposal that aims for a more "informal" set of norms, which do not include treaties, but the generation of trust among states (such as the proposal by Buchanan and Keohane). **Nor have we considered proposals that aim for the creation of "global oversight bodies"**, or a truly new international institutional framework (Urs & Almeida, 2017). Instead, we have called for the creation of a framework of norms that could be accepted by all states, including non-democratic ones, in order to ensure their legitimacy, protect the stability of the international system and offer answers to some of the most pressing issues that AI poses. At the same time, we have advocated the creation of a set of second-order principles for states willing to go a step further in the regulation of AI.

1.5 Setting the rules of the game: the EU as a leader of the global governance framework

The EU enjoys a privileged position in leading the development of these instruments of global governance. The Union has been built on the basis of negotiations, multilateralism and the search for consensus. Throughout its existence, it has dealt with the need to broker between competing visions, in order to extend the territorial and functional scope of the Union. **As a result, it has a great deal of experience when it comes to setting the rules of the game at an international level**, and it can use this as an advantage (Bradford, 2012).

The EU is predominantly associated with two types of power. **Firstly, it is a source of regulatory power**, which consists of producing legislation that is later used as a benchmark for other countries, such as the GDPR. For instance, Canada, Australia and the State of California have adopted privacy laws that were heavily influenced by the EU legislation (Sher-Jan, 2019). **Secondly, it is a source of soft power,** which is the ability to influence the attitudes and values of neighbours and allies. By combining these two powers, the EU could ensure that certain values such as accountability or transparency are enshrined in future treaties underpinning a global governance framework. Its "third-way" approach to AI regulation can also turn the EU into a

[9] As explained in Section 1.2. For more examples of the discrimination bias in AI, see Zou & Schiebinger (2018) and Leavy (2018).

broker between the demands of countries with a more statist approach and those which prefer to put fewer constraints on the activities of business, allowing the market to self-regulate.

Although there are experts who argue that the EU cannot act exclusively as a referee in the development of AI (Wolff, 2020), **the cooperative approach outlined in these pages essentially views the referee as even more powerful than any of the players**. The EU has more experience and expertise in terms of supranational legal frameworks, and consequently it enjoys an advantage when it comes to dictating which principles are to be incorporated into the future governance framework. **By implementing certain standards at a global level, the EU would effectively "win" the race**.

However, if the EU is to use its two powers to lead the way in the adoption of this global framework to good effect, **it will also need to underpin the material bases of its legitimacy**. If it keeps falling behind the US and China in terms of innovation, computing power, research papers and number of successful AI-focused start-ups, the EU will find it difficult to claim the legitimacy needed to propose a global governance framework. **Therefore, all diplomatic efforts must be accompanied by an attempt to retain talent in enterprises and universities** (Horowitz et al., 2018), solve the problems of data fragmentation inside the EU, and invest more and in smarter ways to promote the adoption of AI (Franke, 2019). Only by establishing itself as a true AI power will the EU be able to lead the way in setting the global norms for this new and disruptive technology.

The construction of this global governance framework is important for the future. **In the current climate of competition between superpowers, it is necessary for the EU to embrace a role of broker and rule setter**. By fostering cooperative competition, in order to avoid confrontation, the EU could ensure that some of the principles that it put forward in its "Ethics Guidelines" are fulfilled at an international level. The misuse of AI by governments could lead to more repressive police states and abuses of power by Big Tech, which would go against the aspirations of the EU. Instead, by ensuring the creation of this set of norms, AI can become a useful tool for citizens to hold their representatives accountable and attain new and better standards of living**. It is in the hands of all actors to push for these goals and create cooperative dynamics from which each of them would benefit.**

CHAPTER II
AI adoption in Europe

Artificial intelligence (AI) and its applications could reshape the current economic landscape by fostering productivity gains, improving efficiency, and lowering a variety of costs. The range of benefits is vast, from enhancing decision making to automate completely manufacturing procedures. AI, however, is still in its infancy, and many challenges have to be addressed to promote innovation and inclusive growth.

2.1 European AI landscape

The United States and China hold the lion's share for the majority of AI advance indicators, while Europe's efforts to catch up seem insufficient. **Europe is lagging behind the United States with a digital gap of 35% without a decreasing trend** (Bughin et al., 2019)[10]. More mature companies, in the context of digitization, possess a comparative advantage in AI adoption. However, only two European firms are in the worldwide top 30 digitized companies, and at the end of 2017, none of the ten largest internet companies worldwide was based in Europe (Bughin et al., 2019).

According to McKinsey's 2018 Digital Survey, almost 80% of European enterprises have experienced a positive return from their digital transformation. Still, only 50% realize a return more prominent than their weighted cost of capital. In addition, they insist that the digital gap between Western Europe and the United States remains constant, and it has two main dimensions in the use of digital technologies and the level of diffusion. Specifically, **European companies are less mature not only in their state of diffusion of digital technologies but also in the use of those technologies for new services and business models**. Except for the category of smart robotics, Europe's AI diffusion lags the United States thus far in the types examined (Big data, Advanced neuronal algorithms, and AI tools). Less than 50% of the European firms have implemented one AI technology, and most of them are in the pilot stage (Bughin et al., 2019).

World Intellectual Property Office (WIPO) released in 2019 an insightful report relating to the AI and its technological trends (WIPO, 2019). **Patent and scientific publication statistics indicate that China and the United States lead the AI technologies adoption worldwide. The European**

[10] The digital potential of each country is defined as weighted deployment of digital assets, labor, and practices across all sectors, compared with the most digitised sector. According to McKinsey's Global Institute 2016 research, European countries were capturing only 12% of their full digital potential, two-thirds of the captured potential in the United States (18%).

presence in several rankings is predominated by Germany alone. Starting from AI techniques, Germany is the only European country that appears amongst the top five states for all the examined categories (Probabilistic reasoning, Ontology engineering, Fuzzy logic, Logic programming, and Machine learning) regarding the patent filings. Nevertheless, **from the perspective of the scientific publication, the scene changes, and the United Kingdom, France and Italy appear in the top five countries**. The United Kingdom is in the first five countries for every tested ranking, France stands out in Logic programming and Italy in the Ontology engineering sector. Similarly, in patenting of AI functional applications, Germany is the only European country that is present in the top five countries ranking internationally. Germany and the United Kingdom hold prominent positions in publications for all the examined groups, France appears fifth in Distributed AI and Knowledge representation and reasoning, and Spain is fifth in Predictive analytics. Europe presents a prolific academic sector as the rate of scientific publications shows. However, the commercialisation of inventions, as suggested by the relatively small number of patents, is limited. A possible explanation of this phenomenon could be attributed to the weaker links between the academic institutions and the private sector, especially in the Southern part.

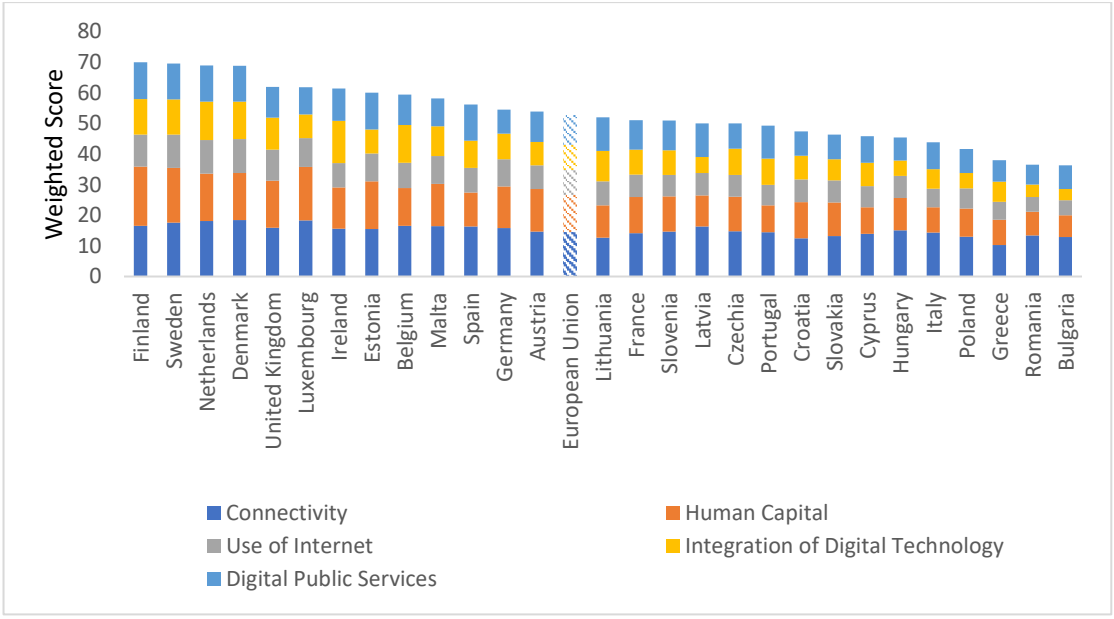
Bughin et al., 2019 developed and analysed a set of indicators by country to estimate their readiness to adopt AI technologies and applications. The AI Readiness Index scores arise from weighting AI enablers according to their relative importance for fostering the economic growth of each country. The components of the index included, among others, the existence of AI startups, the level of automation, Digital readiness indexes, Innovation indicators, Investment capacity, Human capital skills, and ICT connectedness. The report indicated a clear gap in AI readiness in Europe, with Southern and Eastern Europe lagging critically behind. **The United Kingdom, Sweden, Finland, and Ireland seem more ready to welcome an AI transition** since they ranked in the top 25% in AI Readiness Index for at least four categories of the index (eight in total). On the other hand, **Spain, Portugal, Italy, and Greece did not achieve this percentile in any of the examined categories**. Traditionally economies in the South rely more on the tourism sector for fostering economic growth than the advance of the technological or the industrial sectors. Furthermore, SMEs, the dominant firm type in the South, are facing resource shortages in both financial and human capital dimensions. Thus, investing in appropriate infrastructure and hiring ICT experts could be prohibitive. These facts could serve as a perspective for the differences in AI readiness for the two parts of the continent.

The European Union (EU) accounted for 8% of the global AI equity investment in 2017 (OECD, 2019a). However, AI is adopted progressively unevenly in Europe. In terms of investment levels, startups in the United Kingdom received 55% of the EU total investment from 2011 to mid-2018, followed by Germany (14%) and France (13%), when the remaining 25 countries appropriated less than 20% of all private AI equity (OECD, 2019a). **This evidence might suggest that firms in Southern Europe cannot fully appropriate the benefits that AI could offer to their operation and services.** In addition, they might lack the capability to absorb relevant funding, or they cannot attract the appropriate human capital with the proper expertise.

Europe remains behind the United States and China with respect to digital- and AI-based ICT with a share of ICT sector equal to 1.7% of GDP when the portion of China is 2.2% and 3.3% in the United States (Bughin et al., 2019). The top five contributors to ICT sector value added in the EU for 2016 were the EU's five largest economies (Germany, the United Kingdom, France, Italy, and Spain) with their contribution range starting from 20% for Germany to 6%, which is the share of Spain. Together, these five countries accounted for 69% of the total EU ICT sector value added in 2016. **As a proportion of GDP, the ICT sector share was more significant in Malta (6.2%), Sweden (5.8%), Finland (5.5%), and Luxembourg (5.4%). On the other hand, Spain, Italy, Greece, and Portugal ranked below the EU average**. Similarly, those countries were the five most prominent employers in the EU ICT sector in 2016. Still, when the share of ICT to the total employment is considered Southern European Countries are lagging behind. In particular, Portugal and Greece ranked last among EU Member Stated in the 2016, with a percentage inferior to 2% (European Commission, 2019).

