PromethEUs

November 2023

PromethEUs Publication

Artificial Intelligence: Opportunities, Risks and Regulation

The PromethEUs network of think tanks, consisting of Elcano Royal Institute (Spain), I-Com – the Institute for Competitiveness (Italy), IOBE – the Foundation for Economic and Industrial Research (Greece) and the Institute of Public Policy – Lisbon (Portugal) has drawn up a joint paper on the EU Artificial Intelligence Act (AI Act) as the main output of its activity in the second semester 2023. With a view to contributing to the debate around the topic from a Southern European perspective, the paper sets the discussion on the latest developments of the dossier ahead of the agreement that negotiators aim to secure by the end of 2023. As trialogues are reaching their final stage some divisive subjects remain (e.g., the use of copyrighted content by AI, and biometric surveillance).

EXECUTIVE SUMMARY

Chapter 1: Regulatory and Policy Aspects of the AI Act

The first part of this chapter questions **what AI is and why we (should) care about it**. To answer this, the chapter starts by providing a **quick historical overview**, briefly describing the capabilities and milestones of AI to date. Following, a **definitions and classifications** section is included. Finally, a glimpse of **AI current and predicted impacts on the world** is provided, with the intention of making it clear why AI has become an urgent policy debate for the EU and other world institutions and blocs.

In the second part of the chapter, a more **comprehensive background to the present state of the current attempts to regulate AI** is provided. Firstly, the EU is addressed, referring to EU initiatives on AI, with a special focus on the **Commission communication "Artificial Intelligence for Europe"**, dating back to 2018, to the **2020 White Paper on Artificial Intelligence** drawn up by the Commission. Throughout this part, it is outlined, for example, how the EU arrived at the **AI Act's risk-based conception of AI regulation** and the evolution of AI from an important field of action, demanding the Member States' coordination, to a political priority for the 2019-2024 Von der Leyen Commission agenda. Following, the **AI Act proposal consistency with the EU Charter of Fundamental Rights** is analysed This consistency is crucial to understanding the Commission's proposal to regulate AI as the Charter is at the core of what will be considered risky or not. After

addressing the EU take on AI, **four big international initiatives and approaches to AI regulation** are presented. Firstly, the Convention on AI, currently under discussion at the **Council of Europe**, and its complementary nature to the EU AI Act is covered. Then, **three international approaches to AI regulation** by international blocs that diverge from the EU approach - **the USA, the UK, and China** – are described. The latest developments in each bloc's approach to regulating AI is explained, highlighting the core differences in contrast with the European view on AI regulation.

The third part of this chapter analyses the **EU AI Act** itself, namely its **regulatory approach**. The **Act's main structural points** are addressed - namely the Act's attempt to provide a clear definition of AI, the meaning and concrete application of its risk-based approach, the **role of its transparency and accountability**, how it addresses **data governance and privacy concerns**, as well as the Act's **mechanisms to ensure its compliance and enforcement**. Finally, the **main topics on the ongoing trilogue negotiations** on the AI Act (European Commission, Parliament, and Council) are discussed. In this part, the analysis on the debate looks at points such as which **definition of AI** will be adopted in the Act, the legality of **biometrics and real-time biometric surveillance** in public places, what **institutional framework** will serve the Act's enforcement requirements the best (either a Board or an Executive Office), what financial limits and criteria will be considered when the Member States define their **penalties** for Act infringements, the adaptations needed in the Act to regulate **foundational models and general purpose AI** (which is absent in the Commission's first proposal), and how the **classification of high-risk AI systems** in the AI Act will be defined in order to achieve the Act's objectives without undermining innovation and businesses in the EU.

The fourth part of the chapter looks at the **implementation of the EU AI Act**. In this part, the **impact and types of regulatory intervention** in the EU are considered, analysing the regulatory approach adopted in the AI Act in line with the **New Legislative Framework** (NFL) principles, as well as the **role firms will have in the application** of the Act. Secondly, this part addresses the subsidiary structure at the EU level, discussing **the legal form the Act takes on** (a regulation), the **European Artificial Intelligence Board**, the new **EU database for high-risk systems** (to be managed by the Commission), and the definition of **penalty** subsidiary roles in the Act's implementation. Following, the importance **human oversight** will be explored in its impact on the Act's implementation, as well as the main issues that still need to be addressed. Then, the **national competent authorities' identity** and foreseen actions will be looked at more closely. Finally, this fourth part of the chapter ends by discussing a fundamental part of the Act's implementation, namely the **impact of compliance for businesses and SMEs**.

The final part of this chapter on the regulatory aspects of the EU AI Act analyses its **future impact**. For example, one of the core "future-proof" characteristics of the AI Act relies on its **annexes** for AI classification and the possibility of their future modification. Will this work on such a fluid

technology as AI? The **equilibrium between the desired legal certainty** (an objective expressly identified in the AI Act) **and the restrictions and benefits it implies for AI providers and AI users** Is explored**.** What effects can it have on businesses operating within EU borders? Will compliance obligations attract international businesses due to the transparent market rules the AI Act creates? Or, on the contrary, will these businesses prefer other less-regulated markets to develop, use and sell their products? Moreover, is it fair to expect a **"Brussels effect"** for AI international regulation, as happened under the GDPR? Will there be any difference with AI? Or will we expect a **Washington/London/Beijing effect?** Furthermore, what are the main driving forces to believe in a common regulatory approach worldwide? What could hinder it?

Chapter 2: The impact of generative AI

The chapter analyses the **impact of generative AI**, **especially from an economic point of view**. The analysis describes the European competitive position (with a focus on the Southern countries - Greece, Italy, Portugal and Spain) in the field of generative AI, and also offers a global overview of the situation. Finally, the chapter focuses on the potential challenges related to generative AI and how to tackle them.

Generative AI is an advanced form of artificial intelligence that enables machines to learn from existing data to create new data or content, including audio, code, images, text, simulations and videos[1]. Machine learning and deep learning techniques[2] are at its core. **The key difference between generative AI and previous forms of AI lies in the ability to create new content**. Traditional AI may also use neural networks, but these models were not designed to create new content.

Preliminary estimates of the potential impact of generative AI on the world economy are impressive. According to a Goldman Sachs' report, it could drive a 7% (or almost $7 trillion) increase in global GDP and lift productivity growth by 1.5 percentage points over a 10-year period[3]. McKinsey & Company released similar figures. Having analysed 63 use cases across the global economy, the consultancy's analysts reported that generative AI has the potential to generate $2.6 trillion to $4.4 trillion in added value across industries[4]. Moreover, generative AI could result in a labour productivity growth of 0.1 to 0.6 % annually through to 2040, depending on the rate of technology adoption and redeployment of worker time to other activities[5].

[1] McKinsey & Company, What is generative AI?, January 2023
[2] Forbes, Unlock The Potential Of Generative AI: A Guide For Tech Leaders, January 2023
[3] https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html
[4] McKinsey & Company, The economic potential of generative AI, June 2023
[5] Ibidem

However, **automation processes induced by generative AI will impact on knowledge work**, particularly activities involving decision making and collaboration, which previously had the lowest potential for automation. Therefore, a reorganisation of and retraining in work is essential.

At the same time, generative AI may be the first type of automation capable of reducing inequality rather than increasing it, because it is actually based on language and, thus, can mimic higher skills compared to previous innovation waves. However, it is up to us humans to understand how to best use it. If we see it as a substitute for workers, we indeed risk high unemployment or wage compression as salaries would then have to compete with machine costs. If we recognise it as a complement that can enhance overall work performance, we can lay the foundation for a manageable transition, where different tasks than before are performed, but in most cases to the advantage of both workers and companies.

At the same time, with the influx of popular chatbots such as OpenAI's ChatGPT and Google's Bard, **the generative AI market is poised to explode**, growing according to some preliminary estimates to $1.3 trillion over the next 10 years from a market size of just $40 billion in 2022. Value is expected to show an annual growth rate (CAGR 2022-2032) of 42%, driven by training infrastructure in the near-term and gradually shifting to inference devices for large language models (LLMs), digital ads, specialised software and services in the medium- to long-term. Generative AI is on the way to expanding its impact from less than 1% of total tech spending to 12% by 2032[6]. The European generative AI market is also showing rapid growth (though less so). According to some estimates, **EU market value value[7] in the European generative AI market is projected to reach $12.25 billion in 2023** and is expected to show an annual growth rate (CAGR 2023-2030) of 24.52%, resulting in a market volume of $56.85 billion by 2030[8]. Among the European Member States, Germany has the highest value of the generative AI market ($1.90 billion), followed by France ($1.23 billion) and Italy ($0.87 billion). Spain ranks fifth ($ 0.74 billion) while Portugal is in the thirteenth position ($0.17 billion). **Taking country population into account, Denmark appears the largest generative AI market globally, with a market value per 100,000 inhabitants of $ 7.35 million, followed by Finland and Ireland. All countries of Southern Europe (Greece, Italy, Portugal, Spain) are below, displaying lower values, ranging from $1.68 million per 100,000 inhabitants for Portugal to $ 0.68 million per 100,000 inhabitants for Greece.**

As generative AI is already becoming an increasingly prominent part of everyday business activities and our daily lives, it gives rise to **several risks and ethical considerations**. For instance, AI-generated content could be used for malicious purposes, such as spreading **misinformation** or

[6] Bloomberg Intelligence, IDC (2023)
[7] Values are generated by the funding amount in Generative Artificial Intelligence initiatives and projects by companies
[8] Statista (2023)

creating **deepfakes**. It is crucial for developers and platforms to implement ethical guidelines and regulations to mitigate these risks. Other challenges involve **security** and **privacy** in terms of protecting user data and preventing identity theft. Moreover, **copyright issues** are very complex and difficult to manage within the current regulatory framework. This concerns two main aspects - the input materials used for training the models and the outputs produced by AI tools. However, **these risks could be manageable by fittingly adapting the current regulatory system to old and new challenges. The huge potential benefits for Europe, and especially for the Member States currently lagging behind in terms of digital skills, should not be overlooked.**

Chapter 3: The geopolitics of generative AI: international implications and the role of the European Union

Generative AI is a sub-field of AI that is shaking markets and industries due to its capacity to produce and recreate, among other functions, natural language and human-like interactions. This is also impacting the competition between countries to govern this trend. In this report, we analyse why is generative AI substantially different from AI and which countries are leading the AI race in private investments, patents, publications.

The results show that the **dynamics of U.S.-China competition are also being translated in the AI race, with both countries leading the rankings consistently.** While China leads in intellectual property and patents, the United States still leads the way in the Venture Capital and investments in high-risk markets. However, we also show that the most-used current projects in generative AI have a global reach, as they are open-source. This allows their use, and distribution without intellectual property rights and makes it more difficult to the biggest technological companies to *gate keep* developments on generative AI.

What also makes **generative AI so subversive is its strong dual-use component, with the potential to greatly impact both markets and economic activity and national security and defence.** These implications of generative AI are also being tackled in international forums at the multilateral level, and a consistent dialogue about digital rights and social implications is happening.

However, **the EU is specifically concerned about the rise of generative AI due to their poor performance where the U.S. and China lead, such as corporations, Venture Capital investments and patents. However, the EU has the potential to lead the discussion on the rights and security implications of generative AI with the AI Act and with its consistent efforts in establishing partnerships beside other like-minded countries outside the EU like South Korea, India or Japan.**

Chapter 4: AI readiness and the economic potential, with a focus on Southern EU

Artificial Intelligence is a technology with the potential to improve the well-being of people while enhancing productivity across all economic activities, empowering innovation, and helping address key global challenges. It is deployed in numerous economic sectors and its impact is driven by productivity gains of firms with the automation of processes and the support of the workforce with AI technologies but also by the increased consumer demand which stems from the customized and higher quality products and services. **Besides its unquestionable positive effects, it also carries new risks and raises challenges for individuals and the society, mainly concerning data protection, digital rights, and ethical standards, issues that must be addressed through a policy framework with respect to human rights and values.**

OECD has issued five 'AI principles' addressing these challenges alongside five recommendations for the policy makers to foster a suitable environment for the fruitful implementation of AI systems into everyday life. These principles are already adopted by all four members of PromethEUs network in their national strategies and policies. In the same line, EU Artificial Intelligence Act has been proposed which aims to create a balanced and proportionate horizontal regulation through a risk-based approach to safeguard a seamless digital transition that will promote economic growth while respecting firms and citizens' rights and values.

In 2021, only 7.9% of EU enterprises used at least one AI technology. Portugal is one of the pioneering countries in the EU ranking second with 17.3%, Spain's performance is close to the EU average with 7.7%, Italy is relatively lagging with 6.6% whereas Greece is second to last with 2.6%. The usage of AI is currently low, and this presents an opportunity for a mindful deployment of such technologies to rip its benefits as much as possible while keeping under control its pitfalls. The relevant research in all four countries steadily increased in the last two decades while there is a significant surge of investments in AI start-ups in the last three years. In addition, the EU funding for digital technologies in the 2021-2027 Multiannual Financial Framework, especially through the national Recovery and Resilience Plans, is an outstanding complement to national funding of the digital transition of each country.

CONTRIBUTORS

Chapter 1: Regulatory and Policy Aspects of the AI Act

IPP, Institute of Public Policy
Steffen Hoernig, André Ilharco

Chapter 2: The impact of generative AI

I-Com, Institute for Competitiveness
Stefano da Empoli, Maria Rosaria Della Porta

Chapter 3: The geopolitics of generative AI: international implications and the role of the European Union

Elcano Royal Institute
Raquel Jorge Ricart, Pau Álvarez-Aragonés

Chapter 4: AI readiness and the economic potential, focusing on the Southern EU

IOBE, Foundation for Economics & Industrial Research
Aggelos Tsakanikas, Konstantinos Valaskas

Table of contents

Chapter 1: Regulatory and Policy Aspects of the AI Act

1.1 What is AI and why do we care?

1.1.1 Quick historical overview

Artificial Intelligence (AI) has its roots in the mid-20th century, with key contributions from pioneers such as Alan Turing[9] and John McCarthy[10], among others. Early explorations in the field of AI aimed at **constructing intelligent systems capable of emulating human cognitive processes** were primarily directed towards addressing **complex problem-solving tasks**. The evolution of AI since then has been determined by the world's increasing data storage capabilities and computer processing speed, as well as by their cost. Since the early 2000s, AI developed the capabilities of **handwriting and speech recognition** (2000), **image recognition** (2009), **reading comprehension** (2016) and **language understanding** (2018), not to mention that at **least from 2018 all these capabilities were performed above-human level** by AI systems[11].

1.1.2 Definitions and classifications

Defining AI remains a complex undertaking today due to its diverse manifestations. One prominent category encompasses knowledge based and logical reasoning systems built upon structured data and rule-based logic. These **systems aim to emulate human cognition through the use of explicit rules and logical inference to derive conclusions and make decisions**. Another key facet of AI is the domain of **machine learning, big data and pattern recognition**. This branch focuses on developing **algorithms that enable computers to learn from and make predictions or decisions based on data, often in complex and unstructured environments**. Additionally, the emergence of **generative AI** (further explored in Chapter 2 - "The impact of generative AI") has ushered in a new era of creativity and innovation. Generative AI systems, such as deep learning-based models, possess the **capability to produce new content, including text, images and even entire simulations**, by learning the underlying patterns and structures from a given dataset. This type of AI has seen remarkable advances in fields such as art generation, content creation, and even the development of highly realistic synthetic media.

[9] https://redirect.cs.umbc.edu/courses/471/papers/turing.pdf
[10] https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1911
[11] https://ourworldindata.org/brief-history-of-ai

1.1.3 Predicted impacts

The **main breakthrough of AI is in prediction** (Agrawal, Gans, & Goldfarb, 2018) in all of its applications. As for the **benefits**, it is predicted that AI will **boost productivity, innovation and efficiency,** as it can process more data more rapidly. This boost will also be supported by its capability to **target potential customers** more precisely, **customise and personalise services** and predict customers' needs and preferences. **Trade and supply chains** will certainly be adapted due to its **power to optimise logistics, inventory management and predictive maintenance in real-time,** and using large sets of data. AI can perform several tasks of or even fully **substitute human labour** in many jobs (mainly those with repetitive and routine tasks), ranging from manufacturing, transport, justice to public administration. It is therefore predicted that **reskilling** will be needed to accommodate the displaced workforce, but also that new jobs will be created for an AI economy.

On other hand, **AI systems are opaque, complex, data dependent, and can make autonomous decisions impacting citizens' rights**. They have the power to recognise and **predict citizen**, worker and customer **whereabouts, preferences and choices from the data they knowingly or unknowingly provide.** Furthermore, interaction with **AI systems is becoming increasingly harder to distinguish from those with humans**. The capabilities of AI can amplify asymmetries of power and information towards citizens, **restricting fundamental rights and liberties**, and weakening democracy and the rule of law, **if left unregulated and unsupervised**.

To cope with these issues, the European Commission proposed a Regulation called the "AI Act" in April 2021 which is currently in its **final stages of negotiation** with the European Parliament and the Council of Ministers (**the "trilogue"**). These negotiations are expected to be concluded by the end of 2023, and the Act is expected to formally **enter into force in mid-2024**.

1.2 Introduction: EU and international background

1.2.1 The road to the AI Act

In **April 2018,** the Commission issued a communication **"Artificial Intelligence for Europe".** This was the first EU document directly paving the way for EU action on AI. The Commission underlined the **importance of steering both public and private investment into AI initiatives** to keep up with other international actors such as the US and Asia. Investments in these initiatives would be directed towards **research and innovation**, as well as guaranteeing **better data access in all of the EU**. Although a **Regulation on AI has still not been mentioned**, the Commission stated **the urgent need for an appropriate AI ethical and legal framework in the EU**, promoting trust and accountability around its development and use.

Just **one year later**, AI came to be regarded not only as a field demanding coordination and collaboration between Member States, but as a **full policy priority for the EU**. In **von der Leyen's political agenda** (von der Leyen, 2019), AI was taken on as an **EU policy priority for the period between 2019 and 2024**. In 2020, with the **White Paper on Artificial Intelligence**, the Commission mentioned for the **first time the idea of a future regulation of AI with a risk-based approach**. The latter's purpose was the creation of a proportional regulatory framework also promoting AI uptake and avoiding burdensome regulation on SMEs (in line with the EU Data Strategy).

1.2.2 Consistency with the EU Charter of Fundamental Rights

AI can affect fundamental rights in the EU, which is why the Commission proposed an **AI Act** in the first place, designed directly to defend the principles of the EU Charter of Fundamental Rights under AI. As Recital 28 puts it, the **EU Charter of Fundamental Rights acts as a compass when determining which AI systems should be classified as high and non-high-risk or even prohibited**. The main threats identified by the Commission in the Act affect core rights in the Charter such as the **dignity of the human person**, an extreme example of which are AI systems of social scoring (Recital 17).
Consistency implies the defence of the rights enshrined in the Charter, but also the **proportionality and minimising of some limitations on rights in the case of a clash**. Restrictions on the use and development of high-risk AI technology may **limit the freedom to conduct business** (Art. 16) or the **freedom of art and science** (Art. 13). Furthermore, its transparency obligations, such as the conformity assessment (Art. 19), will affect **intellectual property rights** (in compliance with the existing EU legal framework on the matter, such as Directive 2016/943).

1.2.3 International context of AI regulation

a. Council of Europe

The Council of Europe, a Strasbourg-based international forum to **promote human rights, democracy and the rule of law**, has accompanied technological developments and the rise of AI over the last decade[12]. At present, it is **preparing a Convention on Artificial Intelligence** that will aim to ensure that the design, development and application of AI are fully consistent with its values. The preliminary **draft shares many concerns with the AI Act proposal**, committing its signatories to take the measures needed to protect the three values mentioned above against abuse. For example, in Chapter III, the draft states that parties must take appropriate measures to preserve the "ability to reach informed decisions free from undue influence [or] manipulation" (Art. 9). Additionally, it **proposes fundamental principles for the design, development and application of AI, such as equality and anti-discrimination, privacy and personal data protection, transparency and**

[12] https://www.coe.int/en/web/artificial-intelligence

accountability, among others. Parties must develop measures **ensuring the availability of redress and other safeguards**, e.g., the right to human review of an AI decision affecting fundamental freedoms or human rights (Art. 20, 1), as well as the right to be informed that one is being attended by AI rather than by a person (Art. 20, 2; and much like the AI Act, Art. 52, 1).

b. USA

Although the US already has laws and regulations on privacy, security and anti-discrimination, there is **still no comprehensive legislation on AI**. The debate is ongoing, but **until very recently it seemed clear that the US government intended any regulation at a federal level** on this matter **to go slowly and not hinder innovation**. For example, while the EU was discussing the AI Act, **seven big tech firms in AI** (Amazon, Anthropic, Google, Inflection, Meta, Microsoft, and OpenAI) were invited to assume **public voluntary commitments for managing the risks posed by AI** while making its development more safe, secure and transparent. Under these voluntary commitments, companies must ensure the safety of their AI products before releasing them in markets (which includes both internal and external testing), share information on managing AI risks with stakeholders, and invest in cybersecurity measures to protect sensitive data.

President Biden's Administration has taken other steps on AI. Last year in October, the White House Office of Science and Technology Policy issued a **blueprint for an AI Bill of Rights**. This is a **non-binding document** that presents a roadmap for the future development, implementation and use of AI compliant with American citizens' rights. It identifies **five core protection principles**: **safe and effective systems** (citizens should be protected from unsafe or ineffective systems); **algorithmic discrimination protections** (citizens should not face discrimination by algorithms and systems should be used and designed in an equitable way); **data privacy** (citizens should be protected from abusive data practices via built-in protections and should have agency over how data about them is used); **notice and explanation** (citizens should know that an automated system is being used and understand how and why it contributes to outcomes that impact them); **alternative options** (citizens should be able to opt out, where appropriate, and have access to a person who can quickly consider and remedy problems they encounter)[13].

On 30 **October 2023**, President Biden presented a new **Executive Order (E.O.) on Safe, Secure, and Trustworthy Artificial Intelligence**. Diverging from recent meek developments on AI regulation, with this E.O., President Biden places the US on a similar path to the EU. **One of the eight core principles indicated by this E.O.** to regulate the development of AI in the US is to promote "**responsible innovation**"[14], a term also used by the Commission in its explanatory memo of the AI Act. Among other policies and political priorities identified in the E.O., President Biden's

[13] https://www.whitehouse.gov/ostp/news-updates/2022/10/04/fact-sheet-biden-harris-administration-announces-key-actions-to-advance-tech-accountability-and-protect-the-rights-of-the-american-public/
[14] https://www.youtube.com/watch?v=fRL1kplm1H4

Administration imposes **new standards for AI safety and security**. According to the White House, **the E.O. "[requires] companies developing any foundation model that poses a serious risk** to national security, national economic security, or national public health and **safety [to] notify the federal government** when training the model and must **share the results of all red-team safety tests"**[15] (Section 4.2 of the E.O). The E.O. also calls for the **US National Institute of Standards to develop "rigorous standards for extensive red-team testing** to ensure safety before public release"[16]. These standards will be applied by several US Government Departments to **critical infrastructures** as well as to address "as well as chemical, biological, radiological, nuclear, and cybersecurity risks"[17]. Additionally, the US Department of Commerce is tasked to develop guidance for **content authentication and watermarking to clearly label AI-generated content**.

The US political system will still have to digest this E.O. (the Constitutional Court may declare it unconstitutional and strike it down, the Congress may pass legislation that supersedes or nullifies the E.O., among other checks and balance mechanisms that may alter the E.O.'s real and full implementation). However, **this E.O. brings the US closer to the EU**, both by showing a **proactive position in terms of AI regulation** and also in **many principles shared between the EU's AI Act and Biden's E.O.**

c. United Kingdom

Unlike the EU approach to AI regulation, the **UK government will not create any new regulators for AI, nor will it opt for horizontal legislation on AI**. Instead, **existing regulators were given five core principles**[18] - safety, security and robustness; transparency and explainability; fairness; accountability and governance; and contestability and redress - **to guide actions on AI**. The UK's idea is to **leverage the expertise of each regulator** within their respective sectors.

The other side of the U.K. approach, as revealed in October 2023, is to **promote a global discussion on AI and its risks.** The U.K. government announced the **creation of an AI safety body in the UK** to evaluate and test new technologies[19]. It will also promote an **AI Safety Summit**, at Bletchley Park, in early November 2023. Here, the U.K. Prime Minister hopes to bring together international governments, leading AI companies, civil society groups and experts to discuss the risks of AI and how they can be mitigated through internationally coordinated action[20].

[15] https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/.
[16] https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/.
[17] https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/.
[18] Provided by UK government in policy paper "A pro-innovation approach to AI regulation" Secretary of State for Science, Innovation and Technology (in https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper#:~:text=Our%20framework%20is%20underpinned%20by,Fairness)
[19] https://www.ft.com/content/509012f9-4e08-414c-a97f-dd733b9de6ef.
[20] https://www.gov.uk/government/topical-events/ai-safety-summit-2023/about.

These certainly are important initiatives. However, in the meantime, **UK sectoral regulators are still left to their own devices and to the five core principles suggested by the government**. Regardless of the U.K. government's initiatives, a **comparison between the UK's vertical and the EU's horizontal approaches** to AI regulation as well as their own **sense of urgency** will provide hints regarding **how AI is to be best regulated** - through fluid regulation in each sector, or using a more centralised and concerted approach. **Will both work? Will both show timely results? The future will tell**.

d. China

China is one of the first countries to draft and implement regulations and specific laws on AI. China's authorities (mainly the Cyberspace Administration of China) have directed their attention to **specific AI uses**, enacting legal instruments such as the **Algorithm Recommendation Regulation** (ARR, which came into force in March 2021, and regulates the use of algorithmic recommendation technologies to provide online services in China), the **Deep Synthesis Regulation** (which came into force on 10 January 2023; deep synthesis technology is commonly referred to as "deepfakes"), the **Generative AI Regulation** (published on 13 July 2023, and came into force on 15 August 2023), and the **Draft Ethical Review Measure** (published on 14 April 2023, for public consultation which closed on 3 May 2023, focused on the ethical review of science and technology activities including AI technologies)[21].

China's approach to AI regulation has a different form (fragmented with several laws for several AI uses) **from the EU approach** (horizontal and comprehensive). However, it **shares many of the concerns and obligations** contained in the current version of the AI Act. For instance, the **Deep Synthesis Regulation** states that labels be clearly and visibly placed on synthetically generated content[22], a concern shared with the EU (Art. 53, AI Act[23]). Moreover, the **ARR** holds operators responsible for establishing a management system that checks "for (…) published information, data security, personal information protections, countering telecommunication network fraud, security assessments and monitoring, and emergency response and handling of security incidents" (ARR, Art. 7).

One **core difference** between the two approaches is the **dimension of the political control** envisioned in each of the regulatory approaches. For example, in the **ARR**, China does **not only restrict, for example, algorithms from displaying illegal content** (Digital Services Act in the EU), but it also **demands as an ethical requirement** for algorithm operators to adjust their recommendation

21 https://www.lw.com/en/admin/upload/SiteAttachments/Chinas-New-AI-Regulations.pdf
22 https://www.allenovery.com/en-gb/global/blogs/data-hub/china-brings-into-force-regulations-on-the-administration-of-deep-synthesis-of-internet-technology-addressing-deepfakes-and-similar-technologies
23 Art. 53, 3: "Users of an AI system that generates or manipulates image, audio or video content that appreciably resembles existing persons, objects, places or other entities or events and would falsely appear to a person to be authentic or truthful ('deep fake'), shall disclose that the content has been artificially generated or manipulated (…)".

algorithms in order **to adhere to "mainstream values"** (ARR, Art. 6)[24], compatible with the political and social order, and to **prohibit the setting "up [of] algorithmic models** that violate laws and regulations, or go against ethics and morals, such as by **inducing users to become addicted or spend too much**"[25] (ARR, Art. 8). Evidently, the Chinese government considers the mass surveillance and social ranking of its population a legitimate activity, while these are considered completely inadmissible under the EU framework.

1.3 The AI Act: Proposal and Trilogue Negotiations

1.3.1 The Regulatory Approach in the AI Act

a. Definitions: The AI Act attempts to provide clear definitions of AI systems to ensure proper regulatory application

Due to the fast-changing environment of AI, one of the **main concerns** in the creation of the AI Act was to **provide a clear definition of AI**. Article 3, 1 of the Act states that the **definitions** are made on a high-level technological basis, **listed in Annex 1**. In this Annex, we can see that these technologies include "machine learning approaches (including supervised, unsupervised and reinforcement learning), using a wide variety of methods including deep learning; logic -and knowledge-based approaches (including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems); and statistical approaches, Bayesian estimation, search and optimisation methods. The use of **these techniques is defined in Article 3 as generating outputs oriented to objectives defined by humans**. The definition provided by the Commission **aims to be as neutral as possible, in order to cover techniques which are not yet developed, and future revisions are foreseen**.

The **importance of the definition of AI was strongly underlined by several stakeholders** consulted in the public consultation launched by the Commission in February 2020, and it is **still under discussion in the trilogue**. In fact, the European Council's compromise version of December 2022 removes "statistical approaches, Bayesian estimation, search and optimization methods". Both the Council and the Parliament replace the term "software" for "system" and "machine-based system", respectively.

b. Risk-based approach: The Act focuses on high-risk AI systems and imposes specific obligations on their developers and users

The EU approach to AI regulation is based on **risk assessments**. This means the **AI Act regulates AI systems differently according to the risk associated with their design, development and use**. As

[24] https://www.china-briefing.com/news/china-passes-sweeping-recommendation-algorithm-regulations-effect-march-1-2022/
[25] https://www.chinalawtranslate.com/en/algorithms/

such, the Commission proposed **three levels of risk for AI systems - unacceptable risk, high risk, and non-high risk**. Title II states that any AI system of **unacceptable risk is prohibited**, with a few exceptions for the use of real-time remote biometric identification systems (Art. 5, n 2, 3 and 4). In case of infringement, the AI Act establishes administrative f**ines of up to € 30 million** or 6% of worldwide annual turnover (if the offender is a company). These are the **highest values foreseen in the Act**.

Title III focuses on high-risk AI systems, both as **a component of a product or a standalone product**. **Annex III contains the list of high-risk AI systems, which the Commission can amend** via delegated acts (Art. 7). It mentions AI systems used for the management and operation of critical infrastructure (transport, water, gas, heating, electricity), for biometric identification of natural persons, both real-time and posterior; educational and vocational training; employment, worker management and access to self-employment (e.g., AI systems intended to be used for recruitment and selection processes, as well as for making decisions on promotions, task allocation or performance evaluation); law enforcement; migration, asylum and border control management; administration of justice and democratic processes.

According to Chapter 2 of Title III, **operators of high-risk AI systems must develop risk management processes** capable of identifying and analysing known and foreseeable risks associated with the specific AI system (Art. 9, 2a); **estimate and evaluate the risks that may emerge** when this system is used **in accordance with its intended purpose and under conditions of reasonably foreseeable misuse** (Art. 9, 2b); **evaluate other possible arising risks based on the analysis of data gathered from post-market monitoring** (Art. 9, 2c); and **adopt suitable risk management measures** appropriate to the sector's harmonised standards and common specifications (Art. 9, 2d and 3). The risk management process must be such as to make **any residual risk acceptable, and must be communicated to its user** (or deployers, using the EP's amended term). Additionally, **data sets** used to train these AI systems models **must be relevant, representative, free of errors and complete** (Art. 10, 3), and **must be subject to appropriate governance and management practices** (Art. 10, 2). Before being placed on the market, operators of **high-risk AI systems must create technical documentation** that demonstrates that these systems comply with the AI Act's requirements and must **provide national competent authorities with all the necessary information to assess the compliance** of the AI system with those requirements (Art. 11, 1 and 2). **Operators also have record-keeping obligations.** These systems must be designed and developed enabling the automatic recording of events. Article 12, 2 states that these **logging capabilities must ensure a level of traceability of the AI system's** functioning throughout its lifecycle.

Non-high risk AI systems may be considered either of minimal risk and will **not be subject to obligations**, **or of limited risk** if they interact with natural persons and pose specific risks of

impersonation or deception (Recital 70). **In the latter case, they are subject to the transparency obligations** mentioned below.

c. Transparency and accountability: The Act promotes transparency in AI system behaviour and requires human oversight. Individuals have the right to be informed when interacting with AI.

The first page of the Commission proposal states that **"[t]his proposal aims** to implement the second objective for **the development of an ecosystem of trust** by proposing a **legal framework for trustworthy AI**". In the context of the interaction between citizens and AI technology, **trust is based on transparency and accountability**. When interacting with high-risk systems, **citizens must be capable of interpreting the system's output** and use it appropriately, and these systems must be designed and developed to ensure this (Art. 13, 1). These systems must **give instructions for usage and provide the users with relevant, accessible, and comprehensible information** concisely and in a clear form (Art. 13, 2). As part of the specific information required in the last paragraph, the Commission obliges high-risk system providers to specify the system's intended purpose and the identity and the contact details of the provider (Art. 13, 3). Furthermore, **high-risk AI systems** must be developed so as they **can be effectively overseen by natural persons** during the period in which the AI system is in use (Art. 14, 1). The human oversight might be performed by the system provider and in some cases by the user (Art. 14, 3). Additionally, the Act states that without prejudice to the requirements and obligations for high-risk AI systems (Recital 70), **those interacting with natural persons must inform them that they are interacting with an AI system** (Art. 52). AI systems which make use of emotion recognition or biometric categorisation, as well as systems generating "deep fakes", must inform the natural person (Art. 52, 2) and disclose that the content is artificially created (Art. 52, 3).

d. Data governance and privacy: The Act addresses data governance and privacy concerns, emphasising privacy-preserving measures and data protection principles.

The **AI Act is to be applied without prejudice to the GDPR**. There are, however, **data protection, privacy, and data governance obligations foreseen in the Act.** Training, validation and testing data sets involved in the development of high-risk AI systems are subject to data governance and management practices (Art. 10). These data sets must be relevant, representative, free of errors and complete (Art. 10, 3). The concern for monitoring, identifying and correcting biases in data sets is present in the Act as a "matter of substantial public interest" (Recital 44). Thus, Article 10, 5 of the **Act allows the providers of high-risk systems to process special categories of personal data covered by Article 9 of the GDPR**. This may be applied in cases where it is strictly necessary to fulfil the above-mentioned purpose, where anonymisation may significantly affect it, and "subject to appropriate safeguards for the fundamental rights and freedoms of natural persons, including

technical limitations on the re-use and use of state-of-the-art security and privacy-preserving measures". This **exception is not to be understood as a free permission for high-risk systems to process special data**. In that respect, Recital 41 states that the AI Act shall "not be understood as providing for the legal ground for the processing of personal data, including special categories of personal data, where relevant".

Personal data is also of special relevance in the articles concerning **regulatory sandboxes for AI**. The AI Act defines these regulatory sandbox functions as to "provide a **controlled environment** that **facilitates the development, testing and validation of innovative AI systems** for a limited time before their placement on the market". If innovative AI systems tested in these sandboxes require the use and processing of personal data, the AI Act determines that this data must "be processed in (…) a functionally separate, isolated, and protected data processing environment" (Art. 54, 1, d), must "not be transmitted, transferred or otherwise accessed by other parties" (Art. 54, 1, e), and must be "deleted once the participation in the sandbox has terminated or the personal data has reached the end of its retention period" (Art. 54, 1, f).

e. Compliance and enforcement: The Act establishes a European Artificial Intelligence Board to oversee implementation and includes penalties for non-compliance

Regarding the enforcement of the AI Act, actors at different levels have different roles. From a national point of view, each **Member State must designate one or more national competent authorities** for the purpose of supervising the application and implementation of the Act. Among the designated national competent authorities, **each Member State shall also designate a national supervisory authority** which "shall act as notifying authority and market surveillance authority" (Art. 59, 2). **Each Member State is responsible for providing adequate financial and human resources** to **guarantee** permanently enough **staff and expertise** (in AI and all the areas covered by the Act) **for the national competent authorities** to fulfil their tasks under the Act (Art. 59, 4). National competent authorities are also responsible for the establishment and supervision of regulatory sandboxes (Art. 53, 1). Furthermore, **Member States must define penalties for infringements** of the Act. These penalties are limited in financial value according to the nature of the infringing AI system (Art. 71, 3 and 4). Member States must also collaborate in the Commission's efforts to "encourage and facilitate the drawing up of codes of conduct intended to foster the voluntary application to [non-high-risk] AI systems (…) of the requirements set out in Title III, Chapter 2" (Art. 69, 1).

The Commission plays a pivotal role in the implementation of the AI Act. First, the **AI Act empowers the Commission to adopt delegated acts to update Annexes** I (Art. 4), III (Art. 7), IV (Art. 12, 3), VI and VII (Art. 43, 5). Member States are obliged to report annually the status of the financial and human resources of their national competent authorities and an assessment of their adequacy (Art.

59, 5). Similarly, **national supervisory authorities must report to the Commission the outcomes of relevant market surveillance activities on a regular basis** (Art. 63, 2). One responsibility attributed to the **Commission** is to, in collaboration with the Member States, **"set up and maintain a EU database containing information concerning (…) high-risk AI systems**" (Art. 60, 1). This database will be **public** (Art. 60, 3) and **controlled exclusively by the Commission** (Art. 60, 5).

Finally, and crucially for the implementation of the Act, the regulation stipulates the creation of a **European Artificial Intelligence Board (EAIB)**. The Board will be composed of **high-level officials of the national supervisory authorities**, the **European Data Protection Supervisor**, and **chaired by the Commission** (Art. 57). Its role is to contribute, assist and advise the Commission and the national supervisory authorities and guarantee a consistent application of the Act (Art. 56). However, there is still a lot of debate within the EU institutions on whether the EAIB will assume the form and nature proposed in the Act or if it will need to take on a more permanent and autonomous existence (see the next section).

1.3.2 The Trilogue Negotiations

a. Definitions, again

The versions brought to the negotiation table by the Commission, Council and European Parliament (**trilogue**) differ in many points. To start with, there is **disagreement on the definition of AI**. In the **three proposed versions there is a general consensus on what AI generates-** - predictions, recommendations and decisions. However, while the Commission proposed a definition of AI formalised in the Annex, both the **Council and the Parliament propose a direct definition in the text** of the Regulation (which also **removes the Commission's power to change this definition**). As well, both suggest changing the definition of AI from software to a system. The European Parliament aligned its definition of AI with the OECD's, maybe hoping to increase the international relevance of the AI Act and create another "Brussels effect".

In the end, the **main point of contention is what will be considered AI and, therefore, what will fall under the AI Act's jurisdiction** (e.g., take the case of the inclusion of generative AI in the later versions). **How AI is defined in the Act is important** because it must avoid two defects in its application. Under an **inflexible definition**, the act may not cover crucial AI technology developed in the near future and, under a **definition that is too wide**, it may be opaque and invite litigation, thus not being applicable.

b. Biometrics and real-time biometric surveillance in public places

The topic of real-time biometric surveillance systems is a key issue in every discussion of the political and social uses of AI, and the trilogue negotiations are no exception. While **both the Commission and the Council included exceptions where this practice could be allowed**, such as the search for

victims of crime or missing children (Art. 5, d, i) or terrorism prevention (Art. 5, d, ii), the **European Parliament banned it entirely**. According to the European Parliament's version, remote biometric identification systems will only be allowed when used ex-post and with prior judicial authorisation (Art. 5, dd).

The **Parliament's vision** for privacy and fundamental rights translates into a **stronger concern for biometrics and its effects** in its version, as compared to the Commission's. While both versions refer to the GDPR's Article 4, the Parliament's version includes definitions of "biometric-based data", "biometric identification" and "biometric verification", which do not exist in the Commission's proposal. Significantly, **it expands biometric categorisation systems from AI systems** able to assign natural persons to specific categories based on biometric data to systems also able to infer categories and attributes from biometric or biometric-based data (Art. 3, 1, 35)). Emotion recognition systems can be used not only for the purpose of identifying or inferring emotions or intentions, but also thoughts and states of mind, of natural persons and groups on the basis of their biometric data or biometric-based data (Art. 3, 1 34)).

c. EAIB vs AI Office

The **European Artificial Intelligence Board (EAIB)** will act as a **coordinating and enforcing force of the Act.** Not only will it advise the Commission on the subject, but it also has many responsibilities of coordination with national supervisory authorities. The form, powers and independence it will have as an institution are crucial to the Act's implementation. This is probably why the **European Parliament's version proposes the replacement of the EAIB with an "AI Office".** The replacement stems from the **idea that** the EAIB proposed by the Commission is insufficient, and that **the implementation of the AI Act needs a more permanent, independent and resourceful body**. According to the EP version of the Act, "[t]he **AI Office should have legal personality**, should act in **full independence**, should be responsible for a number of advisory and coordination tasks, including issuing opinions, recommendations, advice or guidance on matters related to the implementation of this Regulation and should be adequately funded and staffed" (Art. 76). Contrary to the EAIB which would be chaired by the Commission, the **AI Office would be managed by an "executive director** (…) responsible for managing the activities of the secretariat of the AI office and for representing the AI office". The Parliament's **view seems quite different in terms of which workload to expect** from the implementation of the Act to the point that it **even proposes** the **creation of an AI agency** in case an AI Office proves to be insufficient.

d. Penalties for Infringements

As mentioned above, the **Commission** transfers to Member States the powers to define the penalties for infringements of the AI Act. However, the Act also defines **maximum penalties in case of prohibited systems** (Art. 71, 3: "(…) administrative fines of up to **30 000 000 EUR or, if the**

offender is company, up to 6 % of its total worldwide annual turnover for the preceding financial year"); and for **high-risk system infringements** connected to data governance (Art. 71, 4: "(…) administrative fines of up to **20 000 000 EUR or, if the offender is a company, up to 4 % of its total worldwide annual turnover for the preceding financial year**"). Additionally, the Act also stipulates "administrative fines of up to **10 000 000 EUR** or, if the offender is a company, up **to 2 % of its total worldwide annual turnover for the preceding financial year**" for the cases when "**incorrect, incomplete or misleading information [is supplied] to notified bodies and national competent authorities** in reply to a request" (Art.71, 5).

Both the versions of the **Council and the Parliament show differences**. For instance, the Council agrees with the limit on penalties for non-compliance of prohibited systems but **lowers the limit for SMEs and start-ups** (3% of the SME total worldwide annual turnover for the preceding financial year). The **protection of SMEs** is also present in the cases of **misreporting to notified bodies and national competent authorities (lower limit of 1%) and of non-compliance** reported in Article 71, 4 (limit of 2%). Contrasting with the Council's focus on SMEs, the **European Parliament's main criterion to define penalties is the risk level.** For **prohibited system infringements** (Art. 5), it **raises the limit** to "40 000 000 EUR or, if the offender is a company, up to 7 % of its total worldwide annual turnover for the preceding financial year". Infringements related to Articles 10 and 13 (high risk system obligations of data governance and transparency) are penalised with "administrative fines of up to EUR 20 000 000 or, if the offender is a company, up to 4% of its total worldwide annual turnover for the preceding financial year". Note that in the **Parliament's version the penalty limit of € 20 million is targeted not only to data governance infringements** (Art. 10), **but also to transparency and provision of information to users** (Art. 13). Any infringement to compliance with any article other than Articles 5, 10 and 13 will be subject to penalties no greater than € 10 million or 2% of the company's total worldwide annual turnover for the preceding financial year (both values are half of what the Commission proposed for these cases). Finally, the **Parliament also cuts by half the penalty limit for cases of misreporting** to notified bodies and national competent authorities **and does not include any benefits to SMEs**.

e. Foundational models and General Purpose AI (GPAI)

The fast growth of AI markets and AI capabilities is a challenge for future regulation. The fact that the European Parliament and Council versions were written one year after the Commission's proposal is not without consequences. One of these consequences is the **absence of references to "foundational models", "general purpose AI", "generative AI"** or any specific AI models **in the Commission version**. The fact that **this absence is important** is because, contrary to the conception of AI that justified the use case approach enshrined in the Commission proposal, **foundational**

models and GPAI have a **multipurpose nature and capabilities** that escape many of the Commission's proposed obligations.

One of the core concerns of both the Council and the Parliament is the **downstream regulation**, i.e., the **regulation across the AI value chain**. To make this part of the regulatory discussion clearer, the Council stated in a recent preparatory document for the fourth trilogue session (held on24 October 2023)[26] that "certain tailored transparency obligations are necessary to ensure that downstream providers can build AI systems (including general purpose AI systems) on foundation models in a way that is safe and compliant with the AI Act, minimising the risk to violate fundamental rights and safety". **This is no minor issue**, because, **if left unregulated**, the AI value chain may suffer from a lack of transparency and information between its constituents. Furthermore, **unclear liability and responsibility attributions may also lead smaller enterprises and SMEs to hold back from optimally developing their businesses** using specific GPAI Although these models may be the best for their businesses, it **may be too risky to rely on non-transparent and potentially non-complying providers of GPAI**[27].

In the above-mentioned document, the **Council also defends the application of obligations for all foundational models,** assuming both a **before-market** (such as documenting the model and training process, including the results of internal red teaming, and carrying out and documenting model evaluation in accordance with standardised protocols and tools) **and after-market nature** (such as providing information and documentation to downstream providers and enabling them to test the foundation models). **These obligations are augmented for what the Council defines as "very capable foundation systems"**[28] and GPAI used at scale, i.e., with **regular external red-teaming**, the deployment of a **risk assessment and mitigation system**, and for the case of very capable foundation models compliance with **additional ex-post market controls**. Moreover, the Council also mentions the need to introduce obligations to support enforcement of copyright protections as well as obligations to ensure transparency of AI-generated content.

The European Parliament restricts its regulation approach to foundation models (instead of the Council's overarching regulation of GPAI). According to the **Parliament's version**, **providers of foundation models will have to comply** with a set of obligations such as **data governance measures**, **performance levels, requirements for energy use, technical documentation, and compliance with certain transparency requirements** (Art. 28b of the EP version).

[26] https://table.media/europe/wp-content/uploads/sites/9/2023/10/2023-10-17-conseil-ia-mandat-de-negociation-10412dc9fadd4e4fa9b0360960fd13af.pdf
[27] This concern has been reported in Bienert et al., 2023.
[28] According to the Council: "Very capable foundation models should be understood as foundation models whose capabilities go beyond the current state-of-the-art and may not yet be fully understood".

Regardless of the results of the trilogue negotiations, **it is likely that foundational models and GPAI will be explicitly regulated.** In fact, the EP version states that "foundation models are a new and fast-evolving development in the field of artificial intelligence, **it is appropriate for the Commission and the AI Office [or the EAIB] to monitor and periodically assess the legislative and governance framework of such models**" (Recital 60h). The Council proposes obligations on GPAI that are similar to those for high-risk AI and calls for the Commission to produce an implementing act that further sets out the specific requirements for GPAI (Art.4b, 1 of the Council version).

f. Goal definition of the systems and classification of high-risk AI systems

Contrary to the Commission version where **high-risk AI systems** are to be classified as such if they match the criteria in Annex III, **the Council and Parliament provide more specifications for this classification.** For example, **AI systems classified as high-risk** by the Commission **are excluded by the Council if** "the **output of the system is purely accessory** (…) and is not therefore likely to lead to a significant risk to the health, safety or fundamental rights" (Art. 6, 3). Additionally, the **Parliament's version** allows for companies developing high-risk AI systems which these do not consider posing a significant risk (according to the spirit of the Regulation) "**to submit a reasoned notification to the national supervisory authority they are not subject to the requirements**" (Art. 6, 2). After receiving the provider's request, "the **national supervisory authority shall review and reply** to the [it], directly or via the AI Office, (…) **if they deem the AI system to be misclassified**". The Parliament's version thus implies the attribution of a filtering power to national supervisory authorities. Both versions of the Council and the Parliament seem to defend a **filter-based system for high-risk AI classification.** The core idea is to prevent a strict and innovation-repelling classification system, too rigid to comply with in many real-life situations. The **downside of a filtering system** such as the one proposed by the Parliament or the Council could be **legal fragmentation and the creation of high-risk friendly jurisdictions within the EU**, which is precisely what this Regulation is trying to fight.

1.4 Implementing the AI Act

1.4.1 Impact and types of regulatory intervention in the EU

According to the EU **"New Legislative Framework"** of 2008, manufacturers must conduct pre-market controls, ensuring product safety and performance through conformity assessments (CA) according to specific (and essential) requirements defined by law. The idea is based on the **understanding that the providers' in-depth understanding of the design and production process makes them the most suitable party to guarantee the conformity of their products** with regulatory requirements. Following this logic, in the AI Act **the identification of an AI service's risk level** and

compliance with regulatory requirements **is left to the responsibility of the provider**[29]. AI firms themselves are therefore called to participate in the process of classifying their systems. As such, **while Chapter 2 of the AI Act sets out the legal requirements** "for high-risk AI systems in relation to data and data governance, documentation and recording keeping, transparency and provision of information to users, human oversight, robustness, accuracy and security", **it leaves the "precise technical solutions to achieve compliance with those requirements"**[30] **to the discretion of the AI provider.** The Act leaves the **CA to internal control checks or notified bodies** (entities which must be involved as independent third parties (Art.33, 4), satisfy specific criteria (Art. 30, 1 and 2) and be designated by national notifying authorities (Art. 30)) depending on the type of high-risk AI system to be assessed. This includes using **internal controls for stand-alone AI systems** and employing **third-party CAs for AI systems** intended to be used as **safety components** of products **regulated under the New Legislative Framework** legislation. The Act also explains in detail the CA procedures for each of these types (Ch. 5). CAs must be performed both before the AI system is deployed in EU markets or before putting the AI system into service (Art. 19) and (for cases where the providers are distributors or importers) when a high-risk AI system is substantially modified (Art. 43, 4). **AI product manufacturers also have compliance obligations** in specific cases (Art. 24).