Digital competitiveness is paramount to embrace AI technology development and applications. **The Digital Economy and Society Index (DESI) for 2019 signals an asymmetry in EUs digital advance**. DESI is a composite index that aims to track Europe's digital performance in five key dimensions, Connectivity, Human capital, Use of Internet, Integration of digital technology, and Digital public services. Figure 1: DESI index 2019, European countries ranking indicates that **Scandinavian countries outperformed among European members**, and the United Kingdom, Luxembourg, Ireland, Estonia, and Belgium managed to rank among the top nations of the total index. **On the contrary, the last three positions were allocated to Bulgaria, Romania, and Greece, though most of the Southern – Eastern European countries ranked below the EU average**.

Figure 1: DESI index 2019, European countries ranking



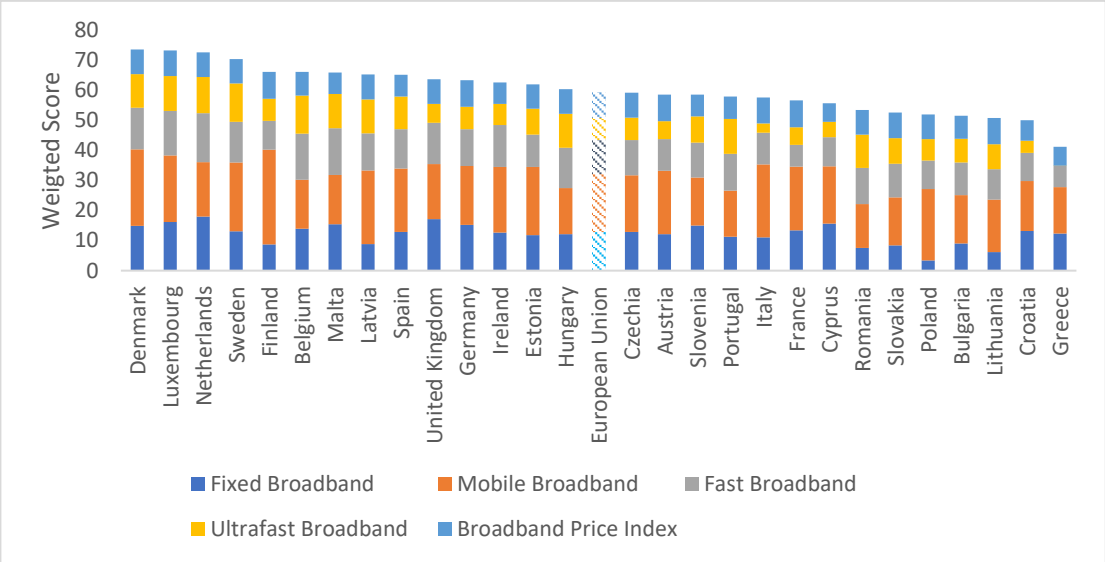
Source: DESI 2019, European Commission

Notes: DESI overall index is calculated as the weighted average of the five main DESI dimensions: 1 Connectivity (25%), 2 Human Capital (25%), 3 Use of Internet (15%), 4 Integration of Digital Technology (20%) and 5 Digital Public Services (15%)

Connectivity in Europe

Connectivity aspect examines the European performance relating to the availability of fixed broadband, mobile broadband, fast and ultrafast broadband, and the range of their prices. Connectedness is critical for a firm or a public service to transit to an AI model. Netherlands and Luxembourg are the best performers to the fixed broadband categories (basic, fast and ultrafast) while Greece, Poland, and Croatia rank last. Finland, Denmark, Latvia and Italy lead Europe relating to mobile broadband connectivity, when Romania and Hungary were the worst performers. Broadband advance has a tendency, with a relatively small correlation, to be lower in states where the cost of access is a greater part of the income. Fixed broadband is most affordable in Finland, France, Germany, and Austria. On the contrary, **Slovakia, Greece, Cyprus, Portugal, Spain, and Bulgaria have no offer belonging to the inexpensive cluster of the Broadband Price Index**. Broadband infrastructure quality and pricing are key concerns for Eastern Europe.

Figure 2: DESI index 2019, Connectivity by Sub-dimensions, European countries ranking



Source: DESI 2019, European Commission
Notes: DESI Connectivity Dimension is calculated as the weighted average of the five sub-dimensions: 1a Fixed Broadband (18.5%), 1b Mobile Broadband (35%), 1c Fast broadband (18.5%), 1d Ultrafast broadband (18.5%) and 1e Broadband price index (9.5%)

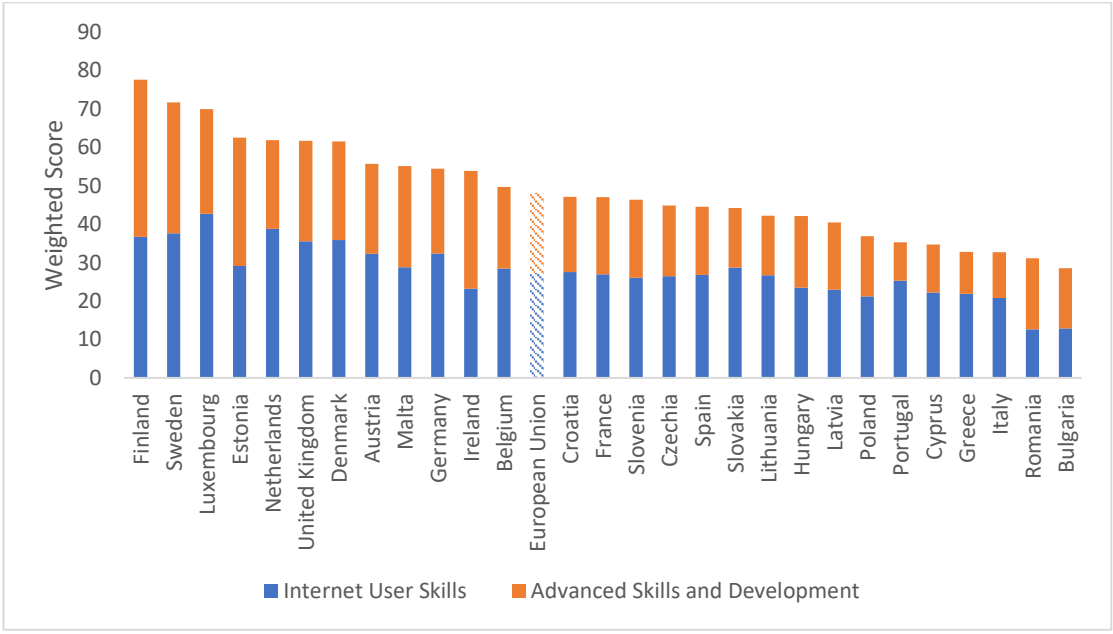
Positioning in the last places of the connectivity category could be a double dimension issue. On the one hand, countries might lack the principle infrastructure to support fast connections. On the other hand, pricing is a key issue, and potential exploitation of increased market power in telecommunication markets should be examined further.

Human Capital Skills

AI adoption is affected considerably by the complementary human skills as AI technology develops. Bughin et al., 2019 find that two critical barriers to AI embracing are the capability to practice ICT tools in work and the ability to provide new AI applications and services, such as AI coding and analytic expertise. The same report states **that Europe is already in supply shortage of the advanced skills needed for AI transformation**. The three categories of skills examined are Advanced cognitive/technology skills, Social and emotional knowledge, and Basic production and processing knowledge. Northern European countries, such as Scandinavian countries, the United Kingdom, and Germany, which are more progressed in the AI adoption, are those that face the higher excess demand without the supply being able to meet it.

The lack of skills is also suggested by the Human capital dimension of DESI 2019, which consists of two sub-categories, internet user skills and advanced skills and development. The first sub-dimension is estimated, utilizing the number and complexity of activities involving the use of digital devices and/or the internet. Then, the dimension of advanced skills and development is comprised of indicators on ICT specialist employment and ICT graduates. For 2019, Finland, Sweden, and Luxemburg are the best performers for this dimension, whereas Bulgaria, Romania, Italy, and Greece rank to the lowest positions. **Apart from Malta, the southern European countries listed below the EU average, indicating a severe insufficiency in digital skills progress**.

Figure 3: DESI index 2019, Human Capital by Sub-dimensions, European countries ranking



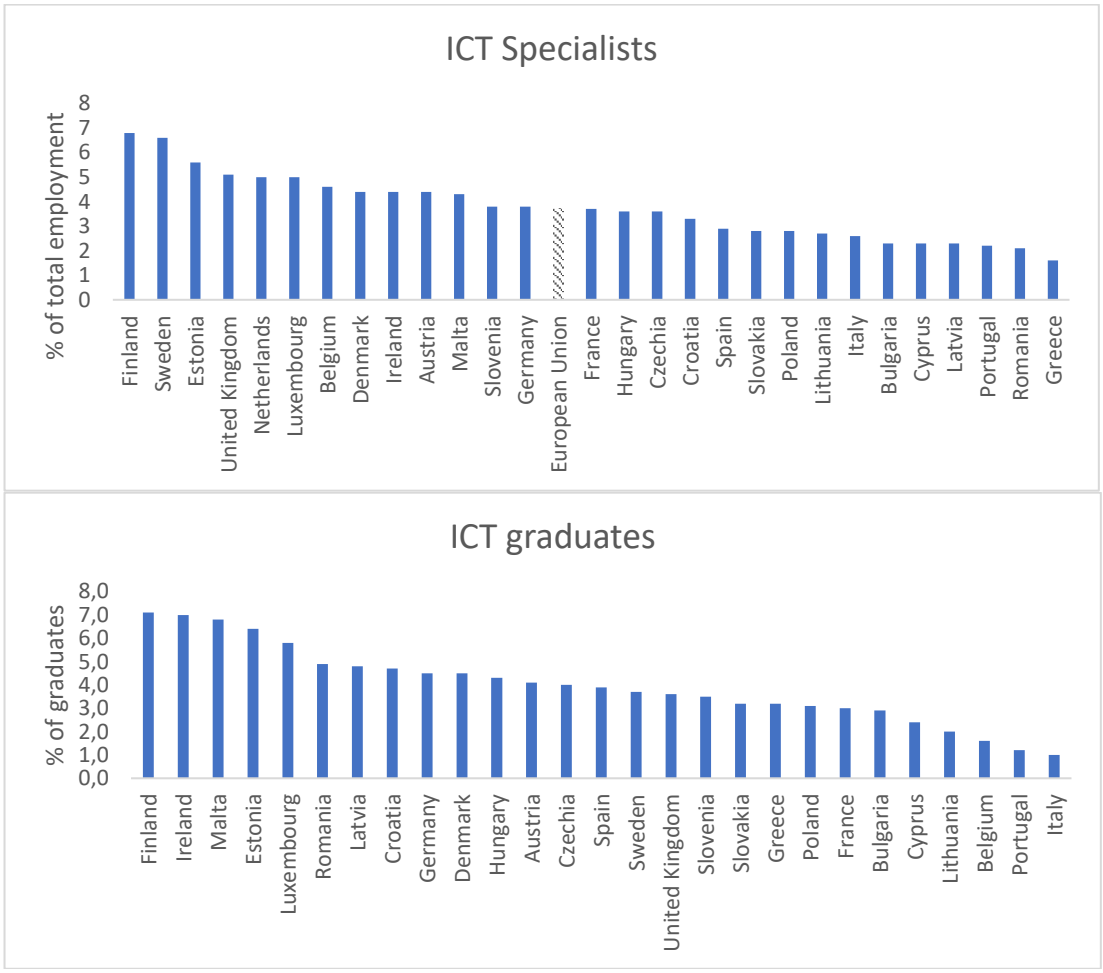
Source: DESI 2019, European Commission
Notes: DESI Human Capital Dimension calculated as the weighted average of the two sub-dimensions: 2a Internet User Skills (50%) and 2b Advanced Skills and Development (50%)

In 2017, 43% of EU citizens had an insufficient level of digital skills. These figures imply a potential digital exclusion of people belonging in this percentage in a case of increased digitization. This issue also has a geographical dimension since the EU Member states differ significantly in the context of digital skills. The share of the population with at least basic digital skills in Luxembourg and the Netherlands is equal to 85% and 79% accordingly, while in Portugal and Greece, approximately half of their citizens are digitally illiterate. Lack of even the basic digital skills required by the job market is still apparent in the EU since 10% of the EU active labour force (employed and unemployed) has no digital abilities, and 35% does not have at least basic digital skills. The vast disparity among EU states also exists in this indicator. The proportion of the labour force that has no digital skills or does not use the internet at all is ranging from 26% in Romania, 25% in Bulgaria, and 18% in Portugal, while in countries like Denmark and Sweden, the percentage is almost negligible.