According to Recital 64, the envisioned **general rule for the CA** process of stand-alone high-risk AI systems consists of **internal controls and checks when applicable** (that is, excluding "AI systems intended to be used for the remote biometric identification of persons"). On the other hand, according to Recital 63, for high-**risk AI systems related to products** covered by existing Union harmonisation legislation, the compliance of those AI systems with the Act should be assessed as part of the conformity assessment already foreseen under that legislation. With this Act, the **Commission wants these types of products** (e.g., machinery, toys, medical devices, etc.) to be **subject to the same ex-ante and ex-post compliance mechanisms** of the products of which they are a component, **but now** ensuring compliance **both with sectoral legislation and AI Act** requirements.

1.4.2 Subsidiarity structure at the EU level

a. Legal form of "Regulation"

The legislative instrument chosen by the Commission is not random. A **Regulation** (instead of a directive) **harmonises the regulatory framework** within the EU and **avoids legal fragmentation**

[29] This must not be mistaken with the idea that solely AI systems designers and developers must comply with these requirements. Even if the provider is not the designer/developer of the system, AI providers must guarantee that Ch. 2 requirements are embedded in the system to be compliant. As we have seen in Section 3. B. 5), part of the trilogue discussion is focused on levelling fairly the compliance burden and responsibility all along the AI value chains, especially in the case of foundation models and GPAI.

[30] AI Act, Explanatory Memo, 5.2.3, p.13.

between EU Member States. There are, however, some issues left to each Member State's discretion, such as the definition of penalties for infringements of the AI Act (explained below) or the application of AI for military purposes.

b. European Artificial Intelligence Board

Both the EAIB and an AI Office will impact the AI Act implementation. As we tried to show in the trilogue discussion part, the EU institutions have different views on the workload this entity will have. However, the **AI Act clearly defines the powers, constitution and decision-making process to define rules of procedure of the EAIB** if it is to attain the "OK" from all EU institutions. For example, the **EAIB adoption of its rules of procedure will be decided by simple majority** (Art. 57). These rules of procedure will "contain the operational aspects related to the execution of the Board's tasks as listed in Article 58". Of course, the nature of the EAIB tasks is mainly advisory and coordination (Recital 76 and Art. 58), however, **is simple majority the right way to decide?** Is this the best way to converge different interests among EU Member States? This may be refuted by the need for the Commission's consent for the adoption of these rules. Besides all the other supervisory tasks the Act attributes to the Commission, its role in the implementation of the Act is again reinforced.

c. Database of High-Risk AI Systems

Title VII of the AI Act establishes the creation of an **EU database of registered high-risk AI systems** to be managed by the Commission (Art. 60, 1). The database **will be publicly accessible** (Art. 60, 3), high-risk AI system providers must register (Art. 51), and national competent authorities contribute as well. The Commission will provide all the technical and administrative support to the providers to carry out their system registration. The data available in this **database will contain the data mentioned in AI Act Annex VIII** (Art. 60, 2), as well as **personal data** regarding the "**names and contact details** of [the] **natural persons who are responsible for registering the system** and have the **legal authority to represent the provider**" (Art. 60, 4).

d. Definition of penalties

As described above, the **definition of penalties** is a topic that is being discussed among the EU institutions. The three institutions seem to agree that risk is the overarching idea that must define the penalty limit (signalling effect), but they must still agree on some points. The **Council version wants to include adaptations to penalties to SMEs** and the **Parliament wants to increase the penalty limits for infringements of transparency and provision of information to user obligations** (Art. 13).

The AI Act system to define penalties, at the national level, seems to be taken from other previous acts, such as the Data Governance Act and the Data Act. However, the problems may still be the same. In the long run, **competition among countries within the EU may arise to soften their AI Act**

penalties, in order to **attract AI providers and developers** and thus promote innovation and value inside their borders. This possibility **may also happen unintentionally** since different Member States have **different financial and human resources**. Of course, the Commission's idea to avoid this is to create the EAIB (or an AI Office) to guarantee consistent application and the dialogue between national supervisory authorities. However, the same Act requires that each Member State must fully fund and equip their national competent authorities to carry out their tasks (Art. 59, 4).

1.4.3 Human oversight

Human oversight is an important requirement of the AI Act **for high-risk systems** (Art. 14) and should be a **crucial component** of the development of **human-centric AI.** However, the requirement **still lacks clarity**. The meaning of human oversight for regulating AI is still under discussion. Core questions remain to be answered: **What and how is it to be supervised**? **When will the supervision take place? By whom?** Take the following example discussed by Enqvist (2023): AI Act article 14, 3 calls not only on the AI designer but also the AI system user for oversight. At the same time, according to Recital 48: "High-risk AI systems should be designed and developed in such a way that natural persons can oversee their functioning. For this purpose, appropriate human oversight measures should be identified by the provider (…) Such measures should guarantee that the system is subject to in-built operational constraints that cannot be overridden by the system itself and [are] responsive to the human operator". The same recital ends by stating these measures must also guarantee "that the natural persons to whom human oversight has been assigned have the necessary competence, training and authority to carry out that role". **These points seem to leave a lot of room for interpretation on the core questions asked above.** How can AI system designers take necessary measures to guarantee that the person "to whom human oversight has been assigned" (Recital 48) has the necessary competence, training and authority to perform the oversight of the AI system? Contrary to what this recital states, Article 9, 4 includes a provision which states that after "eliminating or reducing risks related to the use of the high-risk AI system, due consideration shall be given to the technical knowledge, experience, education, training to be expected by the user". **An obligation of due consideration and guarantees are not substitutes for each other**.

The human oversight provisions for high-risk AI systems **may also conflict with the GDPR.** The latter states that **every citizen has the right not to be subject to decisions made by machines (Art. 22 of the GDPR).** The **European Parliament** concluded that the **AI Act may be declaring a redundant obligation,** if not contradicting the GDPR, and **proposed** the inclusion of a new clause 4a that contains a **general principle of human oversight applicable to all AI systems**. Its proposal

maintained the clause "decisions on **specific areas identified in Annex III** must be **subject to human oversight of at least 2 natural persons**", as did the draft of the Council.

1.4.4 Competent National Authorities' actions and identity

Member States can (but need not) create new authorities or designate existing ones as national competent authorities for the purpose of ensuring the application and implementation of the AI Act. Out of the competent authorities, one will be designated as a national supervisory authority serving as the link between the Member State's competent authorities and EU-level authorities connected to the Act (Commission and EAIB/AI Office). By default, the national supervisory authority "shall act as notifying authority and market surveillance authority" unless a Member State communicates to the Commission "organizational and administrative reasons to designate more than one authority" (Art. 59, 2 and 3).

In terms of **budgetary implications**, the implementation of the **AI Act will require sufficient technological expertise and human and financial resources** which could amount **between 1 to 25 FTE per Member State**. Much of the implementation costs will be directly influenced by each Member State's current institutional setup and nominations of competent authorities.

No specific authorities or types of regulators are indicated in the AI Act as national supervisory or competent authority. Again, this decision is left to each Member State. **However**, two facts may lead one to believe that the task of implementing the act will be left to Data Protection Authorities (DPAs). Firstly, the **European Data Protection Supervisor is represented in the EAIB** (Art. 57, 1), consistently with its responsibility to act as the competent authority to supervise Union institutions, agencies and bodies (Art. 59, 8). Secondly, **in the EU and globally, data protection authorities have defined policies and taken action regarding AI** (e.g., the ban of ChatGPT by Italy's DPA[31], and the French DPA's plan on protecting personal data in the development of AI models[32]).

1.4.5 Impact of compliance for businesses

The limitation of rights in the AI Act are based on what the Commission calls "**responsible innovation**". This means that the Act restricts some fundamental liberties, such as the **freedom to conduct business** or the **freedom of art and science**, in order to prioritise "overriding reasons of public interest", including health, safety, consumer protection and the safeguarding of fundamental rights. These restrictions are designed to regulate the development and use of AI technology, with a primary focus on preventing and mitigating serious safety risks and likely violations of fundamental rights such as those described above. The **Commission attempts to** keep the **limitations**

[31] https://www.euractiv.com/section/artificial-intelligence/news/italian-data-protection-authority-bans-chatgpt-citing-privacy-violations/
[32] https://www.cnil.fr/en/artificial-intelligence-action-plan-cnil

proportional to identified risks, ensuring that its constraints are **reasonable and necessary**, **without unduly impeding innovation** or **damaging the functioning of businesses**.

Moreover, the proposal emphasises the **need for increased transparency requirements**, which are intended to **balance the right to protect intellectual property with the imperative of providing necessary information for individuals to exercise their right to an effective remedy**. These transparency measures are also designed to ensure transparency towards supervisory and enforcement authorities in line with their mandates according to this regulation.

According to the impact assessment published by the Commission in April 2021, **AI compliance costs will be in the range from €1.6 billion to € 3.3 billion in 2025 (p. 12) and over € 31 billion in the next five years (2023-2028)[33].** To compute this value, the **Commission assumed a total of 10% of high-risk AI systems in all of the AI systems landscape**. A **recent study** on the practical perspective of risk classification according to the AI Act **identified, from a sample of 100 AI systems, 18% as high-risk and 40% as unclear.** Thus, the **10% proportion of high-risk AI systems may be way too conservative**, and **AI compliance costs may be much greater**. **Compliance costs only** (excluding, for example, conformity assessment costs) would **represent nearly 17% of total AI investment costs** (p.166). It is therefore not surprising that one of the main concerns of the EU institutions while drafting the AI Act is to minimise its burden.

On the one hand, **SMEs can expect** total compliance costs of up to **€ 400,000 for just one high-risk AI product requiring a quality management system**[34]. To provide an idea of the impact, this cost can represent a **40 % reduction in profit for a European business with a € 10 million turnover** (excluding the cost of the AI system itself). On the other hand, **SMEs are protected in the Act in several clauses.** For example, SMEs will benefit from **priority access to AI regulatory sandboxes** (Art. 55, 1a); will **have their "specific interests and needs** (…) taken into account [by notified bodies] **when setting the fees for conformity assessment** under Article 43, **reducing those fees proportionately to their size and market size**" (Art. 55, 2); **will receive guidance** and advice on the implementation of the Act by national competent authorities (Art. 59); and will have their interests and economic viability taken into account by Member States when the latter define rules on penalties (Art. 71). Part of the compliance burden may be determined by **how heavily other international blocs regulate their AI.** Additionally, the Commission's impact assessment also states that the **compliance "costs for SMEs could be significantly reduced by sharing systems** (e.g., for testing or legal advice)" and that technical and administrative assistance may help to reduce these costs significantly for SMEs (which under Art. 59 should be guaranteed by national competent authorities).

[33] https://www2.datainnovation.org/2021-aia-costs.pdf
[34] https://www2.datainnovation.org/2021-aia-costs.pdf

Businesses other than SMEs are expected to incur **significant costs in complying with the AI Act**. Companies operating with AI high-risk systems will have to establish a quality management system, create and manage technical documentation, conduct (or pay for) a conformity assessment (with regular reviews for significant AI system modifications), implement human oversight and continuous monitoring to mitigate potential risks, and ensure compliance with other sectoral laws, the GDPR, and other relevant regulations. If we **compare with GDPR implementation costs** (**which should be lower, as the AI Act seems to define more requirements than the former**), a recent EU study reports that **34% of large companies spent more than 1 million USD to implement the GDPR**, while 74% of SMEs more than 100.000 USD[35].

1.5 Looking forward

1.5.1 Future-proofness of the AI Act and the Commission's powers

All **three EU institutions understand** that the **future-proofness of the AI Act must be guaranteed and allow for later adjustments** and implementing acts that adapt the Act to current developments. This comes as no surprise. **Other international initiatives on AI regulation are developing in a much less holistic and horizontal fashion.** Both the lack of a comprehensive regulation of AI in the **US**, as well as the **UK's** (leave regulation to each sector regulators) and **China's** approach (specific use cases, such as deep synthesis or algorithmic recommendation) show the **caution these key international players exercise when regulating AI**. Among other factors, one issue is the rapid development of AI technologies and systems that are being produced every day. This speed has already had consequences in the EU legislative process on AI -- the Commission version of the Act did not yet even know of foundation models and GPAI. The Act will for sure require the ability to be updated if it is to endure. The result of leaving important points (such as the list of high-risk AI systems and other compliance specifications) to the annexes is to delegate their updating exclusively to the Commission and its bodies. If this principle is to be maintained, it strengthens the EP's request of establishing an autonomous executive body (such as the AI Office) instead of the EAIB.

1.5.2 Trade-off between legal certainty and restrictions on business models

a. AI providers in the EU

For **AI providers** in the EU, the AI Act represents **two opposing forces**. On the one hand, the Act aims to improve **legal certainty on requirements and compliance. This is positive** – although there are still many points to be decided and clarified in the trilogue, AI may attract a **lot of investment**,

[35] European Commission (2021), p. 161

which may be held back due to uncertain liability and compliance rules. On the other hand, the application of **rigid, quickly outdated and burdensome regulation may incentivise AI providers to seek opportunities outside the EU instead**. The Act and its annexes include several provisions to simplify compliance, specifying its requirements, as well as providing regulatory sandboxes (the possibility of testing AI systems outside these sandboxes is being discussed in the trilogue). In the long term, because the AI Act is globally the first horizontal and comprehensive AI regulation, the **size of this deterrence effect will depend on the virtues and vices of competing regulations in other international blocs**, as well as on the **EU's and Member States' incentives and policies supporting AI**. The future will tell.

b. AI users in the EU

From the point of view of **AI users**, the AI Act will promote their **fundamental rights and liberties**, because the Act guarantees compliance with EU requirements for the AI systems they use. While there are benefits (such as mandatory transparency and information provision requirements for most AI systems), users that deploy AI systems (and, therefore, must comply with the Act) have now a clearer idea of their share of responsibility in the AI value chain.

1.5.3 Brussels effect, or common regulatory approach US-EU-China?

With the AI Act, the Commission hopes to repeat the regulatory success of setting the agenda and the rules with the GDPR: the "**Brussels effect**". Will there be another one? The **benefits** of a Brussels effect would be clear - **rules in international markets harmonised with those in the EU**, and **promotion of EU principles and human rights worldwide in what concerns AI**. This would be a **big deal**, due to AI's potential for both economic growth and political misuse.

There are **two reasons that point towards a future Brussels effect** in AI. First, both the GDPR and the **AI Act have extra-European scope**, that is, both **force companies outside the EU** and international actors to **comply with EU principles** in order to participate in EU markets. Second, although other jurisdictions are already starting to regulate aspects of AI, the AI Act is the first comprehensive AI regulation. This is important because, in the case of **prolonged regulatory inaction of other actors** such as China or the US, the **AI Act may start to impact and even transform companies outside of the EU** and their business models, **increasing the opportunity costs of other actors' regulatory initiatives** if they contradict the EU's AI Act.

There are, however, **several factors that reduce the likelihood of a renewed Brussels effect**. First, in contrast with the GDPR, the AI Act may involve more **burdensome compliance costs**, which can deter international AI companies from acting and investing in EU markets. Second, the **majority of big AI companies are found outside the EU**, such as in the UK, China and the US, whose governments also seem to be quite a bit more receptive to their lobbying against a tighter

regulation. This leverage strongly increases the **likelihood of a London/Beijing/Washington effect** rather than a Brussels effect. It is important to note that the **different blocs still have different ideas on how (and when) to regulate AI**. Recently, the **UK** Prime Minister asked in a public address regarding AI regulation: "**How can we write laws** that make sense **for something that we don't yet fully understand**?"[36]. Less than a week later, in the context of President Biden's Executive Order on AI Safety, **White House Chief of Staff, Jeff Ziently**, declared publicly "given the pace of this technology [AI], we can't move in normal government or private-sector pace, **we have to move fast, really fast – ideally faster than the technology itself**"[37]. These two statements reflect the **two different positions of the US and the UK** but also reveal the **uncertainty** regarding the shape an international regulatory framework for AI will take and who will influence what.

Yet, regardless of the epicentre of regulatory influence, **converging interests could lead to converging international regulatory frameworks**. This **convergence may be delayed** or blocked by **three factors**. First, **competition to attract innovative AI companies** and technologies. Each bloc may feel compelled to regulate as little as possible and to promote public policies and investments to stimulate the development of AI within their jurisdiction. Second, **the role of AI for military purposes**. The absence of compliance requirements for military purposes and the military industry in the EU AI Act is no chance omission. Finally, as we can see if we compare the EU approach to China's, AI regulation in each bloc is designed to protect **different social and political principles** (individual liberty and rights, or social order and patriotism, apart from innovative activity). There may be little room for agreements that can bridge these fundamental differences.

However, **these three factors have potential arguments against them**. First, these **countries share economic interests and benefit from the harmonisation of the markets**, opening their economies to new customers. Second, international convergence of the AI regulatory landscape **will increase transparency and control over foreign AI systems**, which will contribute to a greater control and better implementation of local regulations on AI.

In conclusion, **international regulatory frameworks on AI may converge**, similar to what happened with the GDPR, but this **time it is likely that convergence will not be centred around Brussels**, EU legal requirements, or even the EU's social and political principles.

[36] https://www.ft.com/content/509012f9-4e08-414c-a97f-dd733b9de6ef
[37] https://edition.cnn.com/2023/10/30/politics/white-house-tackles-artificial-intelligence-with-new-executive-order/index.html

Chapter 2: The impact of generative AI

2.1 Introduction to generative AI

Generative AI is an advanced form of artificial intelligence that enables machines to learn from existing data to create new data or content, including audio, code, images, text, simulations and videos[38]. Therefore, it is a new technological frontier that has the potential to drastically change the way each type of content is created[39]. At the core of generative AI are **machine learning** and **deep learning techniques**. These techniques use deep neural networks that are trained on large data sets to recognize patterns and generate new information based on those patterns[40]. Output from generative AI models can be indistinguishable from human-generated content[41].

The key **difference between generative AI and analytical or traditional AI** lies in the ability to create new content. Traditional AI also might use neural networks, but these models are not designed to create new content. They can only describe, predict, or prescribe something based on existing content or data. Companies use "traditional" AI, for example, to predict client churn, forecast product demand, and make next-best-product recommendations[42]. On the contrary, generative AI can be used to create content or data that does not exist yet.

There are **many applications and use cases of generative AI** by different modalities.

[38] McKinsey & Company, What is generative AI?, January 2023
[39] https://medium.com/@makcedward/generative-ai-and-examples-bdb06d6a5ff6
[40] Forbes, Unlock The Potential Of Generative AI: A Guide For Tech Leaders, January 2023 - https://www.forbes.com/sites/forbestechcouncil/2023/01/26/unlock-the-potential-of-generative-ai-a-guide-for-tech-leaders/
[41] McKinsey & Company, What is generative AI?, January 2023
[42] McKinsey & Company, Exploring opportunities in the generative AI value chain, April 2023

Table 1: Many applications and use cases of generative AI across modalities

Modality	Application	Example use cases
Text	Content writing	• Marketing: creating personalized emails and posts • Talent: drafting interview questions, job descriptions
	Chatbots or assistants	• Customer service: using chatbots to boost conversion on websites
	Search	• Making more natural web search • Corporate knowledge: enhancing internal search tools
	Analysis and synthesis	• Sales: analyzing customer interactions to extract insights • Risk and legal: summarizing regulatory documents
Code	Code generation	• IT: accelerating application development and quality with automatic code recommendations
	Application prototype and design	• IT: quickly generating user interface designs
	Data set generation	• Generating synthetic data sets to improve AI models quality
Image	Stock image generator	• Marketing and sales: generating unique media
	Image editor	• Marketing and sales: personalizing content quickly
Audio	Text to voice generation	• Trainings: creating educational voiceover
	Sound creation	• Entertainment: making custom sounds without copyright violations
	Audio editing	• Entertainment: editing podcast in post without having to rerecord
3-D or other	3-D object generation	• Video games: writing scenes, characters • Digital representation: creating interior-design mockups and virtual staging for architecture design
	Product design and discovery	• Manufacturing: optimizing material design • Drug discovery: accelerating R&D process
Video	Video creation	• Entertainment: generating short-form videos for TikTok • Training or learning: creating video lessons or corporate presentations using AI avatars
	Video editing	• Entertainment: shortening videos for social media • E-commerce: adding personalization to generic videos • Entertainment: removing background images and background noise in post
	Voice translation and adjustments	• Video dubbing: translating into new languages using AI-generated or original-speaker voices • Live translation: for corporate meetings, video conferencing • Voice cloning: replicating actor voice or changing for studio effect such as aging
	Face swaps and adjustments	• Virtual effects: enabling rapid high-end aging; de-aging; cosmetic, wig, and prosthetic fixes • Lip syncing or "visual" dubbing in post-production: editing footage to achieve release in multiple ratings or languages • Face swapping and deep-fake visual effects • Video conferencing: real-time gaze correction

Source: McKinsey & Company (2023)

Generative AI will probably affect most business functions over the longer term. Nowadays, information technology, marketing and sales, customer service, and product development are the most fertile fields for the first wave of applications.

Some of **the potential of generative AI** in the areas mentioned above are listed below[43]:

1.Information technology. Generative AI can help teams write code and documentation. According to preliminary estimates, automated coders on the market have improved developer productivity by approximately 50%, helping to accelerate software development.

Marketing and sales. Teams can use generative AI applications to create content for customer outreach. Within two years, 30% of all outbound marketing messages is expected to be developed with the assistance of generative AI systems.

2.Customer service. Natural-sounding, personalized chatbots and virtual assistants can handle customer inquiries, recommend swift resolution, and guide customers to the information they need.

Product development. Companies can use generative AI to rapidly prototype product designs. Life sciences companies, for instance, have already started to explore the use of generative AI to help generate sequences of amino acids and DNA nucleotides to shorten the drug design phase from months to weeks.

In general, the primary **benefit of generative AI** is its ability to quickly produce high-quality content with minimal human effort required compared to traditional methods such as manual coding or writing scripts from scratch. This technology can help reduce costs associated with content production. Generative models can also be used for tasks such as natural language processing (NLP), image recognition/generation and robotics/automation applications, which could lead to improved customer experiences across various industries including healthcare and retail sectors[44].

Not only large enterprises but also SMEs are eager to explore the possibilities that generative AI can offer to accelerate their business growth. Already use cases are emerging that illustrate the potential of generative AI to help small businesses increase efficiency, reduce costs, and improve their marketing and customer services efforts.

In addition, SMEs (like large enterprises) can harness the power of generative AI to improve cybersecurity resilience[45]:

3.Anomaly detection. Generative AI can be used as a tool to discover patterns and behaviours of normal network traffic and user activities or system operations within IT infrastructure.

Rapid monitoring. Generative AI can help a security analyst doing the work to reason over the massive data stores and detect and respond faster.

Automated response. Generative AI can trigger computerized responses, such as isolating affected systems and blocking suspicious IP addresses. It can also guide the user on taking the right action,

[43] Ibidem
[44] Forbes, Unlock the Potential Of Generative AI: A Guide For Tech Leaders, January 2023 - https://www.forbes.com/sites/forbestechcouncil/2023/01/26/unlock-the-potential-of-generative-ai-a-guide-for-tech-leaders/ (Last access: 6 July 2023)
[45] https://www.weforum.org/agenda/2023/07/generative-ai-small-medium-sized-business/

using the right tools, and setting up those types of automation, regardless of which technology the customer has implemented.

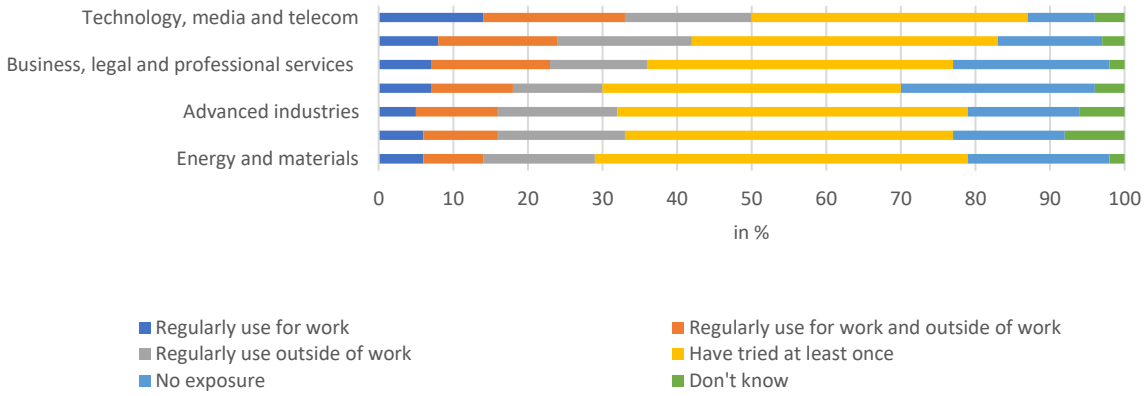
4.Vulnerability assessment and patch management. By simulating potential attack scenarios, generative AI can help prioritize vulnerabilities based on their business impact and recommend effective patch management strategies.

Faster learning. Generative AI can enhance the education and a quicker understanding of the people they have working in IT and security. Generative AI is not doing all the work for them, but enhancing what they can do with the tool.

However, policy actions are needed to ensure SMEs do not fall behind in the adoption of these new tools, further widening the gap between small and large firms. For instance, existing divides, such as the skills gap, will need to be addressed to enable SME adoption. SMEs will need skilled personnel to integrate this tool effectively, and to understand its limits as well as its potential[46].

The latest annual McKinsey Global Survey on the current state of AI confirms **the explosive growth of generative AI**. According to this survey, 33% of respondents from **technology, media and telecom industries** regularly use generative AI for work, or outside of work, while 37% of respondents said they have used this technology at least once. Two other industries using these newer tools are **financial services** and **business, legal and professional services** where nearly one-quarter of respondents regularly use generative AI.

Figure 1: The use of generative AI tools in industries



Source: McKinsey & Company (2023)

The most famous generative AI tool is certainly **ChatGPT**[47], a natural language processing (NLP) application based on generative models. ChatGPT is trained on large datasets and is able to generate

[46] OECD (2022)
[47] https://openai.com/blog/chatgpt

new text based on existing text, making conversations with the chatbot more natural and engaging[48]. It has recently emerged as a powerful tool for various use cases such as content creation, translation, web scraping, text summarization, code debugging, question answering, etc. For example, businesses can leverage ChatGPT for content creation to streamline and enhance their marketing efforts, enabling them to generate engaging, high-quality content more efficiently. By incorporating ChatGPT into their content strategy, companies can automate the production of blog posts, articles, social media posts, and promotional marketing materials tailored to their target audience. Additionally, ChatGPT can assist in optimizing content for search engines through keyword research or content structuring, ensuring increased visibility and improved search rankings that can be especially useful for small businesses to create brand awareness.

Moreover, businesses can harness ChatGPT for translation services, enabling seamless communication across linguistic barriers in today's increasingly globalized market. By integrating ChatGPT into their operations, companies can gain access to real-time, accurate translations for various content types, such as emails, reports, marketing materials, and product documentation.

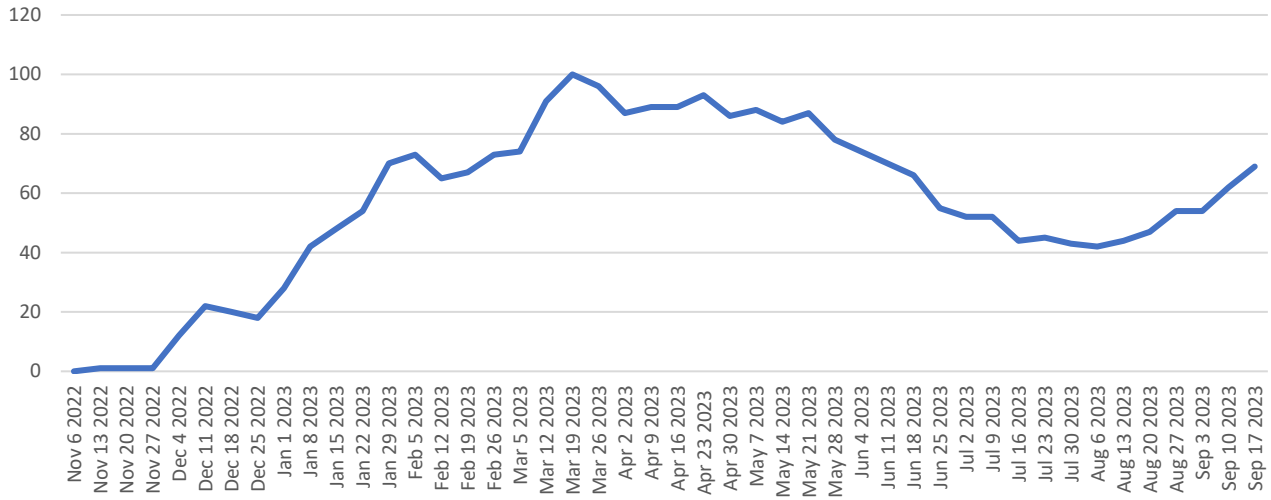
Another top use case of ChatGPT for business purposes is the potential to generate ideas and facilitate brainstorming sessions. By integrating ChatGPT into brainstorming sessions, employees can input their initial ideas or problems, and the model can generate related concepts or potential solutions based on the given context. Furthermore, ChatGPT can assist in refining ideas and proposals, offering feedback and suggestions to enhance the quality and feasibility of those ideas.

Finally, ChatGPT can assist in employee onboarding, providing essential information and guidance to new hires, and facilitating the orientation process. It can also help in forming interview questions for job positions[49].Therefore, companies are interested in researching how to implement ChatGPT in their operations, from marketing to human resources.

Since its launch in November 2022 by the U.S.-based OpenAI, **ChatGPT has sparked a lot of interest and immediately experienced a revolutionary growth** triggering a worldwide chain of innovation in artificial intelligence[50]. As of September 2023, global Google searches for the word "ChatGPT" increased again after a slight decline during the summer months. Interest in the chatbot started rising in the week starting on December 4, 2022. The growing demand for information on ChatGPT made the keyword hit a peak of 100 index points during the week ending on March 19, 2023[51].

[48] https://medium.com/@makcedward/generative-ai-and-examples-bdb06d6a5ff6
[49] https://research.aimultiple.com/chatgpt-for-business/#:~:text=How%20to%20Use%20ChatGPT%20for%20Business%20in%202023%3A,...%208%209-%20Web%20scraping%20...%20Altri%20elementi
[50] https://www.zdnet.com/article/chatgpt-sees-its-first-monthly-drop-in-traffic-since-launch/
[51] Numbers represent search interest relative to the highest point on the chart for the given region and time. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular. A score of 0 means there was not enough data for this term. The term has been analyzed using Google Trends, worldwide, across the past 12 months

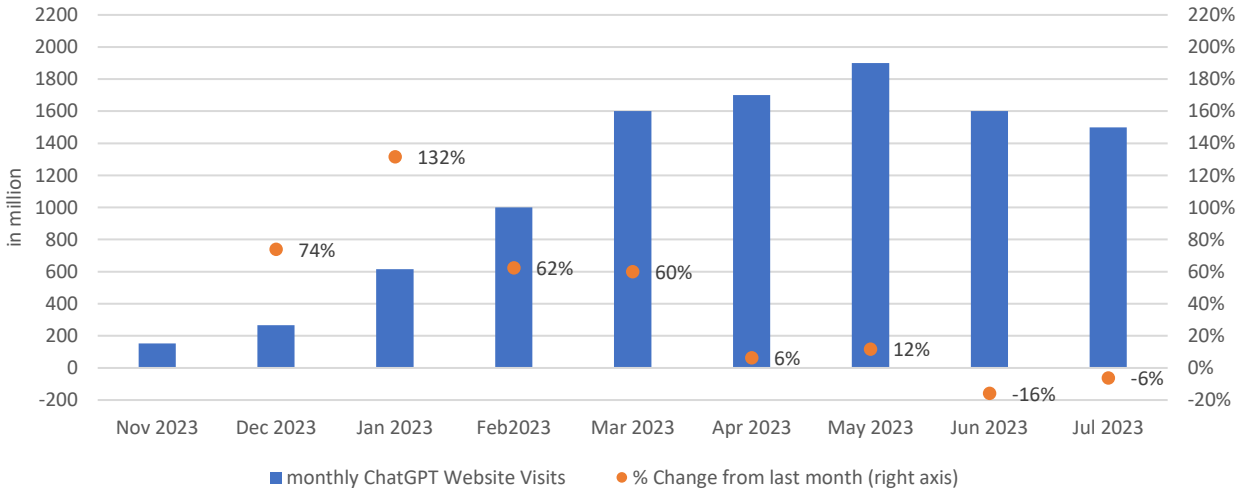
Figure 2: Interest in ChatGPT on Google search from November 2022 to September 2023 worldwide, by week



Source: Statista (2023)

The official ChatGPT website surpassed 1 billion page visits back in February 2023. The highest engagement came in the month of May when the website had 1.9 billion visits. In June 2023, there occurred the first drop ever since ChatGPT's launch[52].

Figure 3: ChatGPT website monthly views

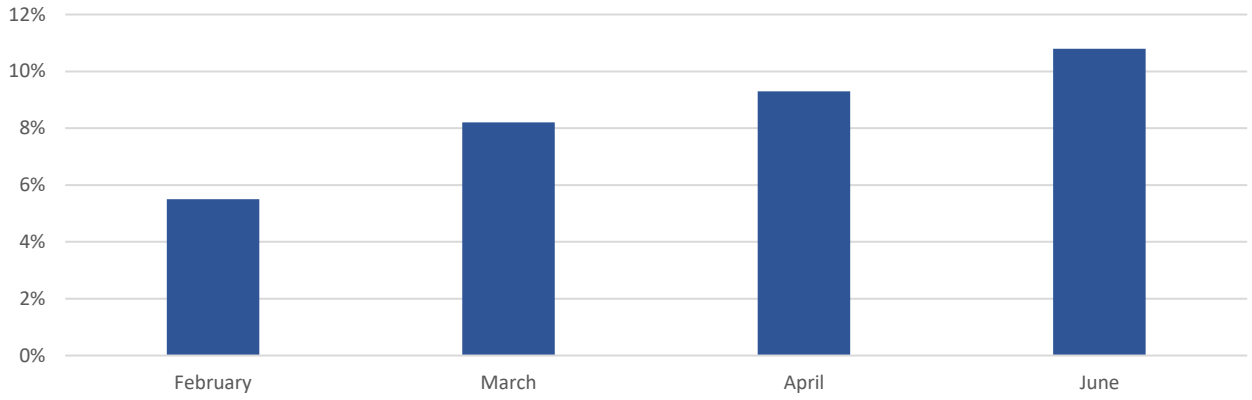


Source: DemandSage, ChatGPT Statistics (September 2023)

[52] https://www.demandsage.com/chatgpt-statistics/ (Last access: 16 October 2023)

In the business environment, the use of ChatGPT is growing rapidly. As of June 2023, it was reported that 10.8 % of employees of worldwide companies had tried using ChatGPT in the workplace at least once, an increase of more than 5 percentage points compared to February of the same year.

Figure 4: % of company employees worldwide using ChatGPT in work environments from February to June 2023



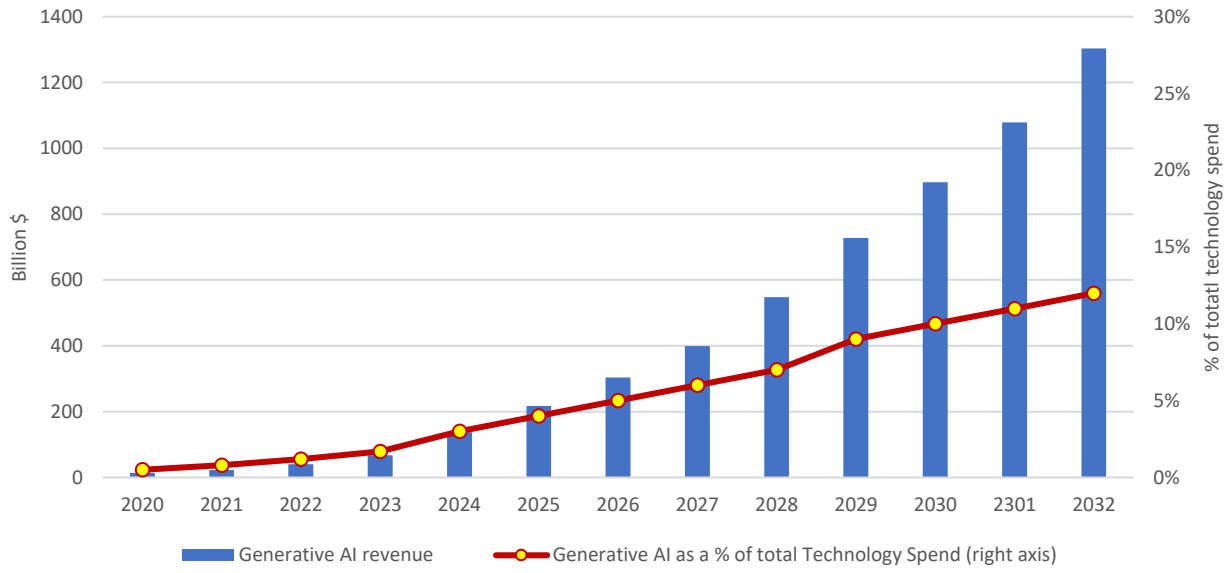
Source: Statista (2023)

2.2 The economic potential of generative AI

With the influx of generative AI programs like Google's Bard and OpenAI's ChatGPT, the generative AI market is poised to explode, growing to $1.3 trillion over the next 10 years from a market size of just $40 billion in 2022. Value is expected to show an annual growth rate (CAGR 2022-2032) of 42%, driven by training infrastructure in the near-term and gradually shifting to inference devices for large language models (LLMs), digital ads, specialized software and services in the medium to long term. Generative AI is ready to expand its impact from less than 1% of total technology spent to 12% by 2032[53] .

[53]https://www.bloomberg.com/company/press/generative-ai-to-become-a-1-3-trillion-market-by-2032-research-finds/ (Last access: 6 September 2023)

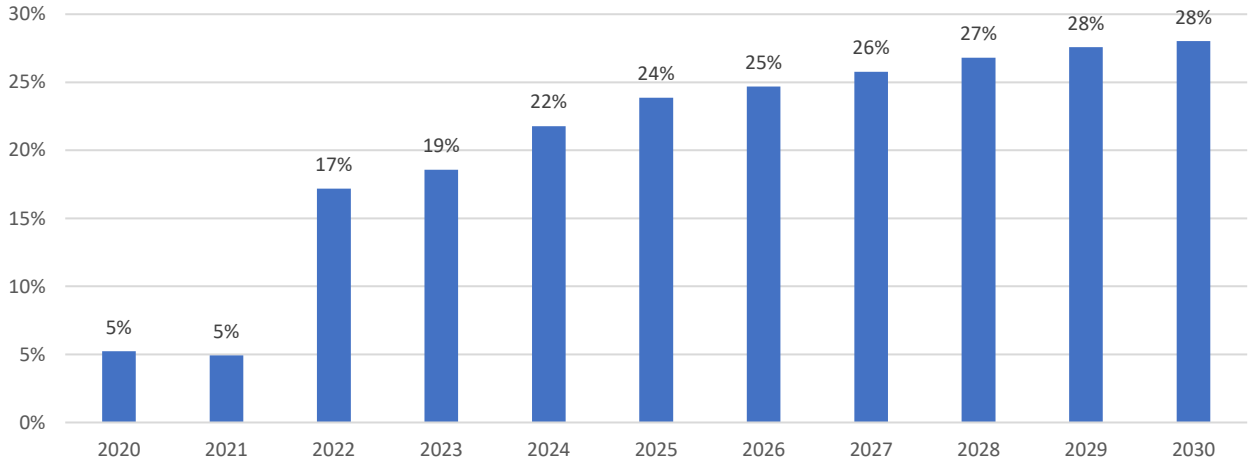
Figure 5: Generative AI worldwide market



Source: Bloomberg Intelligence, IDC (2023)

Moreover, generative AI represents an important share of the total AI market today, which is set to increase in the coming years. In 2023, it represents 19% of the total AI market and is estimated to reach 28% by 2030.

Figure 6: The global generative AI market (as % of total AI market)



Source: I-Com elaboration on Statista data

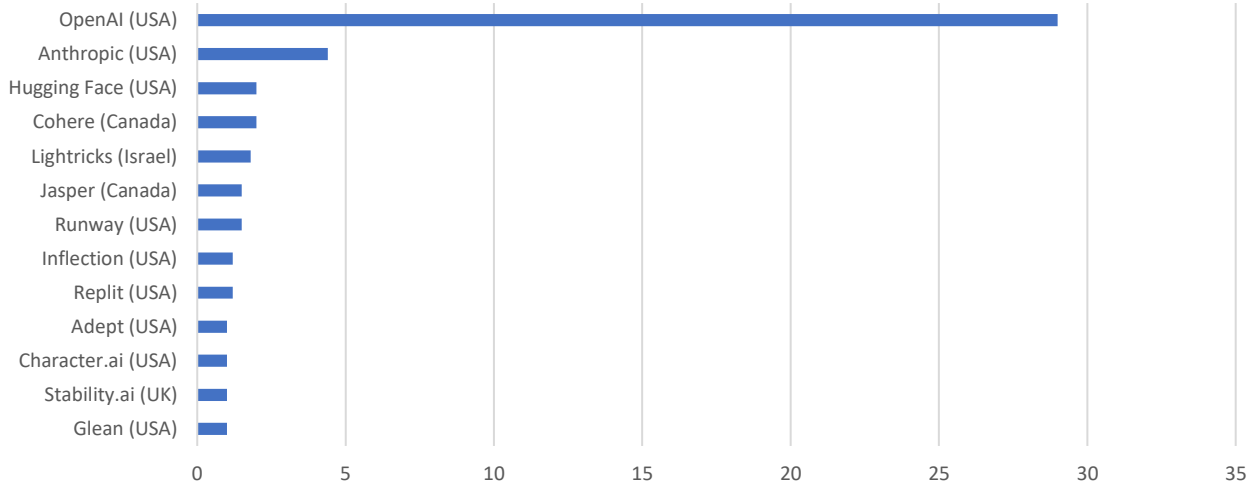
The interest in generative AI is truly astonishing. For instance, venture capitalists have piled billions of dollars into young startups working on the breakthrough technology since the launch of ChatGPT-3 in November 2022.

According to the latest update by CB Insights[54], the generative AI market has already produced 13 **unicorns** (by April 2023).

These startups are taking far less time to join the $1 billion valuation club than most of their unicorn peers. Across the 13 genAI unicorns, the average time to reach unicorn status is 3.6 years whereas for the unicorn club as a whole the average is 7 years — almost twice as long.

In May 2023, OpenAI was reported to have a valuation of $29 billion[55], followed by Anthropic ($4.4 bn), Cohere ($2 bn) and Hugging Face ($2 bn).

Figure 7: Generative AI startups with $1B+ valuations (as of 08/05/2023)

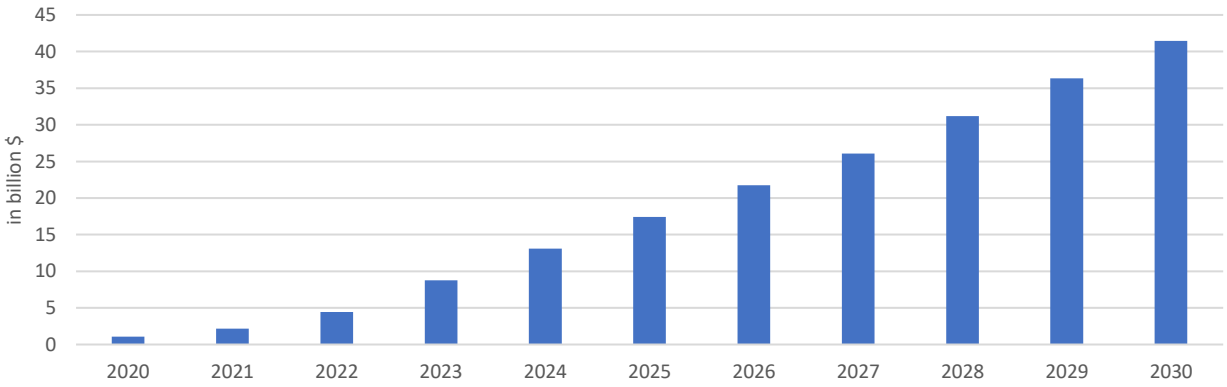


Source: CB Insights (2023)

From this ranking, the **conspicuous absence of EU countries is quite clear**. However, **the EU generative AI market is also showing rapid growth**. According to some estimates, value[56] in the EU generative AI market is projected to reach $ 8.77 billion in 2023 and it is expected to show an annual growth rate (CAGR 2023-2030) of 24.85%, resulting in a market volume of $41.47billion by 2030.

[54] https://www.cbinsights.com/research/generative-ai-unicorns-valuations-revenues-headcount/
[55] According to the Financial Times ("ChatGPT parent OpenAI seeks $86 billion valuation", October 20, 2023), OpenAI would be in talks with investors about selling shares at a valuation of $86 billion, roughly three times what it was worth six months before.
[56] Values are generated by the funding amount in Generative Artificial Intelligence initiatives and projects by many companies such as Open AI, NVIDIA DeepL Learning and Google (Magenta, DeepDream)

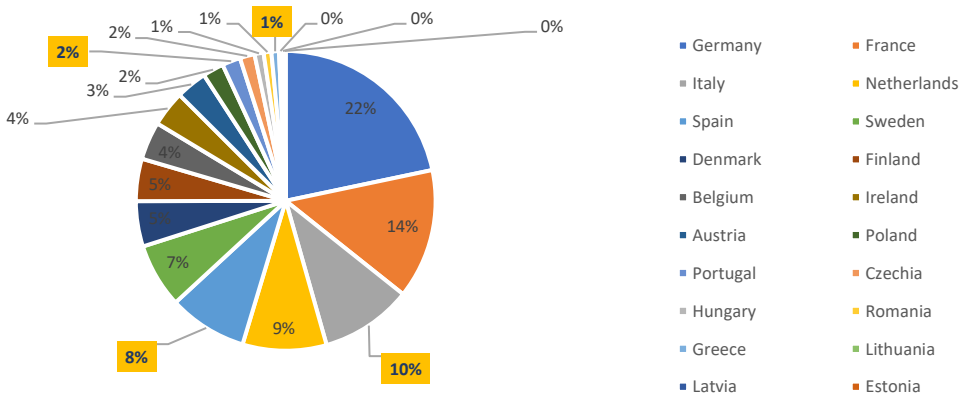
Figure 8: Value of Generative AI market in the EU



Source: Statista (2023)

Among the main Member States, Germany is the largest market for generative AI, representing 22% of the total EU market, followed by France (14%) and Italy (10%). Spain ranks fifth (8%) while Portugal and Greece are at bottom of the ranking with 2% and 1%, respectively, of total generative AI EU market.

Figure 9: Generative AI market in the Member States (%)

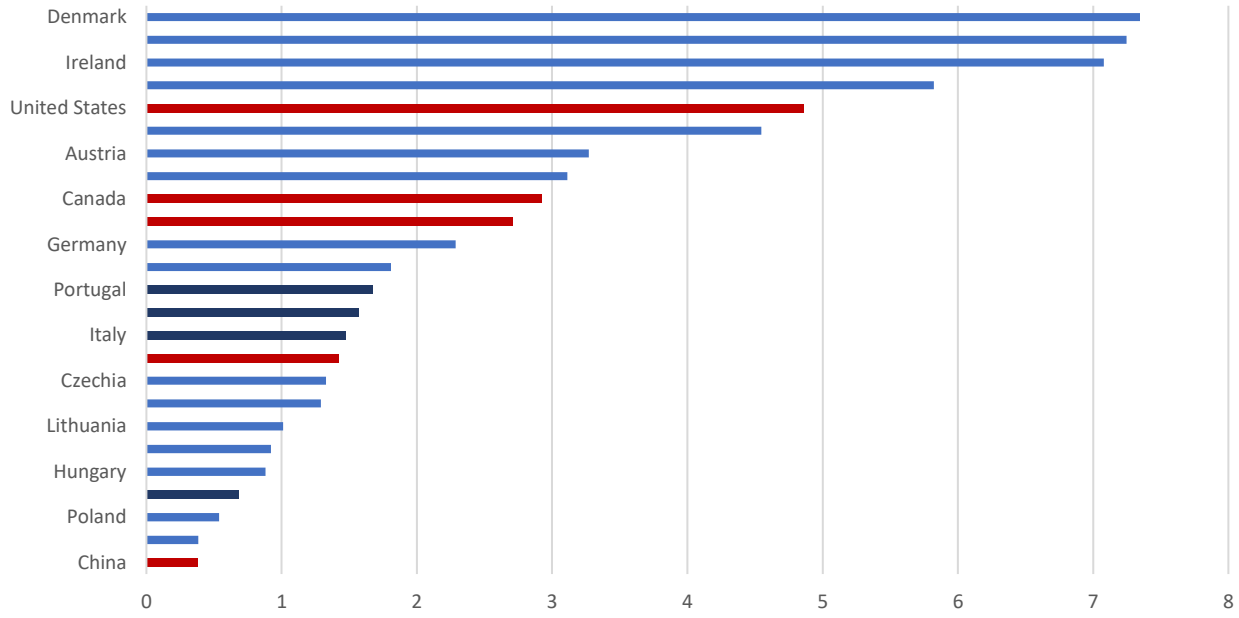


Source: I-Com elaboration on Statista data

Accounting for country population size, **Denmark** appears to be the largest generative AI market at worldwide level, with a market value per 100,000 inhabitants of $7.35 million, followed by **Finland** and **Ireland**. Instead, **Germany** - the biggest AI generative market in Europe in absolute terms - ranks eleventh with a market value of $2.29 million per 100,000 inhabitants.