Germany, the United Kingdom, France, Italy, and Spain were the five most prominent employers in the EU ICT sector in 2016, but Greece presented the lowest share to total employment (1.5%). Furthermore, DESI's Use of the internet services dimension indicates a similar landscape since Romania, Bulgaria, and Greece position last. However, the UK, Italy, Croatia, and Greece, are making significant progress compared to their performance during 2018. In the context of this dimension, e-Commerce usage varies significantly among European countries. **Italy, Spain, Greece, and Cyprus presented fewer online transactions than the EU average while Denmark, Netherland, and Sweden lead the ranking**.

The crucial necessity to transform the current landscape to an AI friendly environment is the availability of ICT skilled labour force. ICT specialists account for 3.7% of total European employment. However, the variation among the Member States has a wide range. **Top countries, concerning the share of ICT experts to the total employment, in EU, are Finland (6.8%) and Sweden (6.6%), whereas the same proportion for the majority of the Southern European countries is less than 3%**. Notably, Greece ranks last with a percentage equal to 1.6%. ICT graduates signal the near future characteristics of the labour market. From the total of graduates, approximately 7% belong to the ICT segment in Finland and Ireland. On the other hand, Portugal and Italy are slightly above 1%, indicating a potential lack of future experts in this field. Prospects are better for Greece, and Spain was the ratio comes to 3.2% and 3.9%, respectively.

Figure 4 DESI 2019 key ICT Individual Indicators, European countries ranking



Source: DESI 2019, European Commission

Corporate AI performance

Taking on AI requires a high level of business digitization. Finland and Denmark are unique examples in the EU possessing more than 10% of highly digitized firms (percentage to total enterprises) according to the Digital Intensity Index (DII), followed by Sweden with 8%. The rest of Europe present statistics are not encouraging, relating to a smooth transition to an AI era. Especially **in Bulgaria, Greece, Spain, and Italy, the majority of firms (over 55%) disposed of low investments in digital technologies**.

Likewise, the DESI index suggested a **vast gap among European countries regarding the Integration of Digital Technology**. The main drivers of Digital transformation in European firms are the fast broadband connections, social media, and mobile applications. In Figure 5: DESI index 2019, Integration of Digital Technology by Sub-dimensions, European countries ranking**, Ireland, Netherlands, Belgium, and Scandinavian countries achieve the highest positions, whereas Cyprus, Greece, and Italy rank below the EU average**. Considering e-commerce, Ireland, Sweden, and Denmark are the top three countries of the indicator and Spain, Italy, Cyprus, and Greece did not manage to rank above the average point of the EU. E-business technologies appear to be more dominant in the Netherlands, Finland, and Belgium. Spain and Cyprus are positioning above the EU average. However, there is much space to improve to catch up with European leaders. On the other hand, Portugal, Italy, and Greece lag behind in the adoption of business digitization.

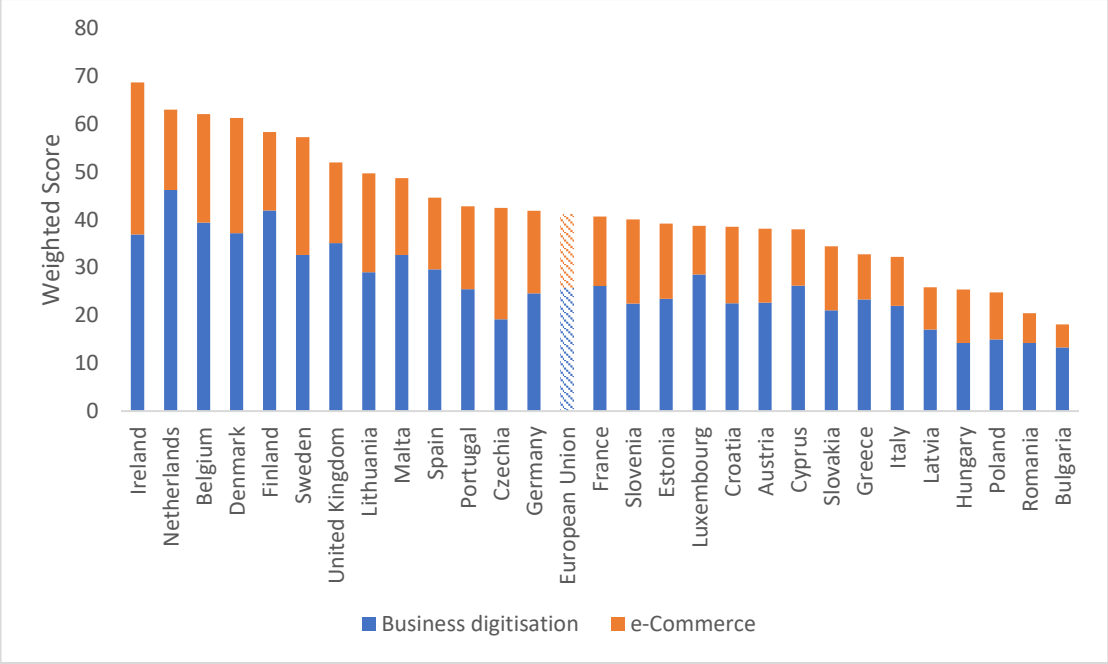
By Defining as an "AI company" "the one that produces a primary product or service utilizing machine learning, deep learning, image recognition, natural language processing, or other frontier AI technology", Berger (2018) aimed to provide systematic evidence relating to AI ecosystem using international data from startups. According to the findings of this study, the majority of AI startups are concentrated in the United States (1,393 enterprises), China is following (383 companies), and the total European figure is equal to 769 new ventures. **The dominant European countries are the United Kingdom, France and Germany (245, 109 and 106 new firms respectively) followed by Sweden (55), Finland (45), Spain (39), Switzerland (28), Netherlands (26) and Italy (22)**. European regional AI hubs in extended urban areas can be found in London, in Paris, in Berlin and Helsinki, as the report indicates. Complementary, Bughin et al., 2019 estimate that the share of Europe to AI startups is about 25%, with the early-stage investment in AI to be less compared to the United States and China.

Large companies are more digitized compared to SMEs since they rank higher for specific vital indicators relating to tracking digitization performance. **For 2018, 90% of large European firms could access ICT specialists, whereas the equivalent percentage for SMEs was equal to 65%**. Large enterprises possess a scale advantage, making it easier for them to employ internal ICT specialists (75% of big firms in EU) while the percentage of SMEs employing ICT specialists decreased marginally (1.6%) from 2012 to 2016 (slightly above 10% and 40% for small and medium companies respectively). However, it is worthy of mentioning that the share of SMEs, which hires external personnel to carry out ICT tasks, boosted at 1.7% (European Commission, 2019).

The proportion of big companies with more than half of the personnel employed using computers and the Internet in 2018 was equal to 52% and for SMEs 43%, increased by 2% and 3% respectively from the last year available. Interesting is the share of firms with the Fastest broadband connection existing in each area (at least 30 Mb/s), which accounts for 75% of EU large enterprises and 43% for the small ones. Buying medium-high Cloud Computing services is usual for 39% of the big European firms, but only 17% of SMEs utilizes such services. Digital transformation for SMEs is at a slower pace, although, even for big players, there is plenty of room for improvement for most of its aspects. Regarding big data computation and analysis, 33%

of large firms appropriate big data techniques while the share of SMEs is approximately 12%. The larger share of SMEs to total businesses in the South compared to Northern Europe could partially explain this lag. SMEs are not able to follow digital trends due to limited resources in terms of financial and human capital. The current entrepreneurial ecosystem, especially in the South of Europe, is characterised by inadequate venture capital, lower levels of entrepreneurial risk and insufficient collaboration between the research institutes and the firms. Taken together, fostering an innovative, productive private sector with respect to digital transformation and AI adoption can be a challenging task requiring effective policy interventions to improve the current governance framework, the legislation and the competitiveness of the firms.

Figure 5: DESI index 2019, Integration of Digital Technology by Sub-dimensions, European countries ranking



Source: DESI 2019, European Commission

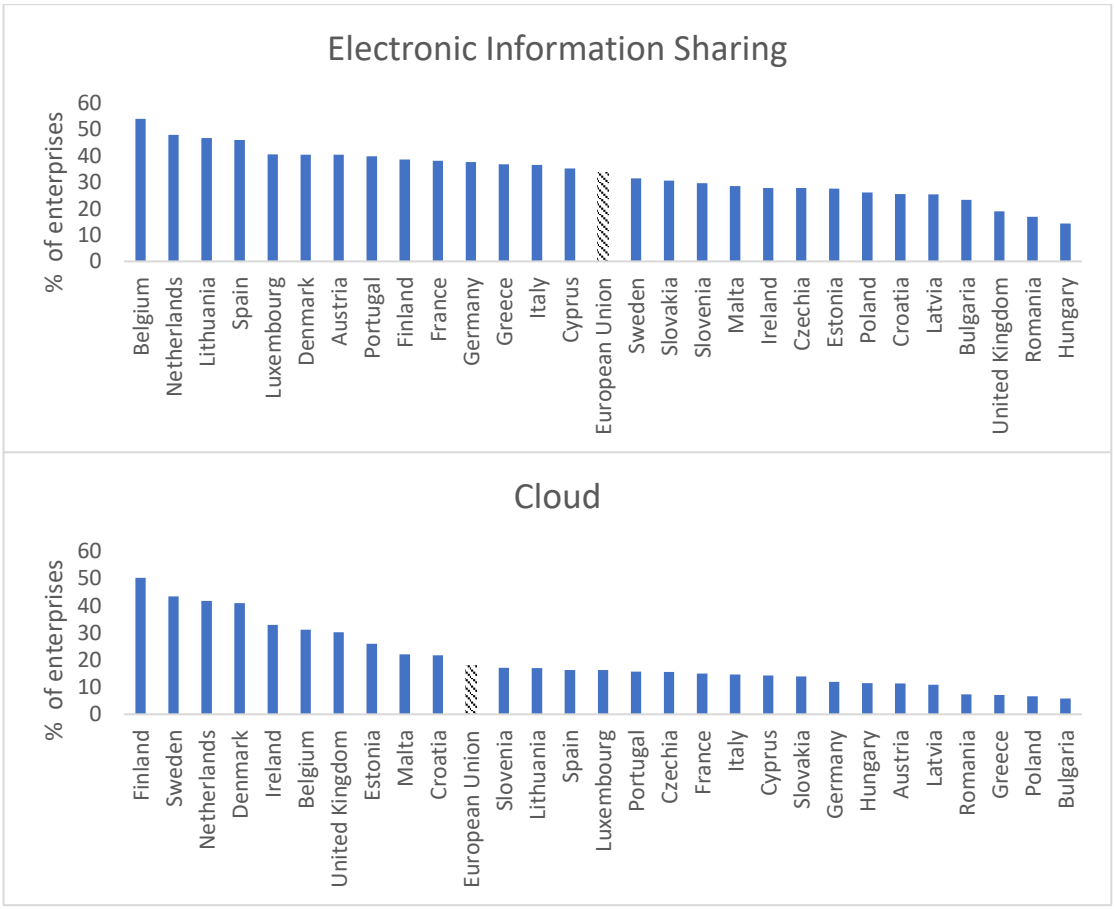
Notes: DESI Integration of Digital Technology Dimension is calculated as the weighted average of the two sub-dimensions: 4a Business digitization (60%) and 4b e-Commerce (40%).

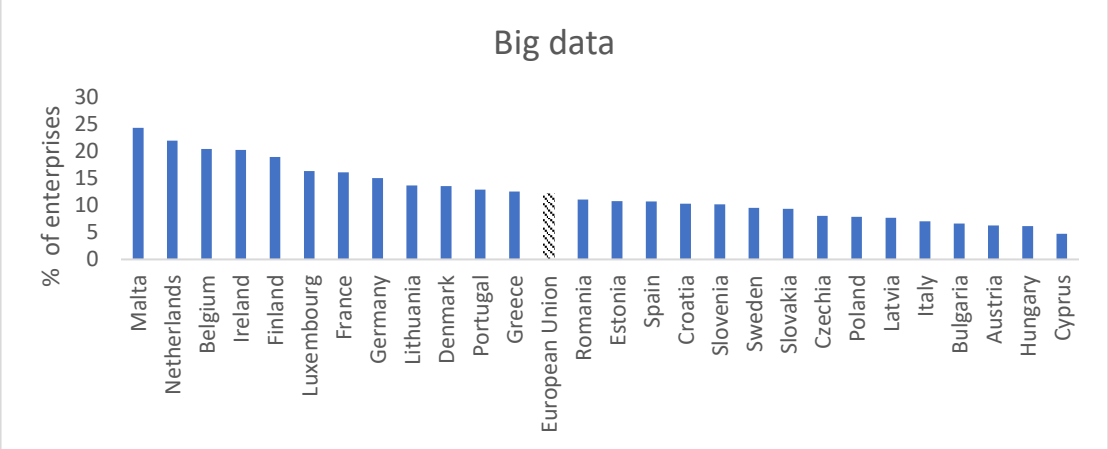
Elaborating further, businesses utilizing an ERP (Enterprise Resource Planning) software package to share information between different functional fields such as the accounting, the marketing, the production, and the strategic planning, are more than the half of total enterprises only in Belgium. Spain (46%), Greece (36.8%), Italy (36.5%) and Cyprus (35.3%) ranked above the EU average (33.8%). Portugal and Greece are slightly higher from the EU average. Spain, Italy, and Cyprus are behind the mean of the EU. The second graph in Figure 6: DESI index 2019, Selection of indicators of business digitization key to AI adoption, European countries ranking depicts businesses that purchase at least one cloud computing service (hosting of the enterprise's database, accounting software applications, CRM software, computing power). A great regional

disparity is present since over 40% of firms located in Scandinavian countries and the Netherlands are using at least one kind of cloud service. In contrast, in Southern European countries, excluding Malta, less than 17.8% of firms are using such products.

Big data collection, storing, and analysis is paramount for the enabling of an AI business model. The 12% of EU companies, on average, are utilizing big data for producing near or real-time results from data that comes in a variety of format types to enhance decision-making and in other business processes. Malta ranks first, among European countries, regarding the share of firms that use big data (24.4%). The Netherlands, Belgium, and Ireland come next, with more than 19% of enterprises taking advantage of big data. Portugal and Greece are near to the EU average, but Spain, Italy, and Cyprus are lagging.

Figure 6: DESI index 2019, Selection of indicators of business digitization key to AI adoption, European countries ranking





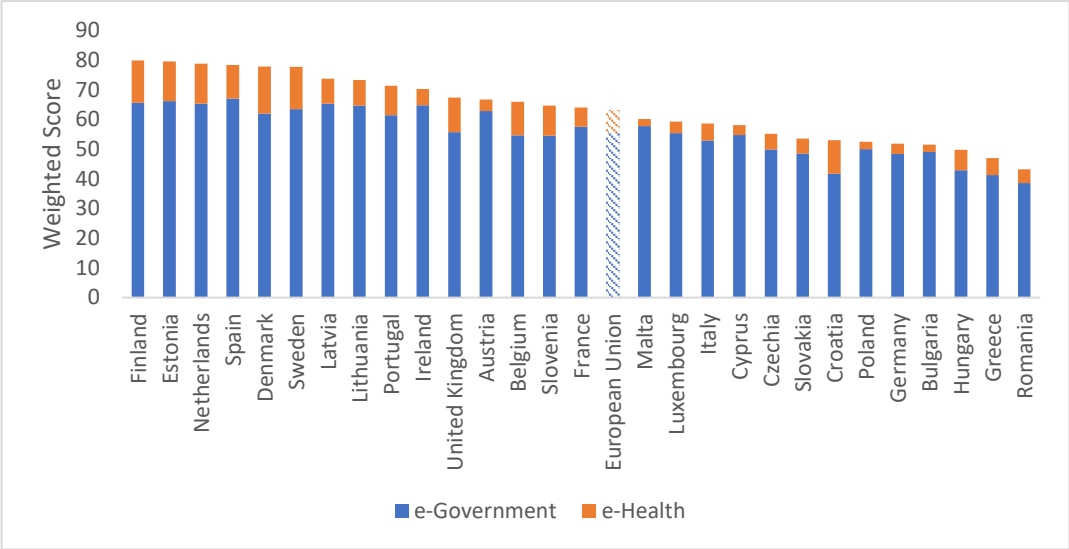
Source: DESI 2019, European Commission

A great variation in digitization is also present at the sectoral level. Economic sectors are advancing following different rates, which could be attributed to the diverse needs and fundamentals of each industry. The segments of the ICT sector, unsurprisingly, have a tendency to be more digitized relative to the other areas. However, travel agencies and the media sectors are also key players in business digitization. Segments of the European economy that lag behind in the digital era belong mainly to the Manufacturing sub-categories such as Construction, and Transport and storage have a relatively low share of firms with high or very high digital intensity index. The dispersion of the DII by economic activity is relatively smooth across EU countries.

AI and the Public sector

Public sector digitized services could be complemented with AI applications and encourage the adoption of AI technologies to the private sector as well. At the moment, 64% of EU citizens used public services online. The highest score in this dimension of the DESI index was achieved by Finland, followed by Estonia and the Netherlands. Spain and Portugal ranked higher than the EU average score, while Italy, Cyprus, and Greece scores have room for improvement. Regarding e-Services demand by Europeans, in terms of usage, Sweden, Estonia, Finland, and Denmark are the leading countries of EU, with more than 90% of internet users (aged 16-74), who have to submit filled forms to the public administration, choosing governmental portals. On the other hand, in Italy and Greece, the percentage was less than 40%.

Figure 7: DESI index 2019, Digital Public Services by Sub-dimensions, European countries ranking

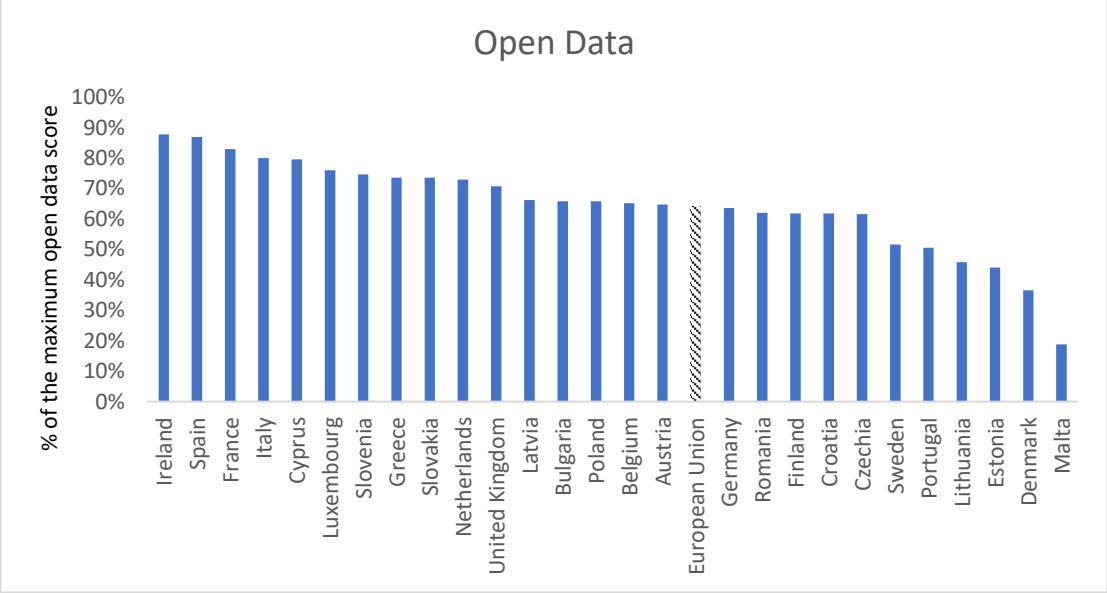


Source: DESI 2019, European Commission

Notes: DESI Digital Public Services Dimension calculated as the weighted average of two sub-dimensions: 5a e-Government (80%) and 5b e-Health (20%)

The development of AI technologies demands the availability of high-quality data. Open data policies could assist in this effort significantly. DESI index for 2019 is trying to estimate the level of maturity of open data taking into consideration four dimensions. The existence of a national standard of policies on open data and other coordinated approaches towards data publication is the first one. Other key factors are the availability of open data portals, the quality of the accessible data, and their impact on policymaking at four extents: political, social, environmental, and economical. When the level of maturity increases, countries also aim to increase the quality of the data available and to maximize the potential impact on decision-making. Best performers, according to the most recent statistics, are Ireland, Spain, and France, having scored more than 80% at the relevant indicator of DESI 2019. Conversely, Malta, Denmark, Estonia, and Lithuania ranked last, having scored less than 50%.