All countries from Southern Europe (**Greece**, **Italy**, **Portugal**, **Spain**) are below with lower values, ranging from $1.68 million per 100,000 inhabitants for Portugal to $0.68 million per 100,000 inhabitants for Greece.

Figure 10: Generative AI market value/100,000 inhabitants ($ m)



Source: I-Com elaboration on EUROSTAT, OECD and Statista data

Europe has more than **150 startups** that have already raised capital so far and are working on generative AI. The **UK** is the first country in terms of generative AI startups, with more than 50 companies working in the field (36% of total European Generative AI startups), compared to **Germany-** in second place, with 19 startups (13%). **Spain** and **Italy** rank well below with 5% and 3%, respectively of generative AI startups.

In terms of the areas that generative AI startups are working in, text generation has seen the biggest explosion of companies working with the tech, with ML platforms, audio and image generation close behind.

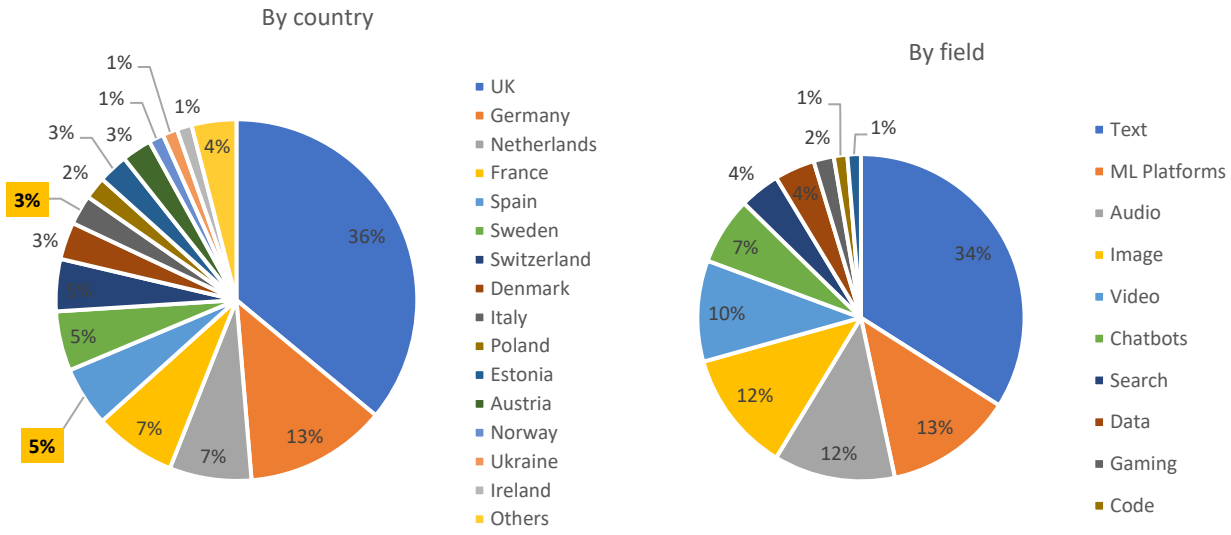
The growth of generative AI could lead to a significant impact on the world economy, thanks to its beneficial effects on productivity.

According to Goldman Sachs, it could drive a 7% (or almost $7 trillion) **increase in global GDP and lift productivity growth** by 1.5 percentage points over a 10-year period[57].

[57] https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html (Last access: 6 September 2023)

As well, according to the estimates of McKinsey & Company analysts, productivity could largely increase thanks to generative AI uses. Across 63 use cases analyzed, generative AI has the potential to generate $2.6 trillion to $4.4 trillion in value across industries.

Figure 11: Generative AI European startups



Source: https://sifted.eu/articles/europe-generative-ai-startups (September 2023)

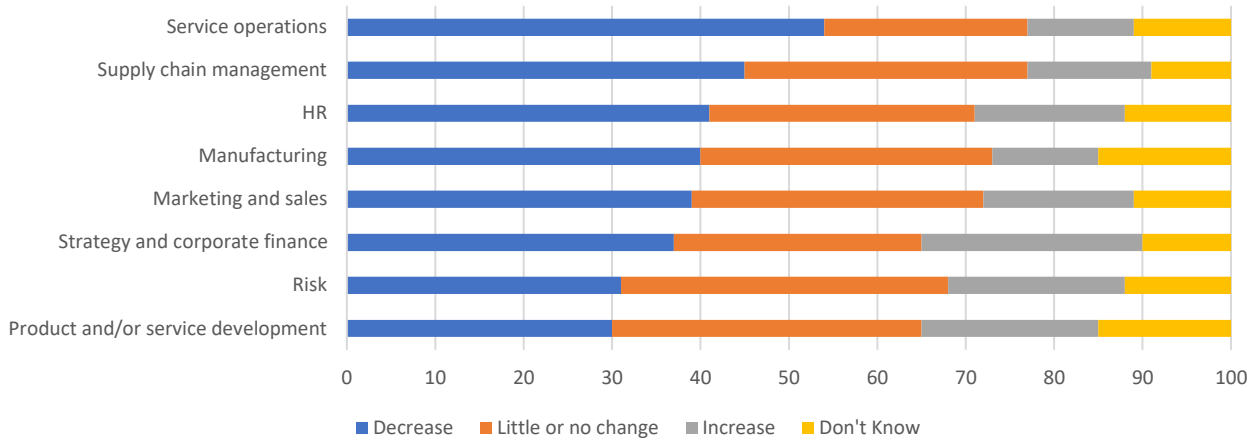
The banking, high tech, life sciences, and retail sectors will reap the greatest benefits on their revenues. In the banking sector, for example, there could be an annual added value of $200-340 billion while in the high-tech sector it could reach $240-460 billion[58].

According to McKinsey analysts, generative AI could enable **labour productivity** growth of between 0.1 and 0.6 percent annually through 2040, depending on the rate of technology adoption and redeployment of worker time to other activities[59]. However, **automation processes induced by generative AI will impact on knowledge work**, particularly activities involving decision making and collaboration, which previously had the lowest potential for automation. Therefore, a reorganization and retraining in the workplace is essential. According to 38% of organizations interviewed by McKinsey & Company in a recent survey[60], more than 20% of their workforce will be reskilled as a result of AI adoption. Moreover, 8% of respondents say the size of their workforce will decrease by more than 20%, while for 30% little or nothing will change. Looking specifically at

[58] McKinsey & Company, The economic potential of generative AI, June 2023
[59] Ibidem
[60] McKinsey, The state of AI in 2023: Generative AI's breakout year, August 2023

generative AI's predicted impact, service operations is the only area where most respondents (54%) expect to see a decrease in workforce size in their organizations (Fig. 12).

Figure 12: Effect of generative AI adoption on number of employees, by business function, next 3 years (% of respondents)



Source: McKinsey & Company (2023)

2.3 Generative AI: risks and issues to be addressed

Generative AI, which, as already underlined, involves using artificial intelligence to create content autonomously is already becoming an ever-more prominent part of everyday business activities and our daily lives. Despite the positive effects, there are also some quite serious risks.

According to initial expert assessments, **the five most immediate risks associated with generative AI**[61] are:

"Hallucinations" and fabrications: including factual errors, these are some of the most pervasive problems already emerging with generative AI chatbot solutions. Training data can lead to biased, off-base or wrong responses, but these can be difficult to spot, particularly as solutions are increasingly believable and relied upon;

Deepfakes: when generative AI is used for content creation with malicious intent. These fake images, videos and voice recordings have been used to attack celebrities and politicians, to create and spread misleading information, and even to create fake accounts or take over and break into existing legitimate accounts;

Data privacy: Employees can easily expose sensitive and proprietary enterprise data when interacting with generative AI chatbot solutions. These applications may indefinitely store information captured through user inputs, and even use information to train other models —

[61] https://www.zdnet.com/article/the-5-biggest-risks-of-generative-ai-according-to-an-expert/

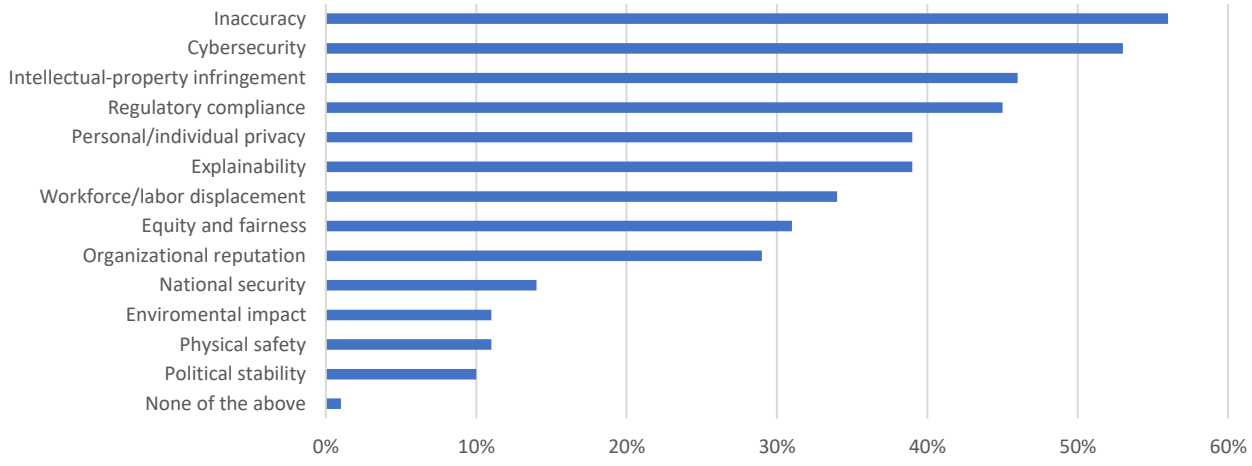
further compromising confidentiality. Such information could also fall into the wrong hands in the event of a security breach;

Copyright issues: Generative AI chatbots are trained on a large amount of internet data that may include copyrighted material. As a result, some outputs may violate copyright or intellectual property (IP) protections. Without source references or transparency into how outputs are generated, the only way to mitigate this risk is for users to scrutinize outputs to ensure they don't infringe on copyright or IP rights;

Cybersecurity concerns: In addition to more advanced social engineering and phishing threats, attackers could use these tools for easier malicious code generation. Vendors who offer generative AI foundation models assure customers they train their models to reject malicious cybersecurity requests, however, they do not provide users with the tools to effectively audit all the security controls in place.

Inaccuracy, **cybersecurity**, **and intellectual property infringement** are also the most cited risks related to the adoption of generative AI by global business organizations, surveyed by McKinsey in a recent survey[62] (Fig. 13).

Figure 13: Generative AI–related risks that organizations consider relevant (% of respondents)



Source: McKinsey & Company (2023)

To mitigate the possible risks of generative AI, organizations should create a company-wide strategy that targets AI trust, risk and security management. Moreover, organizations must establish a governance and compliance framework for enterprise use of these solutions, including clear policies that prohibit employees from asking questions that expose sensitive organizational or personal

[62] Ibidem

data[63]. At the same time, it is urgent that AI developers work with policymakers, including new regulatory authorities that may emerge, to establish policies and practices for generative AI oversight and risk management.

In conclusion, transparency, accountability and safety should be prioritized to ensure a responsible use of generative AI in society.

[63] https://www.gartner.com/en/newsroom/press-releases/2023-04-20-why-trust-and-security-are-essential-for-the-future-of-generative-ai

Chapter 3: The geopolitics of generative AI: international implications and the role of the European Union

3.1 Great power competition: who is winning the generative AI race?

Artificial Intelligence in general, and generative AI in particular, have generated a major public discussion over the future of global competition and international leadership in the key geopolitical vector that technology represents. The competition has become a two-fold race - the race over the development of AI, and the race over its adoption.

While competition is an important element, it does not exclude countries from cooperation. The US and China had the greatest number of cross-country collaborations in AI publications from 2010 to 2021, although the pace of collaboration has slowed. The number of AI research collaborations between the US and China increased roughly 4 times since 2010, and was 2.5 times greater than the collaboration total of the next nearest country pair - the UK and China. However, the total number of U.S.-China collaborations only increased by 2.1% from 2020 to 2021, the smallest year-over-year growth rate since 2010. Both competition and cooperation have increased in recent years, and refer to three levels- governments, private sector and academia.

There is no segregated information on the governmental budget devoted to generative AI with regards to AI in general. However, available data shows that the U.S. federal government is allocating the highest share of funding towards decision science, computer vision and autonomy segments. In the three cases, generative AI plays an increasingly important role. The federal government has increased its budget on AI by more than $ 600 million year on year, up from $ 2.7 billion in 2021. Total spending on AI contracts has increased by nearly 2.5 times since 2017, when the U.S. government spent $ 1.3 billion on artificial technology.

The recently released President Biden's Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence, which has created a landmark in the U.S. Administration to push forward the inclusion of AI in the public sector, makes a reference to generative AI. Concretely, it displays the need to protect citizens from "AI-enabled fraud and deception by establishing standards and best practices for detecting AI-generated content and authenticating official content. The Department of Commerce will develop guidance for content authentication and watermarking to clearly label AI-generated content." However, this reference is only made once.

However, **this statement aligns with the, also released on the same day, G7 Leaders' Statement on the first-ever International Guiding Principles and a voluntary Code of Conduct.** Arising from the Hiroshima AI Process, Germany, Canada, the U.S., France, Italy, Japan and the UK highlights the importance of engaging developers and agreeing on a common baseline of principles to develop AI, including generative AI. At the same time, Western countries have been focusing on generative AI,

although still at a preliminary stage, **with the UK's AI Safety Summit, whose theme is devoted to frontier AI risks and opportunities, including generative AI.**

Alongside these Western-centric approaches to generative AI, China has also been developing certain initiatives. However, no specific data has been publicly presented either, but the proposal for a first-of-its-kind rules governing generative AI, announced by the government in July 2023, shows the level, scope and interest in the issue. **The rules, developed by the Cyberspace Administration of China (CAC),** will only apply to generative AI services that are available to the general public rather than those being developed in research institutions. Generative AI providers will need to obtain a license to operate, conduct security assessments on their product and ensure user information is secure, based on the "core values of socialism", as acknowledged by the CAC.

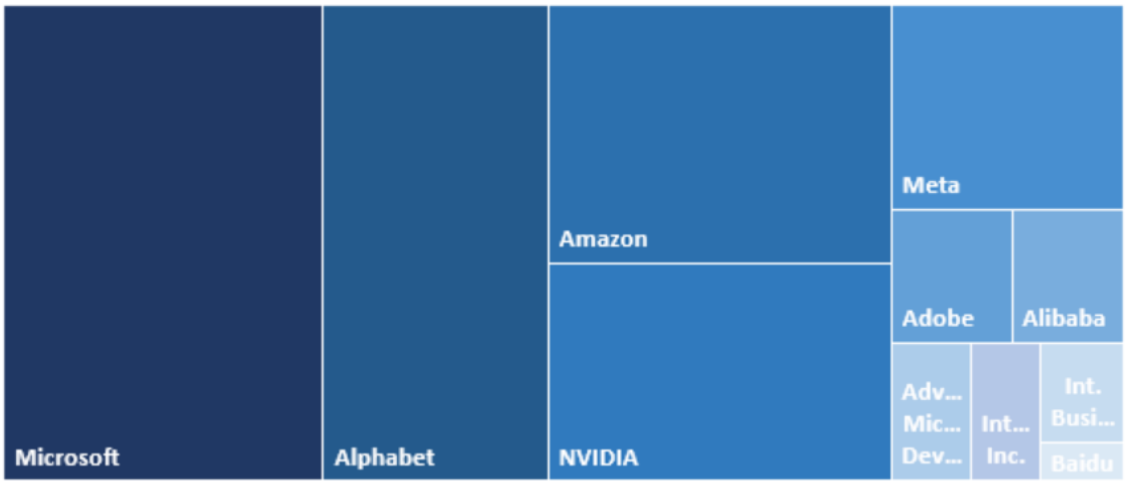
In this scenario of nationally led initiatives or mini-lateral coalitions, the United Nations has decided to take on a voice as shaper of policy discussions at the global level. In particular, the UN Secretariat-General, which already launched the proposal for a Global Digital Compact to be approved during **the Summit of the Future in September 2024, has set up a High-Level Advisory Body on Artificial Intelligence, whose two co-chairs are from Spain - an active EU Member State in AI,** and also the expected final Presidency to the EU Council - and from Zimbabwe to approve the AI Act proposal.

While there is less information on the governmental decisions to address generative AI, the existence of a growing market by the private sector shows the intensity and exponential growth – and, thus, competition - of this technology vertical by companies. In this realm, there is no single Chinese and U.S. leader in generative AI. Depending on the topic, one or the other leads.

a. Top firms using generative AI: the US leads the way

As for top firms using generative AI, U.S. companies are far ahead of Chinese firms and other countries. Microsoft leads the use of generative AI with a market capitalisation of $ 2.442 trillion, followed by Alphabet ($ 1.718 trillion). The leading Chinese company is Alibaba Group ($ 241.97 bn), which is in the top ten worldwide and ranks 7th.

Figure 14: Top 12 of market share of firms using Generative AI



Source: Self-made with data from Insider Monkey, 2023

American firms also have a strong competitive advantage in the design and production of graphic processing units (GPUs). GPUs are a type of hardware that is the main power source for most Large Language Models (LLMs), a deep learning algorithm that can perform a variety of natural language processing (NLP) tasks. This provides major incentives and opportunities for innovation. There are other technologies that have strong interactions with generative AI, **such as the US having a strong leverage in the cloud computing industry,** essential for training and deploying generative AI models. The ranking does not disclose the transformative power of generative AI in abolishing market barriers and strong clusters that end up producing gatekeepers. Even if large companies may have access to the best technology and technologists, generative AI is a field that is growing very quickly and may bring some new players into the field. **Smaller competitors may be building more adaptable and, most importantly, cheaper, and more accessible open-source AI models.** If this document turns out to be true, AI is also fostering a bottom-up competition from smaller AI enablers that can generate even more innovation in the US and hinder Chinese actors from becoming leading top firms in generative AI. This is due to two reasons.

The first is because companies based in China need to abide by a greater number of rules, imposed by the government, and with a greater amount of oversight mechanisms on which type of content flows through generative AI processes on political, societal or cultural issues. The second reason is because Chinese companies might face a similar process as has happened with the **"tech crackdown"** on Chinese companies since 2020. Chinese authorities initiated a regulatory storm against the country's Big Tech firms in late 2020 out of concerns that the country's major internet platforms were experiencing a major growth and monopolistic behaviour. It started with the limitation of Alibaba's AntGroup IPO in the US and was followed by a greater number of norms limiting the international expansion of companies. After three years, the Chinese government

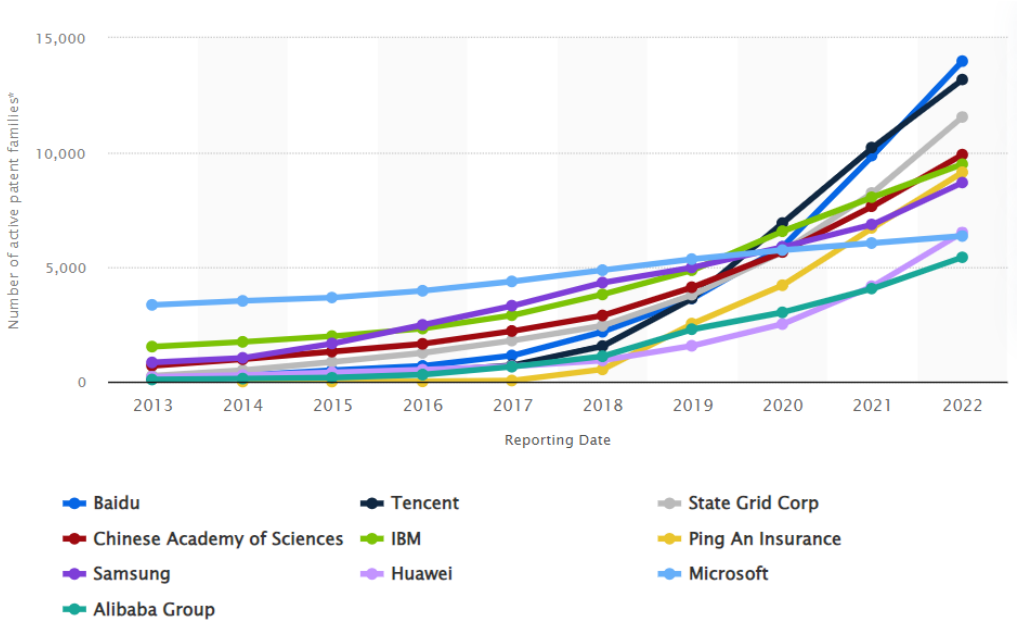
announced in 2023 its strategy to maintain the "bottom line of development security" and strengthen its "linkage effect" with international markets. This might mean that the technology crackdown is declining. However, generative AI companies may still face a greater level of dynamism in the US than in China.

AI has been one of the forefronts of US-China competition, but the effects do remain limited to the competition between these two. **The EU, for instance, is falling behind, as no European firm was in the ranking.** Kai-fu Lee, the ex-president of Google China and prominent venture capitalist and AI expert, **told Sifted that Europe is not even in the running for a "bronze medal" in the AI race.** He stated that Europe has none of the success factors like the US or China, due to a lack of a VC-entrepreneur ecosystem, successful consumer internet companies, social media companies, or big-sized mobile application companies that can drive AI advances, and a lack of governmental support to win the generative AI race.

b. Top countries with the most AI patent applications: issues with generative AI, intellectual property and China's leadership

China has the most AI patent applications, with **4,636 applications or 64.8% of all patents requested globally.** The US comes in second with 1,416 AI patent applications, making up 19.8% of patent applications with offices listed, followed by the Republic of Korea, with 532 applications. In 2022, China filed 29,853 AI-related patents, up from 29,000 the previous year. Beijing accounted for more than 40% of global AI applications in 2022, mainly due to the monetisation of AI products from top tech firms like Baidu and Alibaba. Since 2017, China's patents have surpassed the US ones, and now represent almost double the sum of the US. Concretely, **China is leading the competition in patents with Baidu and Alibaba monetising AI products.** The countries following in the list are Japan and South Korea with a total of 16,700 requests.

Figure 15: Largest patent owners in machine learning and artificial intelligence (AI) worldwide from 2013 to 2022, by number of active patent families



Source: Statista, Worldwide; LexisNexis PatentSight; 2013 to December 31, 2022

The Chinese government considers its domestic patent market as a key economic sector and has been leading in it since 2021, especially now with generative AI becoming a growing market for more patents. Last year, the China National Intellectual Property Administration (CNIPA) released a draft of measures to downgrade the ratings of Chinese patent agencies that were following non-desirable and fraudulent patent schemes.

The Chinese Communist Party (CCP) supports its domestic companies in generative AI, as well as in other strategic industries through subsidies. Even if there are no state-owned enterprises in the AI industry, the CCP still influences market direction and collaborates with private companies through financial and regulatory leverage. For instance, iFlytek has received substantial government subsidies, even exceeding half of the company's annual net profits, namely 258.18 million yuan ($ 37.7 million).