Figure 8 DESI index 2019, Open Data indicator sub-category of, Digital Public Services dimension, European countries ranking



Source: DESI 2019, European Commission

R&D expenditure and AI

R&D expenditure by business enterprises increased faster in the ICT sector than the overall economy in 2016, though ICT R&D intensity equalled 5% in the EU, significantly behind the US and Japan. Examining its allocation in Europe, the aforementioned factor **was significant in the four largest economies in the EU (France, Germany, the United Kingdom, and Italy)**, followed by Sweden and the Netherlands. Together, these six accounted for 74% of total R&D expenditure by business enterprises in the ICT sector, which indicates an acute insufficiency that the remaining countries have to overcome.

Moreover, **R&D personnel in the ICT sector accounted for 19% of total R&D personnel in 2016**. Once again, in absolute numbers, the top employers of R&D personnel in the ICT sector in 2016 were Germany (54,000 full-time equivalents - FTEs), France (FTEs) the United Kingdom (38,000 FTEs), and Italy (30,000). In total, those four states represented 56% of total R&D personnel in the ICT sector in 2016. Surprisingly, **when the proportion of the ICT sector to total R&D Personnel (PERD) is taken into account, Greece (37%) and Cyprus (36%) were among the top countries with the highest concentration of R&D personnel in the ICT sector**.

Public funds aiming ICT R&D expenditures in the EU raised to EUR 6.7 billion in 2017. Estimated public ICT R&D expenditure was less than the Digital Agenda target (doubling publicly funded ICT R&D from 2007 to 2020), which requires an annual growth rate of 5.5%, with a gap of more than 26%. In 2017, public funding of ICT R&D was 7% of the EU total government budget allocations for R&D (GBARD).

2.2 European challenges for successful AI adoption

Europe has to address several crucial challenges when the critical enablers for an AI transition are examined. On the one hand, Europe should put more efforts to catch up with the current AI leaders, the United States and China, since in most of metrics the European continent is lagging behind. On the other hand, **to ensure inclusive and sustainable growth, the geographical disparities regarding AI adoption inside the continent have to be confronted**. An apparent gap in AI readiness inside Europe is evident, with Southern and Eastern Europe being critically behind than the Northern states. **The United Kingdom, Scandinavian countries, and Ireland seem more ready to welcome an AI transition. Conversely, Spain, Portugal, Italy, and Greece have much room to improve their performance**. Germany is the only dominant country from Europe that files for AI patents worldwide. From the perspective of the scientific publication, the scene is altered, and the United Kingdom, France and Italy appear in the top five countries worldwide in several fields of AI technologies. Therefore, **the need to foster the industrial application and the commercialization of academic breakthroughs seems critical for Europe**.

These challenges can be traced to the **notably lower business digitalization in Europe and especially to the southern part of it in contrast to the northern one where it is significantly more developed**. High level of digitization adoption and diffusion are prerequisites to adopt AI technologies and applications. Smaller firms predominate the entrepreneurial landscape of the European South. SMEs digitisation is more arduous in comparison with larger companies' digital transformation. The limited financial resources, the inability to attract senior professionals, the lack of digital awareness and the absence of an enabling general entrepreneurial framework could serve as reasoning, among other factors, to the low digitization levels.

Also, **investment in AI development is substantially lower in the southern firms compared to their northern counterparts**. This is also evident in the small share of R&D investment, where current statistics point out a high concentration in the four largest economies in the EU. Furthermore, **it is vital to promote collaboration among research institutions and encourage synergies with the private sector**. Several barriers hinder such ventures. **Primary concerns are the lack of funding, bureaucracy and cultural differences between academics and business managers**. Simultaneously, providing further motives to fruitful collaborations, such as prizes for excellence, could assist in this scope (Davey et al., 2018). The regional dimension is also critical for fostering synergies and the ability to attract foreign talent and researchers.

Connectivity is vital for ensuring an AI successful expansion. The North EU States possess a comparative advantage in contrast to the South where not only the quality of the connection is inferior, but also Greece, Cyprus, Portugal, Spain, and Bulgaria have no offer belonging to the inexpensive cluster of the Broadband Price Index. North of Europe is quite advanced when the digitization of the public sector is concerned. The South still tries to follow without achieving the magnitudes of the northern part of Europe. The development of AI technologies demands the increased accessibility of high-quality data. Open data policies should be considered thoroughly, providing that the rights to people privacy be secured.

Human skills are imperative for AI adoption since both high technological capabilities and emotional intelligence abilities, which cannot be developed by machines, are equally important. **Human capital skills are already in shortage of more developed countries with respect to AI applications**. Vital regional differences are also present in this dimension**. Southern EU member states demonstrate a severe insufficiency in digital skills progress. Considerable risk of a potential digital exclusion exists**. For instance, in Portugal and Greece, approximately half of their citizens do not possess at least the necessary digital skills. Vast disparity among EU states is present as regards the share of the labour force that has no digital skills. What is more, ICT specialists are not equally shared among EU member states as well. The South experiences a significantly less proportion of them relating to the total level of employment. Slightly more encouraging is the landscape pertaining to the ratio of ICT graduates to the total number of graduates, but still, there is potential to improve.

Enterprises located in Southern Europe do not adopt AI applications at a fast pace for several reasons. Awareness of the benefits that AI could offer to their operation and services is relatively low. The lack of digitized processes also could hinder their efforts significantly to transit to an AI business model. The main drivers of Digital transformation in European firms are the fast broadband connections, social media, and mobile applications. E-commerce and e-business technologies appear to be more dominant, once again, in Northern EU member states where the southern areas are lagging behind.

European AI startups are mainly commencing their activities in the United Kingdom, France and Germany. New ventures in the field of AI are significantly fewer in the Southern and Eastern states. Large companies are more digitized compared to SMEs since they rank higher for specific vital indicators relating to tracking digitization performance. Additionally, big players possess a scale advantage, making it easier for them to employ internal ICT specialists. AI **SMEs shortage in Europe could be attributed to cost-related barriers and insufficient skills to adopt such technologies**. Implementing AI technologies and transition to the AI business model requires extensive investments in technology and infrastructure and increased fixed costs, which could be significant challenges for a small company implying higher risks and opportunity costs. The shortage of available data resources could also hinder small and novel firms.

Simplifying regulations for SMEs to cluster with other firms or with research institutes could pave their way to AI adoption. Creating a facilitating environment that minimizes bureaucratic procedures, motivates participants, focuses on the targets of a cycling economy, open innovation and the AI technologies as a core strategy of such collaborations could serve a double scope. Firms will transform their business model having in its core AI applications, and they will share the risk mitigated by investing in new technologies increasing their productivity significantly, while Universities could be able to commercialize their research amplifying its added value in the production, leading to a more sustainable growth model overall.

Statistics mentioned above imply that SMEs are not equally capable of attracting personnel with increased digital skills compared to larger firms. In addition, access to advanced infrastructures such as fast broadband connection medium-high Cloud Computing services is ambitious for

either small or young firms. Also, big data utilization in European SMEs is still in its infancy. Databases offering quality data and open data policies could act as an essential motive for firms thinking to invest in AI.

Cultural barriers should not be neglected, as well. Southern countries might need to nurture a technological friendly environment that would encourage firms and particularly new ventures to produce, adopt and diffuse innovation. AI should be included in the strategic planning of a firm. Smart specialization focused on technological advance could benefit the scope of AI adoption. Particularly the involvement of SMEs to the chain of value could boost the adoption of AI. In the same vein, the absence of regional specialization to the southern parts of Europe limits the potential to exploit technological and innovation spill overs that could enhance the transition to an AI ecosystem. Such spill overs could favour SMEs since their efforts that could be boosted to a large extent.

On another note, the lack of prominent technology companies in Europe affects the development of novel firms. A new entrepreneur would be more willing to undertake AI investments if larger firms that could buy or invest in this venture were available. This possibility mitigates the undertaken risk substantially, encouraging SMEs' growth. Furthermore, startups and SMEs might be in an adverse position compared to larger firms regarding their ability to absorb public funding, since the bureaucratic and technical procedures are not negligible.

PrometheEUs

CHAPTER III
What rules for an AI-ready Europe

3.1 The DSM strategy: an ambitious European digital policy takes shape

Considering that the global economy has been rapidly becoming digital and the importance of ensuring European position as a world leader in the digital economy, on 6 May 2015, the European Commission launched “A Digital Single Market Strategy for Europe” (DSM).

The Digital Single Market Strategy aims to create a market in which the free movement of goods, persons, services and capital is ensured and where individuals and businesses can seamlessly access and exercise online activities under conditions of fair competition, and a high level of consumer and personal data protection, irrespective of their nationality or place of residence.
The DSM Strategy is built on three pillars, including 16 specific initiatives: 1) better access for consumers and businesses to online goods and services across Europe; 2) creating the right conditions for digital networks and services to flourish; 3) maximizing the growth potential of our European Digital Economy.

The first pillar requires the rapid removal of key differences between the online and offline worlds to break down barriers to cross-border online activity, while the second needs high-speed, secure and trustworthy infrastructures and content services, supported by the right regulatory conditions for innovation, investment, fair competition and a level playing field. The third pillar requires investment in ICT infrastructures and technologies such as Cloud computing and Big Data, and research and innovation to boost industrial competitiveness as well as better public services, inclusiveness and skills.

To achieve the objectives of the first pillar, the Strategy proposed eight key actions: 1) legislative proposals to simplify and make effective cross-border contractual regulations for consumers and businesses; 2) a review of the Regulation on consumer protection cooperation; 3) affordable high-quality cross-border parcel delivery; 4) preventing unjustified geo-blocking; 5) a survey on competitive practices in the e-commerce sector, relating to the sale of goods or the provision of online services, 6) legislative proposals to reform copyright regulation; 7) revision of Directive 93/83/EC for the coordination of copyright laws and related rights applicable to satellite broadcasting and cable retransmission in order to ensure better cross-border access to broadcaster services in Europe; 8) legislative proposals to reduce the administrative burden arising from the current VAT system and to harmonize the different national systems.

In order to create the right conditions for the development of digital networks and services (second pillar), the strategy put forward 5 concrete actions: an ambitious reform of the telecommunications regulatory system, the revision of the Audiovisual Media Services Directive, the analysis of the role and impact of digital platforms and questions regarding the responsibility of online content, a revision of the E-privacy Directive and the creation of a public-private partnership in the field of information security.

Finally, to encourage European companies to make greater use of technology in the production process (third pillar), the Strategy proposed initiatives on the possession and free circulation of data, the adoption of an integrated plan on ICT standards and the extension of the European framework for the interoperability of public services and an action plan for e-government in 2016-2020.

Throughout the Juncker Commission, 30 legislative proposals on the Digital Single Market have been put forward (29 of these have been agreed upon by the co-legislature with the exception of the E-privacy Directive).

3.2 From the 2018 AI Strategy to the 2019 AI HLEG Documents

Artificial intelligence was not explicitly mentioned in the DSM Strategy. It made its (modest) appearance only in the Mid-Term Review.