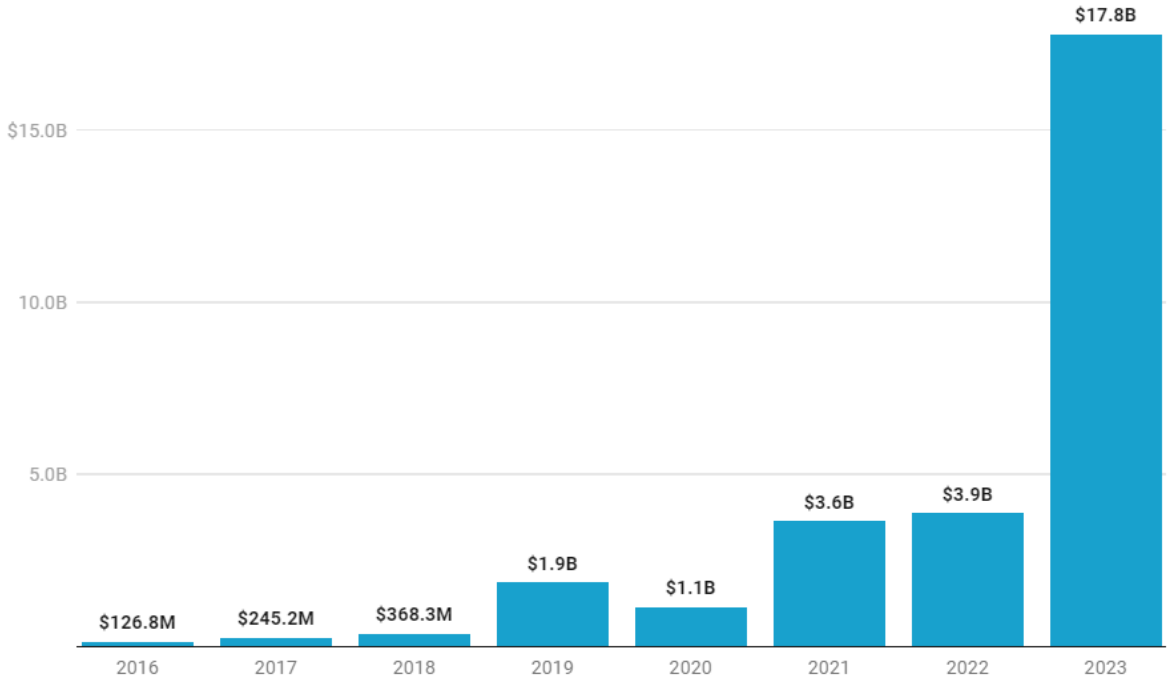
Europe is dramatically underperforming in terms of patents. **According to the WIPO report, out of the top 167 universities and public research institutions for patents, only 4 are in Europe.** Out of these 4 European public research organisations on the WIPO list of top AI patent filers, the highest-placed is the German Fraunhofer Institute, which is ranked 159th, while the French Alternative Energies and Atomic Energy Commission (CEA) is in 185th position.

c. VC Investments in AI: comparison between countries

Since 2019, generative AI startups have attracted over $ 28.3 billion in funding from investors. Only in 2023, did it attract $ 17.8 billion from January to August. Generative AI startups raised over $ 14.1 billion in equity funding across 86 deals in Q2 2023, making it a record year for investment in

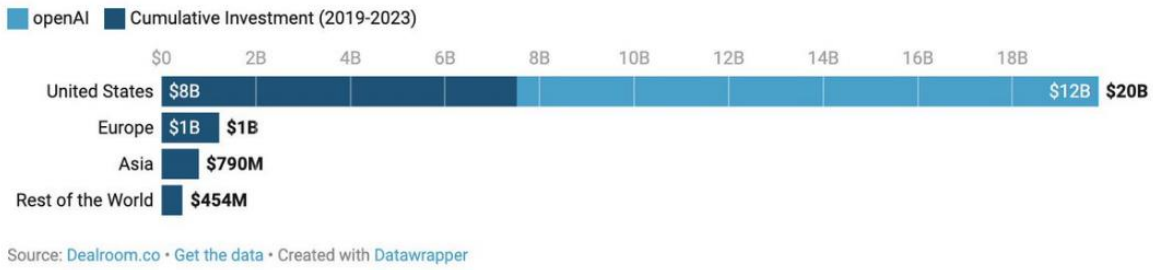
generative AI startups. Concretely, 89% of the global total ($ 20 bn) went to the US, making it the leader in VC Investments in AI.

Figure 16: Attracted funding to generative AI startups



Source: CB Insights, 2023

Figure 17: Generative AI Venture Capital funding by regions



Source: Thenextweb, Dealroom, 2023

The success the US in generative AI startups and venture capital must be be contextualised with the role of the new players in the AI-generative market and the capacity to attract high-risk venture capital, as occurred with the success of the U.S. firm OpenAI. However, this is not the only Company. **According to Forbes' AI 50 list from April 2023, the top 5 generative AI start-up companies in terms of funding come from the US. China,** on the other hand, **has also been investing in several**

generative AI applications. 22 generative AI start-ups in China received funding as opposed to 21 in the US and 4 in the UK. However, US start-ups received more funding in total. 12 of the 18 generative AI companies that received funding totaling more than 100 million yuan (US$ 138,287) in the first half were from the US, while only 3 were from China.

China's government aim is that AI will represent 1 trillion yuan ($ 146 bn) for the country by 2030. While there are 14 unicorns worth a combined $ 40.5 billion in AI companies in China, generative AI will still need a greater level of investment to accomplish its goal of global leadership in this specific technology development.

European startups received just $1 billion out of the € 22 billion that VCs have invested in generative AI since 2019, making up only 5% of the $22 billion to Europe. Asian startups received $ 790 million. However, the assessment of which country may become the leader in generative AI does not only stem from the number of top firms, publications or start-up funding. It also derives from the flow of VC investments to certain high-priority industries, and who is the final recipient of this target investment. Concretely, the main target sector of U.S. investments in generative AI (more than 50%) flow into the media, social platforms and marketing sectors.

The second largest sector is the IT infrastructure and hosting, followed by financial and insurance services, digital security, travel, leisure and hospitality, and government, security and defence to a lesser extent. It is worth noting that US stakeholders invest almost 80% of generative AI funding in these two sectors within their own country, and not abroad.

On the other hand, Chinese stakeholders also invest in generative AI for high-priority sectors within their home country. Mostly, they invest in IT infrastructure and hosting, healthcare and biotechnology, business processes and support services, financial and insurance services, digital security, and media, social platforms and marketing. This scenario reflects that China and the US are similarly targeting media, social platforms and marketing, as well as IT infrastructure and hosting as the main priority sectors for generative AI investments, investing in these to strengthen their home country industries. They diversify other low-profile investments into other sectors, but they are not very significant.

3.2 The implications of generative AI for security, economy and rights

3.2.1 Global security and defence

Generative AI poses both opportunities and risks for security and defence. On the one hand, it has been a valuable source for defence policy and planning. In August 2023, the US Department of Defense (DoD) launched *Task Force Lima* in order to integrate AI into national security by both minimising risks, and leading innovation in future outcomes in defense.

NATO has also initiated preliminary discussions over the potential impact of Generative AI. **NATO's Data and Artificial Intelligence Review Board (DARB),** which serves as a forum for allies and the focal point for NATO's efforts to govern responsible development and use of AI by helping operationalise Principles of Responsible Use (PRUs), hosted a panel briefing on generative AI and its potential impact on NATO to weigh the capabilities and limitations of generative AI. One of the main conclusions was that the generative AI stack is still not highly comprehensive in reasoning and planning to meet the critical functions required by the military.

While this panel briefing has an informational goal, it remains of high interest to assess which will be the final output from DARB on this issue. DARB is in charge of translating these principles into specific, hands-on Responsible AI Standards and Tools Certifications, and provides a common baseline to create quality controls and risk mitigation mechanisms.

Another challenge for security and defence is the impact of generative AI on intelligence analysis. AI enhances situation awareness and geolocation through various open data sets (OSINT), as well as decision-making processes and scenario-planning such as wargaming or foresight analysis. However, one challenge from generative AI is that it has so far tended to be fed by open-source data and platforms, which makes it harder for intelligence communities to make use of it as a trusted tool.

If we focus specifically on the US-China competition, 3 out of 5 of the world's top five commercial drone brands are Chinese, while just one is American. More concretely, DJI's market share is expected to grow from $ 30.6 billion in 2022 to $ 55.8 billion by 2030, while it has 70% of the current market value in the whole drone industry. The growing use of generative AI in drones might create an additional race for the development of these technologies on the battleground. For instance, one of the leading AI firms in the world is DJI from China, which has 70% of the global share of the drone market with a value of $ 33 billion.

However, this potential use of generative AI also gives rise to a number of risks, threats and challenges that might influence the eventual decision of using generative AI to make security- and defence-oriented decisions. **Data overflows may pose a challenge for security and defence policy-makers when working with private companies that provide solutions to their services and projects.** If a company makes use of open-source generative AI, the code sent to the GitHub service

that manages this information flow could contain a company's confidential intellectual property, and sensitive data such as API keys that have special access to customer information.

3.2.2 Markets and economy

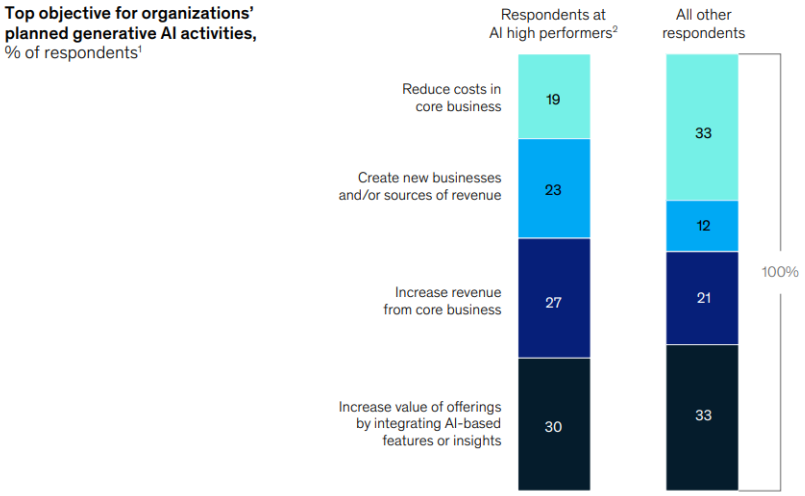
President Xi Jinping considered AI a technology where China had to lead, setting specific targets for 2020 and 2025 that put the country on a path to dominance over AI technology and related applications by 2030. Generative AI is nothing different. The creation of generative AI start-up ecosystems across countries is gaining importance, as has been shown. However, generative AI poses new questions for the global economy.

First, in global trade, generative AI may be understood as both an enabling tool and an end-goal. As an enabling tool, **it may detect large language models of data on transactions, documents and contracts, to identify anomalies, fraud risks, and compliance issues** with regards to money laundering, economic sanctions and tax havens. AI may create a new line of opportunity for global trade, although data shown in previous sections displays that so far countries have been investing in this technological asset on a domestic basis, with limited cross-border flows.

Second, generative AI may lead to two trends. On the one hand, companies aim to leverage generative AI internally in their companies to reduce costs of partnerships with third actors. According to McKinsey, 33% of organisations aim to use generative AI to reduce costs in core business (what might reduce the diversification of partners and suppliers from other industries and countries), and 12% aim to create new businesses and/or sources of revenue.

Figure 18: Top objective for organizations' planned generative AI activities, 2023

Smaller shares of AI high performers see cost reductions as their top objective for generative AI efforts.



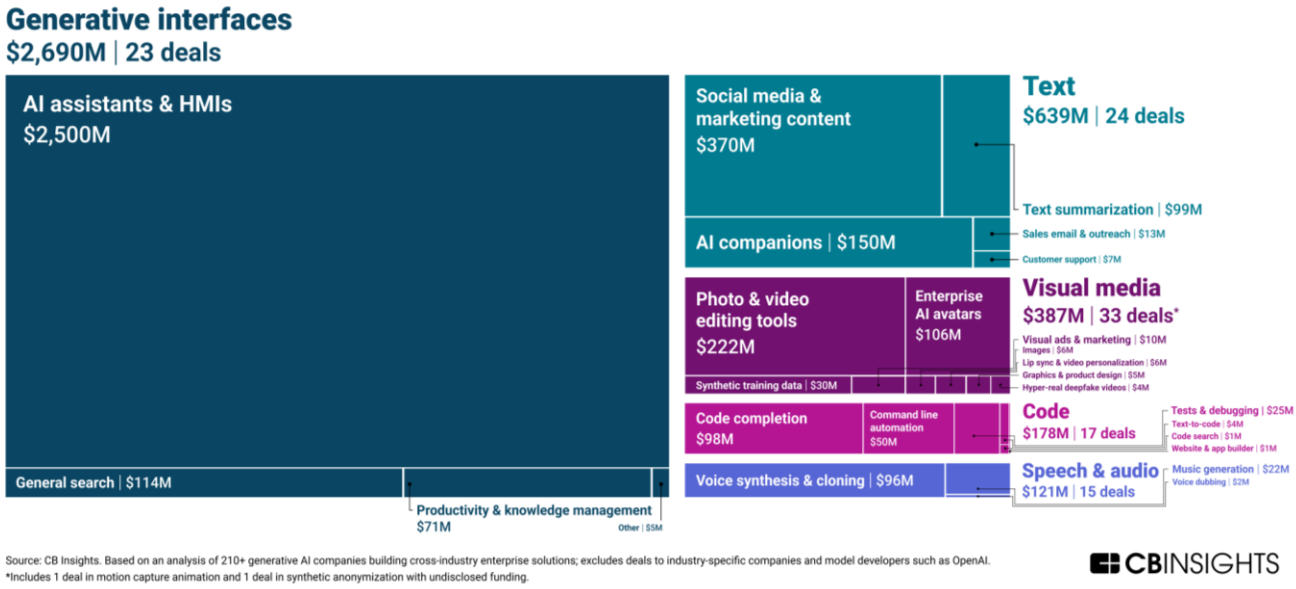
Source: McKinsey, 2023

On the other hand, generative AI infrastructure will require cross-industry collaborations. Even if a company aims to reduce its costs of partnerships with third countries, become much more competitive and create higher revenues on its own, the creation of the infrastructure still requires setting up deals with other industries to make that goal feasible. According to CB Insights, the creation of this infrastructure is leading to cross-industry collaborations ($ 4 bn, 116 deals) more than industry-specific collaborations (only $ 0.3 bn, 45 deals). This two-fold assessment means that generative AI will create new layers of competition across companies but will also undoubtedly lead to greater levels of needed cooperation across sectors to gain the technological edge over others.

Third, generative AI may create new grounds of competition and cooperation across sectors and products. For example, AI-enabled fintech companies are an increasingly important sector. With generative AI, they might see a strong upsurge in their global positioning as an important industry. The more support a country provides to the establishment of existing, yet-to-be-top ecosystems, the more successful and competitive they will be in the global arena. **As of 2023, Tencent's WeChat Pay, the main fintech services provider in China, counts over 1.2 billion users** – of course, only operating in China. In contrast, Apple Pay has over 500 million users worldwide, with no less than 30 million being from the US. The UK is an interesting case as well. Founded in London in 2015, Revolut now has over 30 million retail customers worldwide and 6.8 UK users, having added over five million new users globally since November 2022. **Lastly, Klarna from Sweden, if we move to the EU, has over 4 million monthly users**, however, it reveals that the EU is way behind in AI-enabled fintech compared to other blocs.

Biometrics and facial recognition are another area where generative AI might be used and competition may be present. Biometrics can be understood as the identification of people using their physical or behavioural traits, and it has a strong dual-use aspect which can be used to control population or even use deep fakes and deception techniques through generative AI. Synthetic biometric data like fingerprints or our faces can be produced using generative AI, which can lead to breaches of personal information, data retention and overall trust in AI-enabled security systems.
Fourth, global economic competition, and cooperation, will not only be based on leading industries, but also lie in the specific applications companies may use. The greater a country encourages private companies to invest in AI assistants and HMIs, the more competitive it will become. However, if all companies focus on the same generative interface, they will lose competitiveness. This is why some countries are promoting another interface-specific leadership, for example India in code completion and voice synthesis.

Figure 19: Distribution of generative AI funding, Q3'22 – Q2'23



Source: The generative AI landscape: Top startups, venture capital firms, and more (cbinsights.com)

The U.S. Federal Trade Commission is still struggling to see how they can best achieve both the promotion of fair competition and the protection of citizens from unfair or deceitful practices. Market concentration and oligopolistic tendencies can also happen, as larger firms are those which also have access to more data, talent and capital to deploy AI-generated content. As the ranking shows, the most powerful firms in AI come from the US. These firms are also the technological global markets that seem to control all the necessary "raw materials" for deploying AI - big bulks of data

storage, strong computing power and cloud services, as well as the world's leading AI researchers and investments.

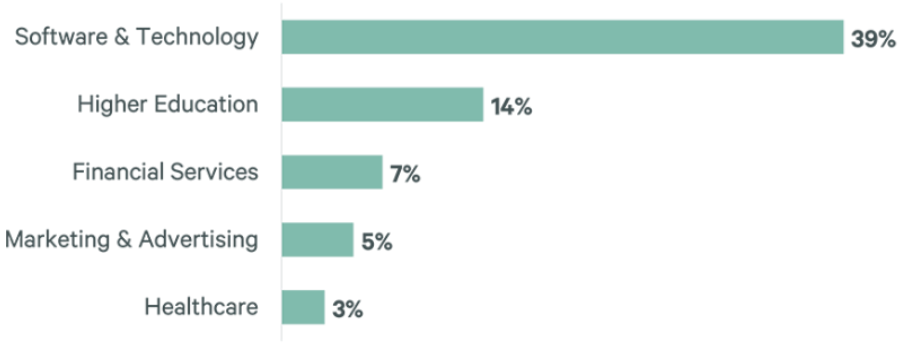
Likewise, an increasingly important, revamped topic is the mechanism of using state aid. China, the US, the EU, India and other technological powerhouses are increasing the amount of subsidies devoted to strategic technologies, either to reduce the dependence on third countries which pose a high-risk in case of supply chain disruption or shock, or to become a frontrunner in this specific technological vector.

China is developing Chinese Guidance Funds (CGFs), among other cases. These public-private investment funds are a meeting point between the Chinese business community and government, as the latter pours in money to mobilise massive amounts of capital in support of strategic and emerging technologies, including AI. The US, similar to the EU, has engaged in state aid promotion for critical technologies such as semiconductors.

However, jointly with subsidies, a critical point is how countries frame export control regimes. For instance, the restriction of exports on a certain technology may hinder the development in other sectors, such as generative AI. **According to Albright Stonebridge Group, U.S. restrictions limit access to some types of advanced semiconductors**, which are necessary for the greater compute requirements that future LLMs will need. If Chinese companies cannot access the Global Processing Units (GPUs), they will face challenges to develop LLMs as quick, comprehensive and complete as Western firms, thus limiting its share in the global marketplace.

The presence of a highly skilled workforce in AI in general, and in generative AI in particular, remains strategic**. In 2021, Quad partners Australia, India, Japan and the US announced the Quad Fellowship, a first-of-its-kind scholarship programme to build ties across STEM experts** from the four countries. AI is one of the top-priority areas for these talent exchanges, which include scholarships for PhDs, Masters, and talent exchanges across companies. This is of particular interest because the US has been retaining international U.S.-trained AI PhD graduates for a long time as they stay in the country, including those from AI competitors such as China. Some sectors in generative AI, on the other hand, employ the greatest share of U.S. talent, and this trend may increase in the next years.

Figure 20: Sectors that employ the greatest share of U.S. AI talent, 2023



Source: CBRE, 2023

By contrast, other studies have found that most China-trained AI talent currently lives outside of China. **Estimates indicate that its need for workers skilled in AI is expected to grow sixfold by 2030 (from one million to six million).** Due to this gap, universities and research centres are participating in a large range of government-backed talent programmes to attract and retain AI talent.

3.2.3. Rights and global governance

Generative AI has received attention regarding rights and global norms. The European Parliament, taking a long time in drafting its AI Act proposal, interrupted one of the final stages of the proposal due to concerns over those loopholes that may remain outside of the AI Act if not covered properly. Likewise, Spain, the US and the UK launched the OECD Global Forum on Technology, a platform for dialogue and cooperation on digital policy issues, where generative AI had a prominent role and was a top priority. The Forum has developed a set of principles on AI from a humanistic approach in order to foster trustworthiness, fairness, transparency, security, and accountability, among other values. The OECD's perspective on generative AI and digital rights is based on the idea that these technologies should be aligned with human values and serve the public interest. The **Global Partnership on AI (GPAI)** has developed some policy reports on the potential impact of generative AI on certain global issues, **such as the future of work.**

The Internet Governance Forum (IGF) addressed generative AI in the last IGF in Japan in October 2023. The IGF's perspective on generative AI and digital rights is based on the global idea that these technologies should be governed by all states. The UN agency, the International Telecommunication Union (ITU), approaches generative AI under the umbrella of the Sustainable Development Goals (SDGs) and bridging the digital divide.

It looks like, even amid global competition for generative AI, most international forums are increasing their positions and views on governing AI under a humanistic view.

3.3 The role of the EU in the global implications of generative AI

While the EU's position in the economy of generative AI is less prominent and significant than that of China's and the US, the EU has caught up on the potential security and rights challenges.

Regarding its economic position, there is no single European generative AI startup or company on the top list of firms worldwide. Venture Capital culture is limited in Europe, and there is an important lack of commercialisation of research and development towards the market.

3.3.1. Generative AI, economic security and critical technologies

The European Commission proposed in the summer of 2023 the first-of-its-kind Economic Security Strategy to address the economic security risks from certain economic flows and activities that may remain vulnerable or threatened in the current scenario of geopolitical tensions and accelerated technological development.

The European Economic Security Strategy is based on a three-pillar approach, or three Ps - promotion of the EU's economic base and competitiveness, protection against risks, and partnership with countries with shared concerns and interests. The four areas that require risk assessment are: resilience of supply chains, including energy security; physical and cybersecurity of critical infrastructure; technology security and leakage; and weaponisation of economic dependencies and coercion.

One of the first deliverables has been the list proposal on critical technologies by the European Commission, which encourages Member States to provide their risk assessments and lead to a collective work to determine which proportionate and precise measures should be taken to promote, protect and partner in specific technology areas. The goal is two-fold - to reduce dependencies on third actors whose supply chain and political security may be of high-risk, and to promote a diversification of strategic assets across the Union and with trusted partners.

Out of this list proposal, which contains 10 technology areas, the second priority is Artificial Intelligence. The technologies listed for this area, that should be assessed, but are not exhaustive and may include new ones, include high-performance computing, cloud and edge computing, data analytics technologies, and computer vision, language processing and object recognition. Most of these listed technologies have an immediate interaction with generative AI, either because generative AI helps them to improve or to optimise their solutions, or because the former depends on the latter to be developed and run.

From a practical perspective, several pressing challenges arise with regards to this list proposal of critical technologies and the embedding of generative AI. First, this list is aligned with President of the European Commission's ***de-risking* strategy.** Its goal is not to promote an inward market of endogenous manufacturing or keep away from global supply chains **(*decoupling*).** The objective is to promote global trade at the same time the EU reduces its dependencies from high-risk third actors and guarantees strategic assets' diversification with trusted partners **(*friend-shoring*).** While the ***de-risking*** approach was initially criticised by some countries that considered that EU's approach to China was not assertive, eventually the US accepted this discourse and the National Security Advisor, Jake Sullivan, has reiterated it during 2023.

However, from a hands-on approach, the reality is that the embedding of generative AI through a de-risking approach may prove to be hard, at least in present times, because this technology application still has several security loopholes and risks (as mentioned in previous chapters) that make a real-time, actual, comprehensive monitoring of the potential challenges it may represent for global security and economic resilience complex. **The EU-US Trade and Technology Council has been working since 2021 on how to develop joint early warning and monitoring systems for certain technologies** such as advanced semiconductors, but this monitoring process requires a high level of existing data, due diligence compliance and the establishment of fluid conversations and partnerships with the private sector that designs, produces and deploys these systems.

Additionally, it remains to be known the actual effect of export control regimes led by the US, and joined by the Netherlands and Japan, on certain semiconductors and AI components towards China. As stated in previous sections, the restriction on semiconductors in China would limit the access for Chinese companies to GPUs that are essential for Large Language Models.

Related to the latter point, one of the reasons why the **Economic Security Strategy** and the list of critical technologies were proposed was due to the identified lack of a comprehensive, fully-fledged coordination and cooperation across EU Member States on technology issues. While the Netherlands' decision to join the US-led export control regime is valid under the European typology of shared and exclusive competences, it showed EU institutions how an actual implementation of collective measures is very much needed.

When it comes down to generative AI, the main question will be how each Member State, when developing their national risk assessments to be sent to the European Commission before the end of 2023, will address generative AI - as a security risk, threat or challenge, as a purely economic issue, or as a topic that needs to be addressed only through regulation (namely, the AI Act proposal). As has occurred with other proposals, such as the **5G Cybersecurity Toolbox,** Member States may have different political, security and market approaches to the same issue.

A key point is how to address generative AI and economic security with regards to China. Neither the strategy nor the list proposal mentions the country explicitly. Some Member States look to a greater assertive approach regarding China, while others prefer not to do so.

3.3.2. Generative AI, the Brussels effect and the EU's regulatory powerhouse

The European Union has become a worldwide benchmark for a large part of technology-related regulations. **Throughout the early development of the AI Act proposal since April 2021**, generative AI was never mentioned. However, the increase in generative **AI use since the end of 2022** has made the actors involved in the proposal include generative AI as an important issue.

The European Parliament has proposed that generative AI systems be subject to three levels of obligations - **specific obligations for generative AI, specific obligations for foundation models,** and **general obligations applicable to all AI systems.**

The EU has been long flagged as the regulatory powerhouse when it comes to technology policy. The 'Brussels effect' phenomenon has been referred to over several years to explain how the EU's regulation influences how third countries apply their own technology legislation using a similar approach. In the case of generative AI, it remains to be seen the scope and depth generative AI receives in the AI Act proposal.

However, an interesting development is how the EU and some of its Member States are taking on a prominent voice in the global technology governance dialogues. Particularly, the Spanish government launched, jointly with the UK and the US, the OECD Global Forum on Technology, where generative AI was one of the two top priority topics. **Likewise, the Spanish Presidency to the EU Council** - which is expected to wrap up the AI Act file during the final semester of 2023 - is pushing towards including generative AI as an economic factor of opportunity and as a topic that needs to guarantee and protect fundamental rights.