The first (partial) step towards a European AI strategy came from the European Parliament, that solicited the European Commission to draft a comprehensive strategy, built on two mainstays. The first was to try to reduce the investment gap with other world areas and countries and, the second, to set up a regulatory framework consistent with European laws and values.
On 16 February 2017, the European Parliament adopted a **resolution with recommendations to the Commission on Civil Law Rules on Robotics**, describing the benefits related to the increasing use of AI in terms, for example, of safeguarding workers in the more difficult or dangerous professions, but also, in general, the impact on the job market and the skills required from workers.
In May 2017, the Commission published its **mid-term review of the Digital Single Market Strategy** underlining the importance of building on Europe's scientific and industrial strengths, as well as on its innovative startups, to be in a leading position in the development of AI technologies, platforms and applications.
On 9 March 2018, the Commission launched a selection for the creation of an **AI high-level group of experts** with the task, among other things, of preparing a proposal for guidelines on ethical development and use of AI in compliance with the EU Charter of Fundamental Rights, considering issues such as fairness, security, transparency and the future of the world of work and democracy. On the same date, the Commission also opened a call for the formation of a **group of experts on damage and new technology responsibility** with the task of advising the Commission on the applicability of the Directive on damage liability regarding defective products to traditional products and new technologies.

Considering the importance of AI and the tremendous opportunities for growth connected to its deployment and use, on 10 April 2018, 25 European countries[11] signed a **Declaration of Cooperation on Artificial Intelligence**. Above all, the Member States agreed to work together on the most important issues raised by AI, to ensure Europe's competitiveness in the research and deployment of AI and deal with social, economic, ethical and legal questions. It was endorsed by the European Council in June 2018.

On 25 April 2018, the European Commission published a **communication putting forward a European approach to artificial intelligence** based on three pillars: 1) being ahead of technological developments and encouraging uptake by the public and private sectors with the Commission increasing its annual investments in AI by 70% under the research and innovation program Horizon 2020, reaching €1.5 bln for the period 2018-2020, connecting and strengthening AI research centers across Europe and supporting the development of AI applications in key sectors and an "AI-on-demand platform" that will provide access to relevant AI resources in the EU for all users; 2) preparing for socio-economic changes brought about by AI supporting business-education partnerships to attract and keep more AI talent in Europe and training and retraining schemes for professionals, also encouraging the modernization of Member State education and training systems and foreseeing changes in the labor market and skills mismatching; and 3) ensuring an appropriate ethical and legal framework.

On 7 December 2018, the Commission published **the Coordinated Plan on AI** resulting from the work of the 25 Member States which signed the Declaration of Cooperation on Artificial Intelligence in April 2018. It details actions to be started in 2019-2020 and prepares the ground for activities in the following years. It will be reviewed and updated annually. Considering that only five Member States had already adopted a national AI strategy with a specific budget (France, Finland, Sweden, the UK and Germany), the document has provided a strategic framework for national AI strategies encouraging the adoption of them by mid-2019. This Plan identifies some goals and actions: 1) reinforcing cooperation with the private sector; 2) strengthening excellence in trustworthy AI technologies and broader dissemination; 3) adapting learning and training programs and systems to better prepare society for AI; 4) building up the European data space essential for AI in Europe, including for the public sector; 5) developing ethics guidelines with a global perspective and ensuring an innovation-friendly legal framework; and 6) better understanding security-related aspects of AI applications and infrastructure.

On 12 February 2019, the European Parliament adopted a **Resolution on a comprehensive European industrial policy on artificial intelligence and robotics.** Considering that AI is one of the strategic technologies for the 21st century, both globally and in Europe, bringing about positive changes for the European economy, enabling innovation, productivity, competitiveness and wellbeing and the necessity to adopt a coordinated approach at a European level to compete with the massive investments made by third countries, especially the US and China, the

[11] Austria, Belgium, Bulgaria, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, the UK, Norway.

resolution underlines the importance of ensuring a trusted ecosystem for AI technology development. It also highlights that there are strong ethical, psychological and legal concerns about the autonomy of robots, their obvious lack of human empathy and their impact on the doctor-patient relationship, which have not yet been properly addressed at EU level. This particularly regards the protection of patients' personal data, liability, and the new economic and employment relationships that will emerge.

On 9 April 2019, the High-Level Expert Group on AI presented the "**Ethics Guidelines for Trustworthy AI**", following the publication of the guidelines' first draft in December 2018 on which more than 500 comments were received through an open consultation. The document sets out a framework for achieving trustworthy AI aiming to offer guidance to all stakeholders identifying a list of ethical principles, by providing guidance on how such principles can be operationalized in socio-technical systems. Guidance is provided in three layers of abstraction, from the most abstract in Chapter I to the most concrete in Chapter III, closing with examples of opportunities and critical concerns raised by AI systems.
It underlines that trustworthy AI should be: 1) **lawful**, complying with all applicable laws and regulations; 2) **ethical**, ensuring adherence to ethical principles and values; and 3) **robust**, both from a technical and social perspective, since, even with good intentions, AI systems can cause unintentional harm.

The Guidelines are divided into 3 chapters: 1) Foundations of Trustworthy AI, setting the ethical principles that must be adhered to in order to ensure ethical and robust AI. Namely, the document highlights the importance to ensure the respect for human dignity, freedom of the individual, democracy, justice and the rule of law, to guarantee equality, non-discrimination and solidarity and the safeguarding of citizens' rights. The Guidelines identify four ethical principles, rooted in fundamental rights, which must be respected in order to ensure that AI systems are developed, deployed and used in a trustworthy manner, specified as ethical imperatives and, namely: i) respect for human autonomy; ii) prevention of harm; iii) fairness; and iv) explicability; 2) Realizing Trustworthy AI, translating these ethical principles into 7 key requirements that AI systems should implement (human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non- discrimination and fairness, societal and environmental wellbeing and accountability) and meet throughout their entire life cycle, offering both technical and non-technical methods for their implementation; 3) Assessing Trustworthy AI, setting out a concrete and non-exhaustive Trustworthy AI assessment list to operationalize the requirements of Chapter II, offering AI practitioners practical guidance.

Finally, on 26 June 2019 the High-Level Expert Group on AI (AI HLEG) presented the report "***Policy and investment recommendations for trustworthy Artificial Intelligence***" representing the second deliverable of the Group. It is a very interesting document which identifies 4 major areas of impact - Humans and Society, Private Sector, Public Sector and Research and Academia - and indicates which impact to achieve with AI and identifies the enablers for trustworthy AI (infrastructures, skills and education, appropriate governance and regulatory framework and a

lucrative funding and investment climate), providing a number of recommendations on how this can be accomplished.

Specifically, it underlines the importance of encouraging investments, research and development on the impact of AI on individuals and society and emphasizes the importance of: 1) increasing knowledge and awareness of AI through digital literacy and courses (e.g. MOOCs) across Europe providing elementary AI training, supporting and further developing basic education on AI and digital literacy, particularly in primary, secondary and tertiary education systems, as well as beyond, creating an AI competence framework for individuals (including a focus on the core skills required), institutionalizing a dialogue between policy-makers, developers and users of AI technology, informing the public at large; 2) protecting the integrity of humans, society and the environment, refraining from disproportionate and mass surveillance of individuals, countering commercial surveillance of individuals (particularly consumers) and society, giving consideration to power asymmetries between institutions, businesses and individuals arising from the growth of digital devices and systems and the rapid expansion of digital data that they generate, introducing a mandatory self-identification of AI systems and fostering the development of AI solutions that address sustainability challenges; 3) promoting a human-centric approach to AI at work, fostering the research, development and deployment of human-centric AI systems in work contexts, encouraging automation of dangerous tasks and when humans are put at risk, applying a process of representation, consultation and, where possible, co-creation, encouraging re-skilling and up-skilling of workers and establishing a fully-fledged European transition fund to help manage the AI transformation in a socially responsible way; 4) ensuring that no one is left behind, introducing a duty of care for developers of consumer-oriented AI systems to ensure that these can be used by all intended users (in particular, users with disabilities), fostering a universal design approach, encouraging the development of AI tools and applications targeted to help vulnerable demographics and establishing a European Strategy for Better and Safer AI for Children.

With regard to the private sector, the document highlights the need for: 1) boosting the uptake of AI technology and services across sectors in Europe, allocating significant resources in the InvestEU program to support the transformation of European enterprises to AI-enabled solutions, creating an easy avenue for startups and SMEs to funding and advice, fostering the availability of legal and technical support to implement Trustworthy AI solutions that comply with the Ethics Guidelines, encouraging companies to form partnerships with training programs addressing all levels of AI training; 2) fostering and scaling AI solutions by enabling innovation and promoting technology transfer, supporting the development and growth of AI technology firms in Europe, facilitating the transition of AI solutions from research labs to testing environments and to commercial markets, creating an EU-wide network of AI business incubators that connect academia and industry and stimulating – through the Horizon Europe Program - beneficial innovation by funding EU hackathons, competitions and industry challenge driven research missions in AI across various sectors; 3) setting up public-private partnerships to foster sectoral AI ecosystems, conducting, in the short term, a sectoral-based in-depth analysis of several selected AI ecosystems and, in the medium term, setting up Sectoral Multi-Stakeholder

Alliances (SMUHAs) for strategic sectors in Europe to build their AI ecosystems with the relevant stakeholders.

With regard to the public sector, the document recommends providing human-centric AI-based services for individuals offering to individuals the possibility, if requested, to interact with a human interlocutor, catalyzing AI development in Europe, fostering digitalization by transforming public data into a digital format, making strategic use of public procurement to fund innovation and ensure trustworthy AI, safeguarding fundamental rights in AI-based public services and protect societal infrastructures, by ensuring the application of the Ethics Guidelines for Trustworthy AI to AI systems deployed by the public sector.

Finally, concerning research and academia, the document underlines the necessity to develop and maintain a European strategic research roadmap for AI, focusing on areas of strategic value and opportunities, ensuring AI solutions that meet the Trustworthy AI principles and requirements, providing dedicated, significant and long-term research funding, creating incentives and support for interdisciplinary and multi-stakeholder research, simplifying and streamlining the structure of research funding instruments, creating the conditions for talents to find Europe attractive as a research environment, creating, strengthening and supporting additional Centres of Excellence (CoEs) that address strategic research topics and become a European level multiplier for a specific AI topic and, finally, encouraging cooperation at all levels.

3.3 Towards an EU AI Law: Von der Leyen Commission's digital proposals and the white paper

The first Von der Leyen Commission digital proposals were published on 19 February 2020, including two Communications (Shaping Europe's digital future & A European Strategy for data), a white paper (Artificial Intelligence: a European Approach to excellence and trust) and two reports (B2G Expert Group Report: Towards a European Strategy on business-to-government datasharing for the public interest and the Commission Report on Safety and liability implications of AI, the Internet of Things and Robotics).

The current EC updated and upgraded the DSM strategy, with its priorities and proposals. In the Communication "Shaping Europe's digital future", published on February 20, the Commission establishes three key objectives to ensure digital transformation complies with European values: 1) a technology that works for people; 2) a fair and competitive economy and 3) an open, democratic and sustainable society.

The AI White Paper aims at setting a framework for trustworthy Artificial Intelligence, based on excellence and trust.
In the so called "ecosystem of excellence", among several planned actions, the Commission aims at proposing to the Member States a revision of the 2018 Coordination Plan, facilitating the creation of excellence and testing centers that can combine European, national and private investments. This involves working with MSs to ensure that at least one digital innovation hub per MS has a high degree of specialization in AI, setting up a new public-private partnership in AI, data and robotics in the context of the Horizon Europe Programme.

For the other ecosystem ("ecosystem of trust"), the Commission assesses the main risks associated with AI in order to ensure a European regulatory framework for a trustworthy AI.
The risk-based approach allows for a proportionate regulatory intervention, heavier for high-risk AI applications than for other lower-risk applications.