3.3.3. Generative AI, the foreign policy of technology and multilateralism

The EU does not only address technology policy through regulation, market and security considerations. Technology has become an asset of foreign policy and an important topic of discussion in multilateral and mini-lateral fora. Two main frameworks help explain how the EU may push generative AI as part of the global discussions.

First, the Global Gateway, launched in December 2021, and which is the EU's major investment plan in infrastructure development with third countries, puts the digital pillar as its first priority. The work with partner countries focuses on digital networks and infrastructures, including Artificial Intelligence. Generative AI is not mentioned in the communication, partly due to the period of time when it was published. With the lengthy time it has taken to get projects released, it is not very

likely that generative AI will be a top priority in the Global Gateway, as there are other technologies with a greater level of capacity maturity, a greater opportunity of investments, due diligence compliance and trusted partnerships with third countries, and with a long-lasting background of public-private partnerships in those technology areas. However, generative AI is a topic not to be overlooked.

Second, the Council of the European Union approved in July 2022 the first-ever framework on EU's digital diplomacy. The European External Action Service received the approval to frame, through a formal action plan and operational initiatives, all international technology partnerships that the EU was involvedin. This would be undertaken with a higher level of coordination, under the same umbrella, and with a much more diplomacy-oriented arm, consisting of networks of flagship EU Delegations with technology policy as a top priority issue in their negotiations and activities.

The EU has been developing several partnerships with third countries. The EU-US Trade and Technology Council, initiated in 2021, has a specific working group on Artificial Intelligence. Both partners announced in December 2022 **the launch of the Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management,** whose implementation plan contains short- and long-term goals. Goals mainly cover the following areas: the establishment of inclusive cooperation channels, advancing shared terminologies and taxonomies, conduct an analysis on AI standards and identify those of interest for cooperation, development of tools on evaluation, selection, inclusion and revision, and the setting up of monitoring and measuring systems of existing and emerging AI risks.

One of the first deliverables has been the **launch of an initial draft of AI** terminologies and taxonomies, by engaging stakeholders. It has identified 65 terms that were identified with reference to key documents. The main goal is to move towards a common vision on how to govern AI and how to move up agendas to the multilateral settings of negotiation. Apart fort the expert working group on terminologies, another two working groups on standards and on emerging risks are ongoing.

Generative AI systems have received attention in the TTC. In May 2023, during the TTC in **Luleå, Sweden, a panel discussion was held on large AI models**, following the "recently witnessed acceleration of generative AI". The main output was for both the EU and the US to foster further international cooperation and a faster global approach to AI.

The EU counts on other international technology partnerships. However, the reference to generative AI has been far limited. First, because the level of depth and scope of topics in other dialogues is not as broad as in the transatlantic TTC. In some cases, the EU addresses the approach to AI with third countries in terms of ethics, definitions and taxonomies - for example**, the Japan-EU Digital Partnership Agreement addresses similar topics,** such as global supply chains, secure 5G,

digitalisation of public services, digital trade, global and interoperable standards, and digital education. It also addresses the topic on safe and ethical applications of AI.

However, there have been no major developments or joint declarations on this issue, except for their common membership to the G20 (the EU as a participant with no voting rights). The declaration on **the EU-Republic of Korea Partnership Agreement** points out to the need to discuss definitions, use cases, high risk AI applications and response measures, and facilitate cooperation towards relevant fora, such as the GPAI and OECD.

In other cases, the EU addresses the cooperation on AI through the lens of technical implementation. For instance, the focus of the **EU-Singapore Digital Partnership Agreement** is on interoperability, cooperation on AI testbeds and testing, cross-border access for AI technologies and solutions, capabilities, and references to trustworthiness, adoptability and transparency.

An important and increasing reference is cooperation on AI in international forums, where both sides may move up common agendas to these settings. Japan and Korea's partnership agreements refer to it. Likewise, the **EU-India Trade and Technology Council** encourages the *coordination* -which is a wording that goes beyond the concept of *cooperation*, and implies a greater level of joint work- within the GPAI. The **EU's interest in India as a technological powerhouse has increased in recent years.**

In some EU technology partnerships with regions and countries, there is no explicit reference to Artificial Intelligence, but the development of technological capacity building in other verticals is an anteroom for future areas of interest that may arise in coming years. To give an example, **the launch of the EU-LAC Digital Alliance in early 2023 focused on regulatory convergence, digital infrastructure, the role of satellites (Earth observation, technology solutions for hazardous climate responses), talent, R&I, data centers, data spaces, but less reference to AI.**

In the case of the **African continent**, the wording "digital" has been present since the 2017 Sixth EU-Africa Business Forum, that highlighted the role of the digital economy as a driver. The extension of digital topics has experienced an upsurge, on digital skills, broadband connectivity, cross-border backbone infrastructure, e-services, but limited references to AI. Digital for Development (D4D) Hubs in both Latin America and the Caribbean and Africa have set up workshops and training activities on Artificial Intelligence.

Overall, the reference to generative AI is still limited across the different global governance regimes, also in the case of the EU's international technology partnerships. This is a trend across all countries, regional organisations and international agencies. However, the launch of the first-ever United Nations High-Level Advisory Body on AI is a window of opportunity to plug generative AI into the discussion as an important issue to be tackled by the diverse groups of representatives on a geographic and cultural basis.

3.4 A much-needed further policy discussion and framing on the impact of generative AI on international affairs

Evidence shows that China and the US are quickly overtaking the global competition on generative AI However, their instruments of power differ. While China leads in intellectual property and patents, and has a large number of generative AI startups and is building up a strong ecosystem of companies, the US still leads the way in the amount of investments in this technology vector, mostly through Venture Capital, with investments in high-risk markets, marked by uncertainty and potential impact.

The impact of generative AI is three-pronged and introduces both opportunities and challenges. In security, it touches on its military applications, the impact on intelligence community's decision-making processes, and the growth of hybrid threats triggered by generative AI, such as deepfakes and foreign information manipulation interference (FIMI). In economy, it may lead to new cross-sector and cross-industry collaborations to foster generative AI development. At the same time, it may become an area of competition with similar trend patterns such as market concentration. In rights and global governance, main forums have started in these two last years (mostly in 2023) to include early discussion panels on the impact of generative AI. Some issues have been touched upon, such as the future of work and the impact on human rights. However, there is still more to do to increase policy discussions which are capable and comprehensive in setting down complete principles, roadmaps and action plans.

In this scenario, the European Union should play a key role. Evidence shows that, from the economic perspective, it lags behind with regards to large corporations and startups alike. In terms of Venture Capital, it is also way behind and may fail to be in the top three worldwide. Where security and rights are concerned, the EU is moving forward, with generative AI falling under the umbrella of the AI Act with a three-level set of obligations and contributing to discussions in international fora. Still, the EU's bilateral technology partnerships with trusted countries and like-minded partners should deepen the conversation on generative AI. This is an opportunity to have a first-mover advantage in this area of foreign policy of technology, and also to agree on common principles and be more influential in the international agendas.

Chapter 4: AI readiness and the economic potential, focusing on Southern EU

Artificial Intelligence is a field of study in computer science that uses datasets as input to provide problem-solving output. The question if machines can think was first posed by Alan Turing in his paper[64] in 1950 in which he introduced the 'Turing Test' to determine whether a computer can demonstrate human-like intelligence. In 1956, John McCarthy[65] was the first to use the term AI and later that year Allen Newell, Herbert A. Simon, and Cliff Shaw developed the first AI program, the 'Logic Theorist', which performed automated reasoning. Since then, the field of AI has been slowly evolving mainly because of its inability to handle large problems due to combinatorial explosion. This limitation was surpassed in the early 2010s with the evolution of deep learning, a sub-field of machine learning, which is based on artificial neural networks with multiple layers and representation learning. Today, **AI technology has displayed capabilities of reasoning, perception, decision making, knowledge representation and natural language processing.** These capabilities have enabled applications such as search engines, visual assistants, targeted advertising, language translation and recommendation systems but also dedicated applications like autonomous vehicles, medical diagnosis, or supply chain management. The use of AI can provide a competitive advantage to firms and economies with financial and societal benefits for all industry sectors and social activities, and especially in high-impact sectors, such as healthcare, financial services, retail, the public sector, automotive and transportations, agriculture, and the energy sector. However, the same elements that drive the socio-economic benefits of AI can also create new risks or negative consequences for society, mainly concerning data protection, digital rights, and ethical issues.

4.1 Artificial Intelligence SWOT analysis[66]

The implications of AI are vast and cover the entire economy and society functions since its possible applications refer to every aspect of everyday life. Here we present a non-exhaustive SWOT (Strengths, Weaknesses, Opportunities and Threats) analysis of AI.

[64] Turing, A. M. (1950). Computing Machinery and Intelligence. Mind, 59(236), 433–460.
[65] McCarthy, John; Minsky, Marvin; Rochester, Nathan; Shannon, Claude (1955). "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence"
[66] I. Palomares, E. Martínez-Cámara, R. Montes, P. García-Moral, M. Chiachio, J. Chiachio, F. Herrera, A panoramic view and swot analysis of artificial intelligence for achieving the sustainable development goals by 2030: Progress and prospects. Appl. Intell., 1–31 (2021). https://doi.org/10.1007/s10489-021-02264-y

Strengths

- Smart cities and intelligent transportation systems enhancing efficient commuting and flexible working.
- Intelligent sensors and 5G network for real-time infrastructure monitoring.
- Data mining on credit card transactions for fraud detection.
- Intelligent personal assistants, sensing multi-functional drones and autonomous vehicles are significantly facilitating citizen lives.
- AI-based personalisation to provide content adapted to learners' individual needs.
- Well trained AI systems remove gender bias in recruitment and similar decisions.
- Smart grids favour energy efficiency and its timely supply at an optimal cost.
- AI models help in making better emergency or disaster recovery decisions.

Weaknesses

- Replacement of non-qualified workers by robots or algorithms in least developed countries.
- Absence of integrated data platforms hamper the use of intelligent systems for infrastructure design.
- Open-AI regulations insufficiently extended, owing to lack of information exchange.
- AI helps detecting fraud but still lacks solutions to raise awareness against it.
- More attention needed in adapted and location-based teaching systems to yield equity, efficiency, and quality education.
- Biased machine learning without data that intentionally reinforce vulnerable collectives.
- Data centres account for an estimated 1% to 8% of global energy consumption.
- Political resistance and economic cost of large-scale AI systems to optimise pollutant emissions in urban areas.

Opportunities

- Digital labour and external outsourcing as an engine to create employment.
- AI for designing environmental risk maps helps in recovering from disasters.
- Integrating AI in urban design and planning guided by e-government.
- Blockchain and cryptocurrencies as a driving force for secure digital transactions and administrative processes.
- Intelligent tutors scale-up possibilities towards one-to-one education tailored to the individual and supporting students with special needs.
- AI detection of dishonest, bullying or harassing behaviours would help in mitigating legal laws by providing evidence to prevent fatal consequences.

- Smart, energy efficient and sustainable buildings to reduce energy consumption.
- Prediction of energy needs and traffic helps in reducing pollutants with ecological impacts.

Threats

- Increased inequalities by emergence of AI and robotics in work.
- Industry 4.0 brings socio-economic risks in developing countries due to job losses.
- Barriers to bridge gaps between public and private sectors and achieve data interoperability.
- A wrong use of AI for fighting fraud could aggravate it and reveal new security breaches.
- Equal access to technologies and AI training is still not a universal right.
- New forms of digital harassment in social media.
- Smarter digital energy systems are also more vulnerable to cyberattacks.
- Climate change implies obsolescence in AI models to predict natural catastrophes on data.

4.2 OECD Artificial Intelligence Principles

The OECD has issued the 'AI Principles', that were adopted in 2019, to promote the use of AI that is innovative, trustworthy and respects human rights and democratic values[67,68,69]. The PromethEUs network countries are OECD members and their relevant national policies and framework adhere to these principles.

Principle 1.1 - Inclusive growth, sustainable development, and well-being

This principle recognises that a trustworthy AI can be used for social good and support the Sustainable Development Goals (SDGs) in areas such as education, health, transport, agriculture, environment, and sustainable cities, among others. It addresses the potential shortcomings of unequal technology access between developed and developing countries and the disparate impact on low- and middle-income countries. An AI system could sustain the existing social biases and have a negative effect on vulnerable populations so, it should be designed to prioritise the prosperity of all society members and help reduce inequalities.

Principle 1.2 - Human-centred values and fairness

AI should be based on human-centred values, such as fundamental freedoms, equality, social justice, data protection and privacy, as well as consumer rights and commercial fairness. This principle recognises the importance of measures such as human rights impact assessments (HRIAs) and human rights due diligence, codes of ethical conduct, quality labels, certifications, and human determination (i.e., a "human in the loop"). An AI system should include safeguards to ensure a fair

[67] OECD/LEGAL/0449 - Recommendation of the Council on Artificial Intelligence
[68] OECD (2019). Artificial Intelligence in Society. doi.org/10.1787/eedfee77-en.
[69] OECD/LEGAL/0463 - Recommendation of the Council on Enhancing Access to and Sharing of Data

society and allow for supervision, and intervention if needed, to protect and promote human rights and values as well as to reduce discrimination or other unfair outcomes.

Principle 1.3 - Transparency and explainability

In this principle, transparency refers to the need of disclosure when AI is being used. It will assist people make informed decisions by understanding how an AI system is developed, trained, operates, and is deployed. In addition, transparency will help to raise awareness and understanding of AI systems paving the way for social acceptance. On the other hand, explainability is about conveying information to people that are affected by the AI system on how the outcome was reached. Here there are privacy and security concerns that should be considered, as well as possible increases in complexity and costs that could disproportionately affect SMEs that are AI actors.

Principle 1.4 - Robustness, security, and safety

The safety and security of AI systems is essential to instil trust in people and robustness is critical to ensure that an AI system can persevere in the face of digital security risks. Today there are laws and regulations concerning safety risks and these should extend to AI. AI actors could take a risk management approach to identify and fortify against misuse, that is, use of AI systems for purposes other than those for which they were originally designed. Besides the risk management approach, another way for robust, secure, and safe AI systems is to foster traceability which enables analysis and inquiry into the outcomes and promotes accountability by maintaining records of data characteristics but not necessarily the data themselves.

Principle 1.5 - Accountability

Organisations and individuals developing, deploying, or operating AI systems should be held accountable for their proper operation. Accountability here pertains to ethical, moral, or other expectations, that guide individuals' or organisations' actions or conduct and allows them to explain reasons for which decisions and actions were taken. AI actors are expected to warrant the proper functioning of the AI systems that they develop and/or operate while respecting the regulatory frameworks.

Furthermore, governments are encouraged to implement the following Principles, consistent with the above AI Principles, in their national policies and international co-operation, with special attention to Small and Medium-Sized Enterprises (SMEs).

Principle 2.1 - Investing in AI research and development

Governments should consider long-term public investment, and encourage private investment, in research and development to promote innovation in trustworthy AI that focuses on challenging technical issues and on AI-related social, legal, and ethical implications and policy issues. In addition, these investments should also refer to open datasets that are free of inappropriate bias and respect privacy and data protection.

Principle 2.2 - Fostering a digital ecosystem for AI

Governments should foster accessible AI ecosystems with digital infrastructure and technologies, and mechanisms to share data and knowledge, considering their national frameworks. The necessary digital technologies and infrastructure include access to affordable high-speed broadband networks and services, computing power and data storage as well as supporting data-generating technologies such as the Internet-of-Things (IoT). The appropriate mechanisms for sharing AI knowledge, include data, code, algorithms, models, research, and know-how, and they must respect privacy, intellectual property and other relevant rights.

Principle 2.3 - Shaping an enabling policy environment for AI

Governments should promote a policy environment that supports the transition from the research and development stage to the deployment and operation stage for trustworthy AI systems. To this effect, they should consider using experimentation to provide a controlled environment in which AI systems can be tested, and scaled-up, as appropriate. In addition, these policies should be reviewed and adapted accordingly to encourage innovation and healthy competition for trustworthy AI.

Principle 2.4 - Building human capacity and preparing for labour market transformation

Governments should cooperate with stakeholders to prepare for the transformation of the world and to that end, empower people to effectively use and interact with AI systems, including by equipping them with the necessary skills. A fair transition of the workforce should include social dialogue, training programmes throughout the working life, support for those affected by displacement, and access to new opportunities in the labour market. In addition, the responsible use of AI at work should be promoted to enhance the safety of workers and the quality of jobs, to foster entrepreneurship and productivity, and aim to ensure that the benefits from AI are broadly shared in a fair way.

Principle 2.5 - International co-operation for trustworthy AI

Governments, including developing countries, and with stakeholders, should actively cooperate to advance these principles and to progress on responsible supervising of trustworthy AI. They should foster the sharing of AI knowledge, and encourage international, cross-sectoral, and open multi-stakeholder initiatives to garner long-term expertise on AI. Governments should promote the development of multi-stakeholder, consensus-driven global technical standards for interoperable and trustworthy AI, as well as encourage the development, and their own use, of internationally comparable metrics to measure AI research, development, and deployment, and gather the evidence base to assess progress in the implementation of these principles. Portugal has issued 36 initiatives regarding policies on AI, Spain 28, Italy 10, and Greece 3.

Table 2: Number of relevant initiatives to each principle is presented for each country.

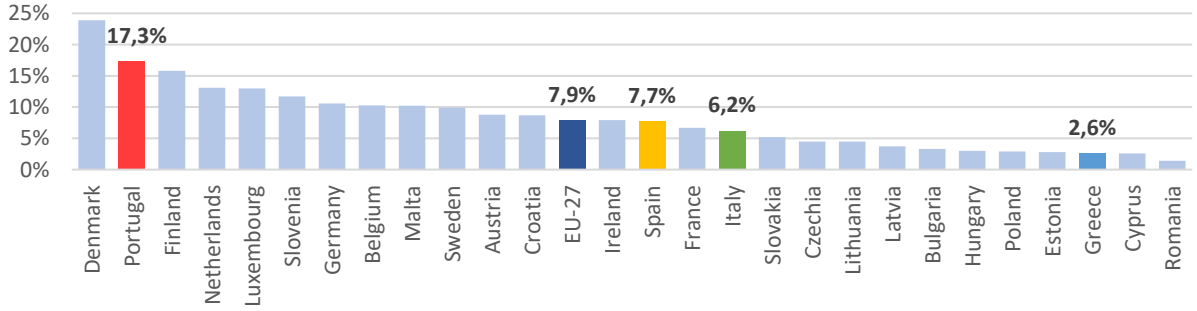
	Principle 1.1	*Principle 1.2*	*Principle 1.3*	*Principle 1.4*	*Principle 1.5*	*Principle 2.1*	*Principle 2.2*	*Principle 2.3*	*Principle 2.4*	*Principle 2.5*
Portugal	24	14	12	19	8	15	19	12	11	8
Spain	16	8	6	9	3	11	15	6	12	8
Italy	3	3	1	3	1	5	7	2	2	0
Greece	2	2	0	1	0	1	2	2	1	1

Source: OECD

4.3 Artificial Intelligence in PromethEUs network countries

In 2021, only 7.9% of EU enterprises used at least one AI technology. Portugal is one of the pioneering countries in the EU ranking second with 17.3%, only behind Denmark (23.9%). Spain's performance is almost equal to the EU average with 7.7%, Italy is relatively lagging with 6.6% whereas Greece is second to last with 2.6%, tied with Cyprus and ahead only of Romania (1.4%).

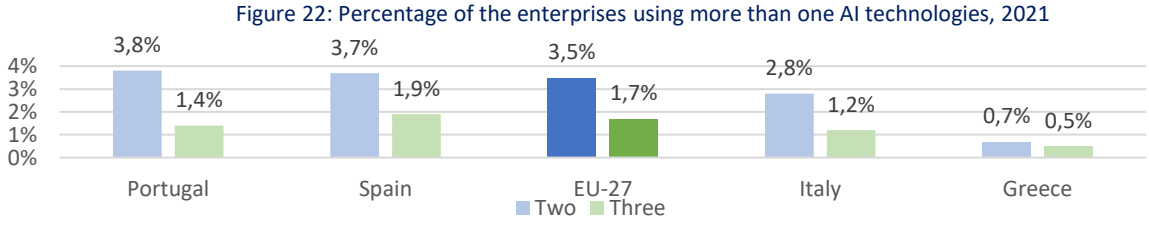
Figure 21: Percentage of the enterprises using at least one AI technologies, 2021



Source: Eurostat

The percentage of enterprises in Spain with a higher level of AI technology penetration is greater than the EU average and Portugal trails behind only when it comes to three implemented AI technologies. Italy performs relatively better for higher adoption while enterprises in Greece fall behind even more for two or more AI technologies.

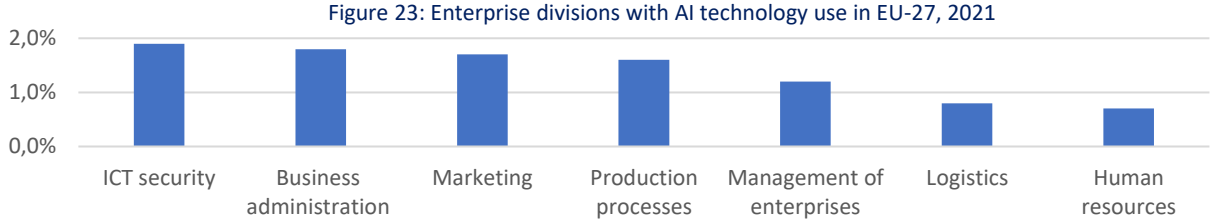
Figure 22: Percentage of the enterprises using more than one AI technologies, 2021



Source: Eurostat

In 2021, the EU enterprises mainly used AI in ICT security and business administration processes followed by marketing division and production processes.

Figure 23: Enterprise divisions with AI technology use in EU-27, 2021



Source: Eurostat

a. Greece

In Greece, **the Digital Transformation Strategy 2020-2025[70], which was adopted in 2021,** consists of **interventions in the digital infrastructure, in the education and training of the population regarding digital skills as well as in the digital technology implementation in all sectors of the economy and the public administration.** It describes the principles upon which the model of the digital transformation is built, and involves 475 projects, classified in short- and medium-term, horizontal and sectoral, of which 146 are currently underway. **It is structured on 6 strategic axes - connectivity, digital public services, digital skills, digital business, digital innovation and advanced technologies.** The latter includes a national strategy for the development and utilisation of AI which describes the national priorities and analyses the actions that will lead to pilot applications per policy area as well as the possibilities of using AI in PAs. Digital challenges for Greece include a shortfall in connectivity, a lack of digital skills, a slow uptake of digital technologies, especially by SMEs, and a low level of digital public sector services[71]. Greece's RRP supports the digital transition with investments and reforms in the digitalisation of PAs and private sector companies, in connectivity, and in digital skills. It has allocated € 130 million in investments for the deployment of fibre optic infrastructure in buildings, € 1.3 billion in the digital transformation of the public sector and € 375 million for the digitalisation of businesses. In addition, € 500 million will be invested in the promotion of the digital transformation in the education and health systems, and € 750 million in digital upskilling to enhance basic digital literacy across the entire population.

The Artificial Intelligence market is projected to reach a volume of US$ 387m in 2023. It is expected to have a Compound Annual Growth Rate (CAGR 2023-2030) of 17.17%, corresponding to a volume

[70] digitalstrategy.gov.gr/en/

[71] https://commission.europa.eu/business-economy-euro/economic-recovery/recovery-and-resilience-facility/country-pages/greeces-recovery-and-resilience-plan_en

of US$ 1.17bn by 2030[72]. **There had been no venture capital investments in AI startups before 2020 and in the last 3 years these investments were less than € 20 million.** Research on AI has steadily increased over the last two decades with a noticeable increase in the last five years.