According to the white paper, an AI application should be considered high-risk when it meets the following two cumulative criteria: 1) it is employed in a sector where, given the characteristics of the activities typically undertaken, significant risks can be expected to occur (for instance, healthcare, transport, energy and parts of the public sector); 2) the AI application in the sensitive sector is used in such a manner that significant risks are likely to arise (based on the kind of impact on presumably affected parties). Moreover, the use of AI applications for employment processes, biometric identification and other intrusive surveillance purposes would always be considered as high-risk.

Mandatory requirements for high-risk applications would cover the following areas: 1) training data; 2) data and record-keeping; 3) information to be provided 4) robustness and accuracy; 5) human oversight; 6) specific requirements for certain specific applications, such as biometric identification.
These requirements would be at least in part verified under prior conformity assessments, in line with already existing mechanisms for a large number of products being placed on the EU's internal market.

Of course, ex post controls could be still enforced by competent national authorities.
For non-high risk applications, the Commission envisages a voluntary labelling scheme, allowing the economic operators to signal the trustworthiness of their products or services.
As data is the essential enabler for AI, the European Data Strategy aims at Europe emerging as a leader in the data economy, providing for a single market for data and a larger role for European companies.

The Commission starts from acknowledging that the EU has the potential to be successful in the data-agile economy, thanks to its technology, its know-how and its highly-skilled workforce. However, several issues are holding the EU back from realising its potential in the data economy, mainly due to the fragmentation between Member States (compared to the small number of US and China-based Big Tech firms). Among the most important issues, the strategy lists: 1) availability of data; 2) imbalances of market power; 3) data interoperability and quality; 4) data governance; 5) data infrastructures and technologies; 6) empowering individuals to exercise their rights; 7) skills and data literacy; 8) cybersecurity.

Included in the actions envisaged by the strategy, the Commission aims at supporting business-to-business data sharing, investing in a High Impact Project on European data spaces and federated cloud infrastructures, by the establishment of EU-wide common, interoperable data spaces (in manufacturing, environment, mobility, health, finance, energy, agriculture, public administration and skills) and the setting up of a cloud service marketplace, empowering individuals regarding their data and investing in skills and general data literacy.

Both documents (the AI White Paper and the European Strategy for Data) provide specific policies for SMEs and elements for a proactive international approach.

3.4 Commenting on the AI white paper: how to foster EU innovation and a thriving ecosystem based on SMEs

An ecosystem of excellence

While the EU should strive to improve its current standing in research and innovation, increasing public and private investments but also better coordinating existing initiatives, most companies, especially SMEs, would be either only or mainly AI users. Therefore, for a competitive economy, the EU regulatory framework should lead the vast majority of companies to adopt AI easily and at a cost to be competitive.

EU objectives to increase R&D and productive excellence should not jeopardize the possibility for EU citizens and companies to have access to the best available AI technologies at a competitive price. A balanced approach should be used taking into full account the interests of all the concerned parties, including the vast majority of citizens and companies that would be adopters rather than R&D and/or commercial producers in the AI ecosystem.

For instance, while AI requires vast amounts of computing power, data and expertise, a wide gulf exists between the few companies that can afford these resources and everyone else. We need a close partnership between academia, government and industry providing affordable access to those high-end inputs. This means democratizing access to the essential tools and ensuring that they remain open and potentially shared by all.

Both training and advice to SMEs should be key activities for AI specialized digital innovation hubs (DIHs). For this reason, foreseeing only one DIH per Member State may involve a sizeable geographical barrier for SMEs, especially in larger countries. A more distributed network of DIHs providing expertise to SMEs in different regions should be pursued, possibly involving trade associations and larger AI technology players.

An ecosystem of trust

To start with, a EU-wide regulatory approach is preferable in order to avoid major risks of internal market fragmentation. Therefore, Member States should refrain from unilateral moves and look for agreements and alliances at EU level.
Although some new legislation is certainly required and a EU-wide regulatory framework is surely preferable to national, current legislation should apply whenever possible in order to avoid excessive market fragmentation and uncertainty and increase compliance costs for companies, especially SMEs.
Whenever possible, a clearer interpretation of current legislation to be applied to all products, including those embedding AI, should be chosen instead of new legislation reserved to AI products.

Although many mentioned concerns deserve a high level of scrutiny and sometimes need to be addressed by ad hoc regulation, it would be fairer to compare AI applications with a human-

based benchmark. It would not be realistic to expect AI achieving an error-free perfection where, in the same field, the same standard is not currently applied. This requirement could significantly stifle innovation, especially from SMEs and new entrants.

If the two proposed cumulative criteria to determine "high-risk" AI applications seem quite logical and could help provide legal certainty, exceptional additional instances should be better defined and limited to specific cases in order to avoid any unnecessary ambiguity, where the aim of the risk-assessment approach is rightly the opposite.

Concerning enforcement, the regulatory framework should mostly rely on ex-ante self-assessment, instead of an external procedure, in order to speed up the innovation process and ensuring a thriving European AI ecosystem, setting low compliance costs for SMEs, and ex-post enforcement, paramount to guaranteeing full compliance by AI developers and deployers.

PrometheEUs

CONCLUSIONS

The Global Race for AI

Although we are still far away from achieving a “broad” form of AI, capable of mimicking human intelligence, the far-reaching socio-economic implications of this new technology are already being felt at national and international level. Without proper governance frameworks in place, it will be impossible to guarantee a level playing field in the development and implementation of AI, and we will be at risk of drifting towards a more confrontational international stage, where the most technologically advanced nations could impose new forms of colonialism.

It is obvious that a global governance framework is not a silver bullet, and building it will not be a straightforward task given today’s adverse geopolitical winds. But we believe our two-tier proposal has the potential to garner widespread support, by establishing a set of basic norms, as well as envisioning more ambitious rules that a “coalition of the willing” could sign up to. These efforts must be accommodating enough for national regulations to flourish, as each society will deal with different problems regarding the implementation of this technology. AI poses global challenges that demand global solutions, but national and local institutions are better equipped to deal with other types of AI-related issues, such as how to regulate AI investment.

The Industrial Revolution brought us a “Great Divergence” between early adopters and those who lagged behind. We have inherited enormous differences between countries and power blocs in terms of readiness, capabilities and economic power. Similarly, in the current context, some countries will profit massively from AI in the short term, while others will depend on processes of diffusion and adopt this technology at a much slower pace. However, ensuring that everybody has a fair chance in the ongoing AI race is a more sustainable, positive-sum approach. By reducing tensions on an international level, and building a technology that works for everybody and not just a privileged few, even those who lead the way will secure more peace and prosperity in the long term.

In this respect, the EU will need to play a dual role. On the one hand, it should leverage its regulatory and soft power to help to craft this global governance framework, while being respectful towards the nuances in the positions of all countries. On the other hand, it should address the existing divergences between its own member states with regard to AI development. Attaining the latter objective is a necessary condition for pursuing the former: by putting forward

effective and inclusive solutions at home, the EU could increase its global standing and spearhead the adoption of an AI global governance framework that paves the way to a future of shared prosperity.

AI adoption in Europe

Following the aforementioned, we need to analyze challenges and develop well-structured policy endeavours to tackle with them, for fostering an optimal AI adoption in Southern Europe. These challenges can be traced to the notably lower business digitalization and e-commerce development in the southern countries in contrast to the northern ones where it is blooming. Also, investment in AI development is significantly lower in the southern firms compared to their northern counterparts. This is also evident in the small share of R&D investment, where current statistics point out a high concentration in the four largest economies in the EU. In general, firms located in the South of Europe appear to be less capable of developing, adopting and transferring innovation and technology breakthroughs. Furthermore, small business size, which is a structural characteristic of the business demography in the South, is a significant barrier and affects any efforts to improve the technological advancement and invest heavily in AI.

Specifically, SMEs have to deal with a series of issues for being able to follow the current trends in AI adoption. Cultural barriers and a fear of change and transition to a new business model are affecting negatively such decisions. The transition to an AI model, in most cases, is not included in the strategic business objectives. The shortage of technology experts and the difficulty for SMEs to attract them is another critical issue that delays the AI escalation. From the infrastructure perspective, digitization of the firm and its services is key for AI enabling. However, endeavours of SMEs in this direction are evolving at a slower pace relating to bigger players.

Understanding the reasoning behind this phenomenon is complicated since it could be correlated to various factors such as insufficient infrastructure and connectivity levels, lack of specialized human capital which appears to be concentrated in northern Europe, possible lag in R&D investment incentives, lack of resources or even lack of awareness in the potential benefits of AI. These differences matter. Countries that are AI leaders could have faster growth and higher productivity than the rest, making the lag even more substantial and difficult to catch up. Primarily, if the potential digital gap among the EU and the other digital leaders (China and the US) is taken into account, the in-house challenges should be addressed holistically and instantaneously. Finally yet importantly, several issues remain unaddressed relating to AI in a more general context. The transformation of the needed human capital in the industry could lead to a severe loss of employment positions, issues of increased inequality between leading and following countries, or infringing on human rights and values, such as privacy are, among others, critical concerns.

Strengthening the AI ecosystems of excellence and trust: suggestions from Southern Europe

After an initial lack of attention on AI from EU institutions and Member States, starting from 2018, the European Commission, together with the European Parliament and national governments, began to increasingly invest in an ambitious AI Strategy. This effort has still not resulted in decreasing the economic and technological gap with the most advanced countries, such as the US and China. However, it is a preliminary milestone based on two mainstays - to try to reduce the investment gap with other world areas and countries and to set up a regulatory framework consistent with European laws and values.

According to the Commission's approach, achieving a trustworthy and human-centric AI would allow Europe to improve its competitiveness, spurring excellence and long-term investments.
The AI White Paper, published in February 19, follows the same approach and needs to be assessed together with the European Data Strategy.
We believe that both documents are on the right path and should be followed by legislative measures, as well as other non-legislative actions.
However, from a Southern European perspective (that probably could be applied to many other countries), we need to ensure that some elements are not missed or overlooked.

First, while the EU should strive to improve its current standing in research and innovation to achieve top level excellence, most companies, especially SMEs, would be either only or mainly AI users. Therefore, the EU regulatory framework should encourage the vast majority of companies to adopt AI easily and at a cost so they can be competitive.

EU objectives to increase R&D and productive excellence should not jeopardize the possibility of EU citizens and companies having access to the best available AI technologies at a competitive price. For instance, while AI requires vast amounts of computing power, data and expertise, a wide gulf exists between the few companies that can afford these resources and everyone else. A closer partnership between academia, government and industry providing affordable access to those high-end inputs is required. This means democratizing access to the essential tools and ensuring that they remain open and potentially shared by all. European AI should be based on interoperability, open standards and competitive infrastructures.

Second, strengthening the digital skills of SMEs and society at large would be pivotal for a viable AI transformation. Digital innovation hubs (DIHs) supplying both training and advice to SMEs should be key activities. However, foreseeing only one AI specialized DIH per Member State may involve a sizeable geographical barrier for SMEs, especially in larger countries. A more distributed network of DIHs providing expertise to SMEs in different regions should be pursued, possibly involving trade associations and larger AI technology players, as well as massive programs addressed to the general population.

Third, a EU-wide regulatory approach is preferable in order to avoid major risks of internal market fragmentation. Therefore, Member States should refrain from unilateral moves and look for agreements and alliances at EU level.

Fourth, although some new legislation is certainly required and a EU-wide regulatory framework is surely preferable to national, current legislation should apply whenever possible in order to avoid excessive market fragmentation and uncertainty, as well as increasing compliance costs for companies, especially SMEs. Whenever possible, a clearer interpretation of current legislation to be applied to all products, including those embedding AI, should be chosen instead of new legislation reserved to AI products.

Concerning enforcement, the regulatory framework should mainly rely on ex-ante self-assessment, instead of an external procedure, in order to speed up the innovation process and ensure a thriving European AI ecosystem, setting low compliance costs for SMEs, and ex-post enforcement, paramount to guaranteeing the full compliance by AI developers and deployers.

PrometheEUs

REFERENCES

Abrahms, M., & Mierau, J. (2017). Leadership matters: The effects of targeted killings on militant group tactics. Terrorism and Political Violence, 29(5), 830-851.

Allison, G. (2017). The Thucydides trap. Foreign Policy, 9. Available online at: https://foreignpolicy.com/2017/06/09/the-thucydides-trap/

Anderson, K., & Waxman, M. C. (2017). Debating Autonomous Weapon Systems, their Ethics, and their Regulation under international law. University of Columbia.

Annoni, A., Benczur, P., Bertoldi, P., Delipetrev, B., De Prato, G., Feijoo, C., ... & Cobo, M. L. (2018). Artificial Intelligence: A European Perspective (No. JRC113826). Joint Research Centre (Seville site).

Bakerjian, L. (2019) International Law and Drone Warfare. Publicgoods.eu, available online at: https://publicgoods.eu/international-law-and-drone-warfare

Barnes, J. E., & Chin, J. (2018). The new arms race in AI. The Wall Street Journal, 2. Available online at: https://www.wsj.com/articles/the-new-arms-race-in-ai-1520009261

Baum, M. A. (2013). The Iraq coalition of the willing and (politically) able: Party systems, the press, and public influence on foreign policy. American Journal of Political Science, 57(2), 442-458.

Berger, R. (2018). Artificial Intelligence-A Strategy for European Startups. Recomendations for policy makers, Asgard–Human Venture Capital, Berlin.

BIJ (2019). Drone Warfare research. Bureau of Investigative Journalism. Available online at: https://www.thebureauinvestigates.com/projects/drone-war

Bob, C., & Nepstad, S. E. (2007). Kill a leader, murder a movement? Leadership and assassination in social movements. American Behavioral Scientist, 50(10), 1370-1394.

Boyle, M. J. (2013). The costs and consequences of drone warfare. International affairs, 89(1), 1-29.

Bradford, A. (2012) The Brussels Effect. Northwestern University Law Review, 107, 1.

Brundage, M. et al. (2018) The Malicious Use of Artificial Intelligence: Forecasting, Prevention and Mitigation. Future of Humanity Institute. University of Oxford.

Brzozowski, A. (2019). No progress in UN talks on regulating lethal autonomous weapons. Euractiv. 22nd November 2019. Available online at: https://www.euractiv.com/section/global-europe/news/no-progress-in-un-talks-on-regulating-lethal-autonomous-weapons/

Buchanan, A., & Keohane, R. O. (2015). Toward a drone accountability regime. Ethics & International Affairs, 29(1), 15-37.

Bughin, J. et al. (2019) Tackling Europe's gap in digital and AI. McKinsey Global Institute. Available online at: https://www.mckinsey.com/featured-insights/artificial-intelligence/tackling-europes-gap-in-digital-and-ai

Bughin, J., Seong, J., Manyika, J., Hämäläinen, L., Windhagen, E., & Hazan, E. (2019). Notes from the AI Frontier: Tackling Europe's Gap in Digital and AI. McKinsey&Company: New York, NY, USA.

Calo, R. (2017). Artificial Intelligence policy: a primer and roadmap. UCDL Rev., 51, 399.

Castro, D., McLaughlin, M., & Chivot, E. (2019). Who is winning the AI race: China, the EU, or the United States? Center for Data Innovation.

CB Insights (2017) Research report: Artificial Intelligence Trends to Watch in 2018. CB Insights research team.

CB Insights (2020) Research briefs: The Artificial Intelligence Start-ups re-defining Industries. CB Insights research team.

Cave, S., & ÓhÉigeartaigh, S. S. (2018, December). An AI race for strategic advantage: rhetoric and risks. In Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, pp. 36-40.

Craglia, M. et al. (2019). Artificial Intelligence: an European perspective. Publications Office of the European Union.

Dafoe, A. (2018). AI governance: A research agenda. Governance of AI Program, Future of Humanity Institute, University of Oxford.

Davey, T., Meerman, A., Galán-Muros, V., Orazbayeva, B., & Baaken, T. (2018). The state of university-business cooperation in Europe.

European Commission (2020). A Europe fit for the digital age.

European Commission (2019). The Digital Economy and Society Index (DESI). Recuperado de https://ec. europa. eu/digital-single-market/desi.

Executive order 13859 (2019). Artificial Intelligence for the American People.

Farrell, H., & Newman, A. L. (2019). Weaponized interdependence: How global economic networks shape state coercion. International Security, 44(1), pp. 42-79.

Floridi, L. et al. (2019). Designing AI for Social Good: Seven Essential Factors. University of Oxford.

G20 (2019) G20 AI Principles. Available online at: https://www.g20-insights.org/wp-content/uploads/2019/07/G20-Japan-AI-Principles.pdf

Gady, F. (2019). Elsa B. Kania on Artificial Intelligence and Great Power competition. The Diplomat. 31st December 2019. Available online at: https://thediplomat.com/2020/01/elsa-b-kania-on-artificial-intelligence-and-great-power-competition/

Gasser, U., & Almeida, V. A. (2017). A layered model for AI governance. IEEE Internet Computing, 21(6), 58-62.

Gill, I. (2020) Whoever leads in Artificial Intelligence in 2030 will rule the world until 2100. Brookings. Available online at: https://www.brookings.edu/blog/future-development/2020/01/17/whoever-leads-in-artificial-intelligence-in-2030-will-rule-the-world-until-2100/

Gigova, R. (2017). Who Vladimir Putin thinks will rule the world. CNN. 2nd September 2017. Available online at: https://edition.cnn.com/2017/09/01/world/putin-artificial-intelligence-will-rule-world/index.html

Greitens, S. C. (2016). Dictators and their secret police: Coercive institutions and state violence. Cambridge University Press.

Hacker, P. (2018). Teaching fairness to artificial intelligence: existing and novel strategies against algorithmic discrimination under EU law. Common Market Law Review, Vol. 55, pp. 1143-1186.

Harari, Y. (2019) Who Will win the Race for AI? Foreign Policy.

High-level Expert Group on Artificial Intelligence (2019). Ethic guidelines for trustworthy AI. European Commission publication.

Hogarth, I. (2018) AI Nationalism. Blog. Available online at: https://www.ianhogarth.com/blog/2018/6/13/ai-nationalism

Horowitz, M. C., Kreps, S. E., & Fuhrmann, M. (2016). Separating fact from fiction in the debate over drone proliferation. International Security, 41(2), 7-42.

I-Com (2020), Leveraging AI in the Age of Pandemics

I-Com (2020), Artificial Intelligence Fit for Europe. Building Trust in Times of Emergency

I-Com (2020), Innovative Europe: The Way Forward. Taking Stock and Thinking Ahead

I-Com (2019), The New Europe in the Global Digital Era: Current Rule-Maker, Future Investor?

I-Com (2018), Data-driven innovation & Artificial Intelligence. Which strategy for Europe?

IOBE (2018), Measuring the economic impact of digital skills in Greece: challenges ahead

Jordan, J. (2014). Attacking the leader, missing the mark: Why terrorist groups survive decapitation strikes. International Security, 38(4), 7-38.

Latici, T. (2019) Civil and Military Drones: Navigating a disruptive and dynamic technological ecosystem. European Parliament Research Service.

Leavy, S. (2018, May). Gender bias in artificial intelligence: The need for diversity and gender theory in machine learning. In Proceedings of the 1st international workshop on gender equality in software engineering (pp. 14-16).

Lee, K. (2019) AI superpowers: China, Silicon Valley and the New World Order. Houghton Mifflin.

Mazzucato, M. (2013). The Entrepreneurial state: Debunking Public vs Private sector myths. Anthem Press.

Moore, A. in LeVine, S. (2018). The stakes for who wins the AI race. Axios. Available online at: https://www.axios.com/the-stakes-for-who-wins-the-ai-race-0363d9cd-0d97-4a5a-9ee6-36a7fb03ff44.html

Nye Jr, J. (2017). Deterrence and dissuasion in cyberspace. International security, 41(3), pp. 44-71.

Nye, J. (2011). Nuclear lessons for cyber security? Strategic Studies Quarterly, 5(4), pp. 18-38.

Nye, J. (2019) Eight norms for Stability in Cyberspace. Project Syndicate. Available online at: https://www.project-syndicate.org/commentary/eight-norms-for-stable-cyberspace-by-joseph-s-nye-2019-12

OECD (2019) Recommendation of the Council on Artificial Intelligence. Available online at: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

OECD (2019a), Artificial Intelligence in Society, OECD Publishing, Paris, https://doi.org/10.1787/eedfee77-en.

Oxford Insights (2019). Government Artificial Intelligence Readiness Index 2019. Oxford Insights. Available online at: https://www.oxfordinsights.com/ai-readiness2019

Rodrik, D. (2019) Putting Global Governance in its place. NBER Working Paper Series.

Rodrik, D. (2020). Globalisation after Covid-19: my plan for a rewired planet. Prospect Magazine. May 4th 2020. Available online at: https://www.prospectmagazine.co.uk/magazine/dani-rodrik-globalisation-trade-coronavirus-who-imf-world-bank

Schenker, J. (2019) Europe's place in the AI Race. The Innovator. Available online at: https://innovator.news/europes-place-in-the-ai-race-c3e8c9dd0f10

Sher-Jan, M. (2019) Global impacts of the GDPR: one year later. RadarFirst. Available online at: https://www.radarfirst.com/blog/global-impacts-of-the-gdpr-one-year-later/

State Council of China (2017) New Generation Artificial Intelligence Development Plan.

Thompson, N. (2018). Emmanuel Macron talks to Wired about France's AI strategy. Wired. 31st March 2018. Available online at: https://www.wired.com/story/emmanuel-macron-talks-to-wired-about-frances-ai-strategy/

Thornhill, J. (2020). Trusting AI too much can turn out to be fatal. Financial Times. Available Online at:

Walch, K. (2019). Why Cognitive Technology may be a better term than Artificial Intelligence. Forbes. 22nd December 2019. Available online at: https://www.forbes.com/sites/cognitiveworld/2019/12/22/why-cognitive-technology-may-be-a-better-term-than-artificial-intelligence/#3714da79197c

Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). The role and limits of principles in AI ethics: towards a focus on tensions. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (pp. 195-200).

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

Wladawsky-Berger, I. (2018). The Impact of Artificial Intelligence on the World Economy. Wall Street Journal. Available at: https://blogs.wsj.com/cio/2018/11/16/the-impact-of-artificial-intelligence-on-the-world-economy/

Wolff, G. (2020). Europe may be the world's AI referee, but referees don't win. Bruegel. Available online at:

Zou, J., & Schiebinger, L. (2018). AI can be sexist and racist—it's time to make it fair. Nature