Figure 24a: Venture capital investments in Greece

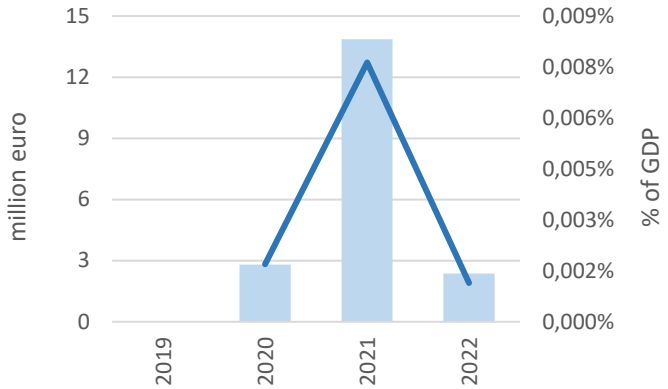
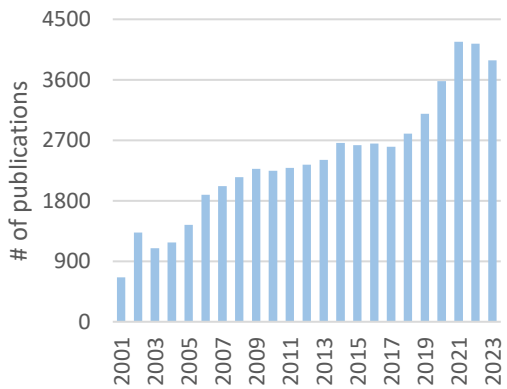


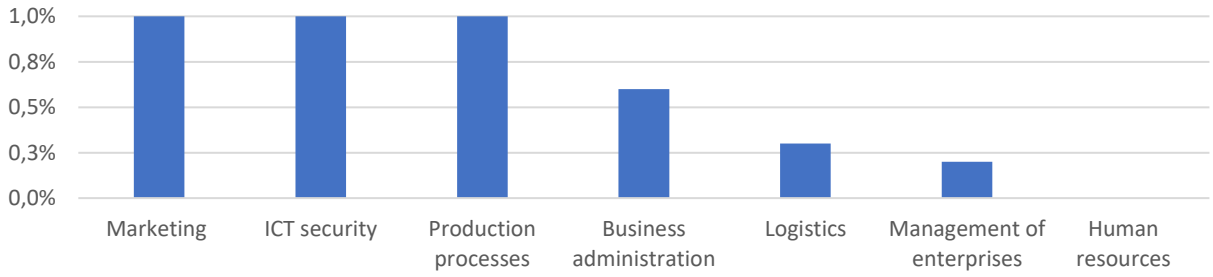
Figure 24b: AI publications in Greece



Source: OECD

In Greece, AI technologies are mostly used in marketing, ICT security, production processes and are notably absent from the human resources divisions of enterprises.

Figure 25: Enterprise divisions with AI technology use in Greece



Source: Eurostat

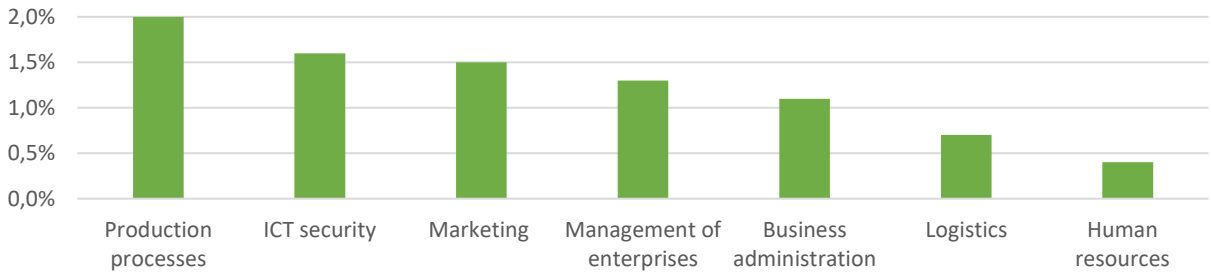
b. Italy

In 2021, Italy adopted the **"AI Strategic Programme"[73] for 2022-2024 which aims at the strengthening of skills and attracting talent to develop an Artificial Intelligence ecosystem, increasing the funding for advanced research in AI and promoting the adoption of AI and its**

[72] https://www.statista.com/outlook/tmo/artificial-intelligence/greece
[73] https://assets.innovazione.gov.it/1637937177-programma-strategico-iaweb-2.pdf

applications both in PAs and in the private sector. The strategy's objectives are to strengthen research in AI, to reduce the fragmentation of AI research, to develop and adopt human-centric and trustworthy AI, to increase AI-based innovation and AI technology development, to develop AI-based policies and services in the public sector and to create, retain and attract AI researchers in Italy. In Italy, the digital transformation faces challenges in improving the digital skills of the population, the workforce and the digitalisation of businesses while, the digital public services must be improved, and key e-government projects should be accelerated[74]. **Italy's RRP supports the digital transition with investments of € 6.7 billion in connectivity with 5G and optic fibre network, € 13.4 billion for the digital transition and innovation of the production system, through incentives for investments in cutting-edge and 4.0 technologies, RDI and 4.0 training activities, as well as € 6.1 billion for the digitalisation of the PA.** These investments are supported by a set of reforms to procure ICT solutions in a more timely and efficient way by the PA, to support the digital transformation of central and local administrations, as well as to support the adoption of cloud solutions by PAs and eliminate the bureaucracy hurdles that slow down the data exchange processes between administrations. In Italy, AI technologies are mostly used in production processes, ICT security, marketing, management and business administration processes.

Figure 26: Enterprise divisions with AI technology use in Italy



Source: Eurostat

The Artificial Intelligence market is projected to reach a volume of US$ 4.66bn in 2023. It is expected to have a Compound Annual Growth Rate (CAGR 2023-2030) of 17.86%, corresponding to a volume of US$ 14.72bn by 2030[75]**. In 2022, the venture capital investments in AI startups surpassed € 375 million, more than the cumulative investments in the previous 10 years, 2012-2021.** Research on AI has steadily increased over the last two decades with a noticeable increase in the last four years.

[74] https://commission.europa.eu/business-economy-euro/economic-recovery/recovery-and-resilience-facility/country-pages/italys-recovery-and-resilience-plan_en
[75] https://www.statista.com/outlook/tmo/artificial-intelligence/italy

Figure 27a: Venture capital investments in Italy

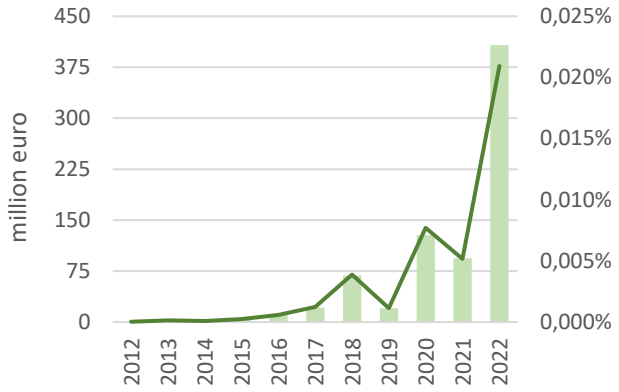
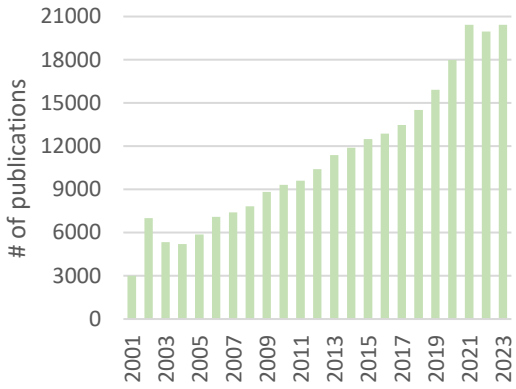


Figure 27b: AI publications in Italy



Source: OECD

c. Portugal

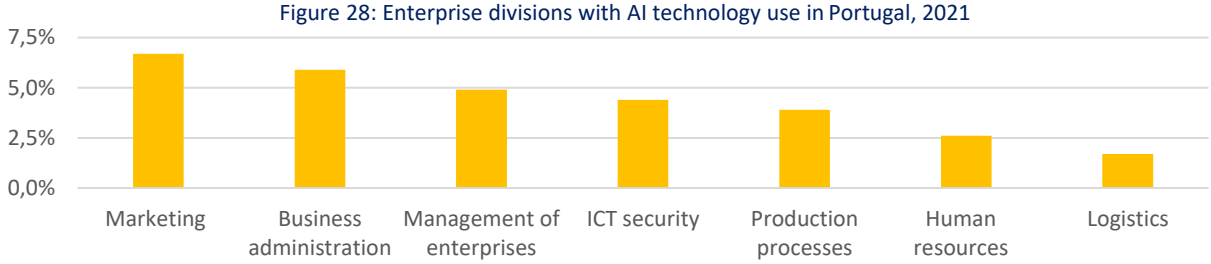
Portugal implemented a national strategy for AI in 2019, the **"Estratégia Nacional para Inteligência Artificial" which is aligned with the Coordinated Action plan of the EU and is included in INCoDe.2030[76], the Portuguese initiative to foster digital skills.** The AI Strategy aims to promote the engagement of citizens and key stakeholders to build up a knowledge intensive labour market and create a community of forefront companies producing and exporting AI technologies supported by research and innovation. It is structured on seven axes: to promote sustainability, resources management and employment, to foster AI skills, to promote AI skills and assimilate AI services into the economy, to promote experimental new developments, to solidify AI niche markets through key services, to generate new knowledge and developments through AI research and innovation as well as to provide better public services for citizens and businesses. **Digital challenges for Portugal include the need to invest in the digital transition, particularly in the development of digital skills, in the use of digital technologies for equal access to education and training, and to boost firms' competitiveness since the Portuguese economy includes many small-sized enterprises concentrated in traditional sectors[77].** Portugal's Recovery and Resilience Plan supports the digital transition with investments and reforms in the areas of skills, digitalisation of education and businesses as well as in the public sector. **The reform is supported by investments of € 710 million for the modernisation of vocational education and training institutions.** In the health sector, the € 300 million investments aim to modernise the computer systems of the National Health Service and to increase the digitalisation of medical records while, € 650 million will be invested in the business

[76] https://www.incode2030.gov.pt/

[77] https://commission.europa.eu/business-economy-euro/economic-recovery/recovery-and-resilience-facility/country-pages/portugals-recovery-and-resilience-plan_en

sector for the digitalisation of small and medium enterprises and for their workers digital skill training. In Portugal, AI technologies are mostly used in marketing, business administration processes and management while the ICT security and production processes trail behind.

Figure 28: Enterprise divisions with AI technology use in Portugal, 2021



Source: Eurostat

The Artificial Intelligence market is projected to reach a volume of US$ 933 m in 2023. It is expected to have a Compound Annual Growth Rate (CAGR 2023-2030) of 18.83%, corresponding to a volume of US$ 3.12 bn by 2030[78]**. In 2021-2022 the venture capital investments in AI startups surpassed € 600 million, more than the cumulative investments in the previous 9 years, 2012-2020**, while € 313 million were invested in 2018 alone**.** Research on AI has steadily increased over the last two decades with a noticeable increase in the last five years.

Figure 29a: Venture capital investments

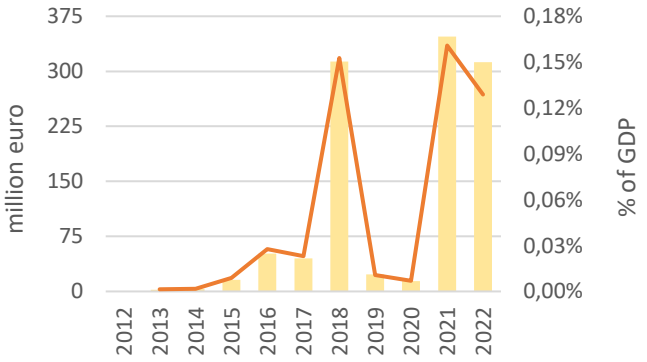
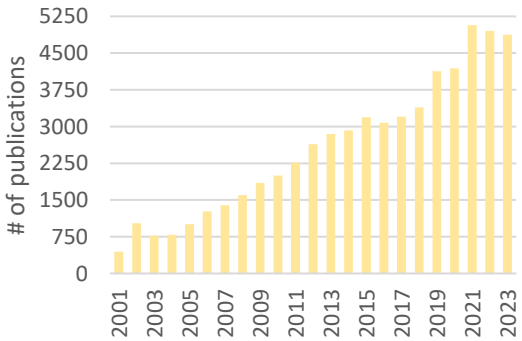


Figure 29b: AI publications in Portugal



Source: OECD

d. Spain

Spain introduced in 2020 the national AI strategy, **"Estrategia Nacional de Inteligencia Artificial (ENIA)"[79],** revolving around six strategic axes - to foster scientific research, technological

[78] https://www.statista.com/outlook/tmo/artificial-intelligence/portugal
[79] https://www.lamoncloa.gob.es/presidente/actividades/Documents/2020/ENIA2B.pdf

development and innovation in Artificial Intelligence, **to promote the digital capacities, the development of national talent and the attraction of international talent, to develop data platforms and technological infrastructures that support AI, to integrate AI into the value chains, to support the use of AI in the PA and in national strategic missions, to establish an ethical and regulatory framework that guarantees the protection of individual and collective rights, with social welfare and sustainability.** The main objectives include the promotion of scientific excellence and innovation in AI, the projection of the Spanish language, the creation of qualified employment, the transformation of the productive model, to foster an environment of trust with humanistic values in inclusive and sustainable Artificial Intelligence. **In addition, the Spanish Digital Agenda 2025[80] was approved in 2020, which seeks to advance the digital transformation and foresees plans to develop AI, including the National Strategy on AI, resulting in the creation of the Spanish Artificial Intelligence Council to assist in public policy design and implementation.** In Spain, a large share of the population has a low level of digital skills, and the workforce lacks in workers with specialist digital skills. This hampers digitisation in general and is one of the barriers to investment in Spain[81]. The RRP includes the Digital Spain Agenda 2025, the 5G cybersecurity law, the Artificial Intelligence Strategy, the Digital Skills Plan and a law on telecommunications to upgrade the regulatory framework with the development of new regulatory and enforcement instruments. **The RRP supports the digital transition with investments of € 3.6 billion in digital skills training, € 4.5 billion in the digital transformation of the public administration, with a special focus on the justice, health care, employment, educational and social services systems. It also includes € 10.2 billion in investments to promote the digitalisation of industry and SMEs, in artificial intelligence and investments in artificial intelligence while, € 15.4 billion are meant to support fixed and 5G connectivity, data infrastructure and the related ecosystem.** The RRP was recently modified to further strengthen the digital transition with € 40.4 billion of the available funds, from € 19.7 billion in the original plan, for measures that support digital objectives. In Spain, AI technologies are mostly used in production processes, marketing, ICT security and business administration processes.

[80] https://espanadigital.gob.es/sites/agendadigital/files/2022-01/Digital-Spain-2025.pdf
[81] https://commission.europa.eu/business-economy-euro/economic-recovery/recovery-and-resilience-facility/country-pages/spains-recovery-and-resilience-plan_en

Figure 30: Enterprise divisions with AI technology use in Spain, 2021

Source: Eurostat

The Artificial Intelligence market is projected to reach a volume of US$ 4.00bn in 2023. It is expected to have a Compound Annual Growth Rate (CAGR 2023-2030) of 18.22%, corresponding to a volume of US$ 12.91bn by 2030[82]. In 2021-2022 the venture capital investments in AI startups surpassed € 2 billion, more than the cumulative investments in the previous 9 years, 2012-2020. Research on AI has steadily increased over the last two decades with a noticeable increase in the last six years.

Figure 31a: Venture capital investments in Portugal

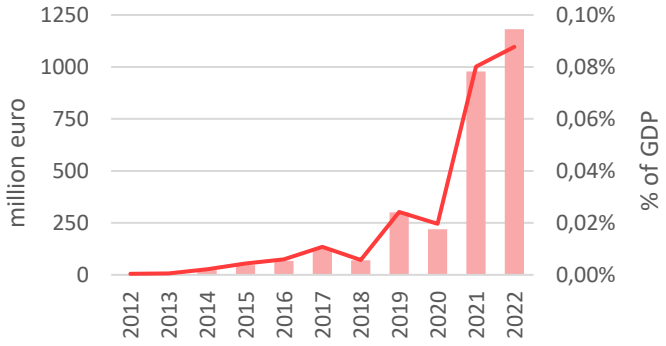
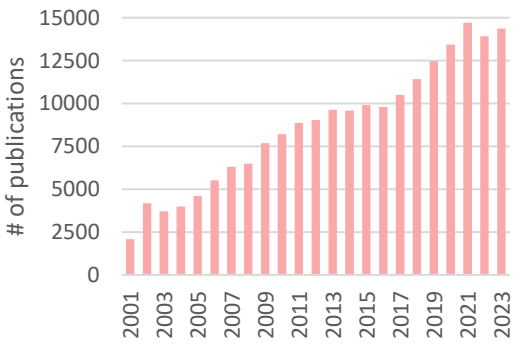


Figure 31b: AI publications in Portugal



Source: OECD

[82] https://www.statista.com/outlook/tmo/artificial-intelligence/spain

Chapter 1:
Regulatory perspective on AI Act

(**IPP –** Steffen Hoernig, André Ilharco)

The AI Act is a consistent proposal for a regulatory framework, **but crucial issues still need to be agreed on, i.e., those under discussion in the trilogue and the terms and role of human oversight.** Even after these clarifications, AI regulation in the **EU will need to be updated regularly** in the future (as shown by the absence of foundation models and GPAI in the Commission's initial proposal) **due to:**

- the **everchanging nature and capabilities of AI**
- potential unintended **legal fragmentation resulting from the AI Act subsidiarity dispositions** (penalties, high-risk classification)

There **is little reason to expect a Brussels effect this time round**. However, it seems reasonable to hope for international convergence and at least some impact of EU regulatory principles on the future international framework.

The AI Act is a consistent proposal for a regulatory framework, **but crucial issues still need to be agreed on, now in the trilogue and further down the road.** First, the European Parliament and Council versions of the AI Act contain conflicting visions on key issues, **such as the definition of AI itself** (essential to outline the scope of the Act's application), **the use of biometrics, the types and attributions of the institutions that will coordinate and assist the Act's implementation** (whether a Board is enough, or whether a more resourceful and independent structure is needed), **which limits on penalties shall apply and what they will depend on** (AI risk, AI transparency, economic development), **how the Act will address foundation models and general-purpose AI**, and finally how the risk-classification AI systems will be adjusted over time.

Other problems need to be fixed, but most likely not during these rounds of trilogue, such as **human oversight obligations** (including the meaning of what this is), **potential legal fragmentation** (role of Member States both in defining penalties and their power to classify AI systems' risk-level). Furthermore, the ever-changing nature of AI developments makes it clear that future adjustments will be needed. **All EU institutions agree that the Act will be re-evaluated in the future, making the AI Act a work-in-progress even after its entry into force.**

The EU has strong reasons to quickly come to an agreement on the AI Act. As we have seen in parts 1 and 2 of the IPP chapter, the predicted impacts of AI are significant, and other countries have started work on their own regulatory frameworks (or at least, engaged in the international debate). **While China and the US are showing signs of a proactive position in terms of AI regulation**, the UK. still retains a "wait and see" position. **Although the international regulatory discussion does not involve unbridgeable contradictions, coordination and agreement still have a long way to go.** This creates much uncertainty for EU businesses and SMEs, which will also feel the impact of the regulatory approaches of international actors other than the EU. Lastly, the comparison of EU regulation with other approaches will also help to evaluate the Act's impact on EU businesses, more specifically, it will let us understand better whether legal certainty will outweigh potentially burdensome compliance costs.

Finally, we highlight that there is **little reason to expect a Brussels effect this time around, since the biggest AI companies and investments are outside the EU. However, the EU's early proposals may serve as an example and have a coordinating effect on others' regulations** (as it seems to be having on the U.S.). While, instead, there may be a Washington-Beijing-London effect, common economic and safety interests leave reasonable hope for international convergence and at least some impact of EU regulatory principles on the future international framework.

Chapter 2:
The impact of Generative AI

(**I-Com -** Stefano da Empoli, Maria Rosaria Della Porta)

Generative AI is an advanced form of artificial intelligence that enables machines to learn from existing data to create new data or content, including audio, code, images, text, simulations and videos. **Preliminary estimates of the potential impact of generative AI on the world economy are impressive.**

However, **automation processes induced by generative AI will impact on knowledge work**, particularly activities involving decision making and collaboration, which previously had the lowest potential for automation. Therefore, a reorganization of and retraining in work is essential.

At the same time, generative AI may be the first type of automation capable of reducing inequality rather than increasing it, because it is actually based on language and, thus, can mimic higher skills compared to previous innovation waves. It is up to us humans to understand how to best use it. If we see it as a substitute for workers, we indeed risk high unemployment or wage compression as

salaries would then have to compete with machine costs. However, if we recognize it as a complement that can enhance overall work performance, we can lay the foundation for a manageable transition, where different tasks than before are performed, but in most cases to the advantage of both workers and companies. **Policy actions are needed to ensure that current and future workers and companies, particularly SMEs, do not fall behind in the adoption of these new tools**, further widening existing gaps.

As generative AI is already becoming an increasingly prominent part of everyday business activities and our daily lives, it gives rise to **several risks and ethical considerations, such as misinformation and deepfakes, cybersecurity, privacy and copyright issues**. However, **these risks could be manageable by fittingly adapting the current regulatory system to old and new challenges. Notwithstanding the potential risks, the huge potential benefits for Europe, and especially for the Member States currently lagging behind in terms of digital skills, should not be overlooked.**

For this reason, the EU should aim not only at adopting these technologies but also to play a part in their development. With much less frequency, among the many concerned statements often made recklessly about AI, a fact emerges that should disturb the sleep of the European decision-makers - **Europe is completely excluded from the leading group that is making AI history**. However, many of the researchers and some managers working for companies and research centers located outside of Europe are of European origin, and in some cases retain the passport of one of the EU Member States, where they often come for a holiday or conference. Brussels, which, in 2018, had started a bit late but with the necessary clarity regarding an AI strategy and coordinated plan, envisioning rules and investments as two pillars that reinforced each other, ended up betting almost entirely on the former rather than the latter. **R&D investments in AI not only should be increased by Member States but coordinated and centralized as much as possible if Europe wants to be a key player and not only a spectator in this technological change.**

Chapter 3:
The geopolitics of Artificial Intelligence

(**Elcano –** Raquel Jorge Ricart, Pau Álvarez-Aragonés)

Evidence shows that China and the United States are racing ahead in the global competition on generative AI. However, their instruments of power differ. While China leads in intellectual property and patents, and has a large number of generative AI startups and is building up a strong ecosystem of companies, the US still leads the way in the total of investments devoted to this technology vector, mostly through Venture Capital, which invests in high-risk markets, marked by uncertainty and potential impact.

The impact of generative AI is three-pronged and introduces opportunities and challenges. In **security, it touches on its military applications,** the impact on intelligence community's decision-making processes, and the growth of hybrid threats triggered by generative AI, such as deepfakes and foreign information manipulation interference (FIMI). **In economy, it may lead to new cross-sector and cross-industry collaborations to foster generative AI development.** At the same time, it may become an area of competition with similar trend patterns such as market concentration. **In rights and global governance, main forums have started in these last two years -mostly in 2023- to include early discussion panels on the impact of generative AI.** Some issues have been touched upon, such as the future of work and the impact on human rights. However, **more is still required to set underway policy discussions which are capable and comprehensive enough to set down thorough principles, roadmaps, and action plans.**

In this scenario, the European Union should play a key role. Evidence shows that, from the economic perspective, it lags behind with regards to large corporations and startups alike. In terms of Venture Capital, it is also losing the race and may fail to be in the top three worldwide. Where security and rights are concerned, the EU is moving forward, with generative AI falling under the umbrella of the AI Act with a three-level set of obligations, contributing to discussions in international fora. Still, the EU's bilateral technology partnerships with trusted countries and like-minded partners should deepen the conversation on generative AI. This is an opportunity to gain a first-mover advantage in this area of foreign policy of technology, and also to agree on common principles and be more influential in the international agendas.

Chapter 4:
AI readiness and the economic potential, focusing on the Southern EU

(**IOBE -** Aggelos Tsakanikas, Konstantinos Valaskas)

The evolution of artificial intelligence in recent years has already had a significant impact on society and it is set to profoundly alter everyday life for individual and firms. The rapid growth of technological advances in previous decades has paved the way for data analysis but due to computational restrictions we were unable to tap in its true potential until the emergence of AI in recent years. Today, AI is deployed in numerous economic sectors and its impact is driven by productivity gains of firms with the automation of processes and the support of the workforce with AI technologies, but also by the increased consumer demand which stems from customised and higher quality products and services.

The deployment of AI has altered business models, the decision-making and risk management processes, and the overall performance of firms with substantial economic and social benefits, but it also has major implications for society. The swift transition to this new paradigm has many advantages but also brings new risks and possible negative consequences for individuals or society, mainly concerning data protection, digital rights, and ethical standards. It matters how policy issues are addressed to resolve ethical and legal conflicts, and how much transparency is required in AI and data analytic solutions. The software development depends on human choices and these immensely affect the AI outcomes and their integration into operational processes.

There is an urgent need for a responsible deployment of AI regarding human rights and values which must be ensured through a carefully designed framework. To this end, the AI principles of the OECD have set the general guidelines, which were accepted by all its members in 2019, and describe the way that AI should be developed and implemented in national policies to mitigate its drawbacks. Similarly, the EU Artificial Intelligence Act has been proposed to create a balanced and proportionate horizontal regulation through a risk-based approach to safeguard a seamless digital transition that will promote economic growth while respecting firms and citizens' rights and values. The PromethEUs network members have already national strategies and policies in effect, which adhere to the OECD principles, but they need to continuously adjust to the meteoric rise of AI applications and ensure their smooth implementation into society functions. **The potential economic growth of AI utilisation is difficult to assess due to its widespread implications throughout all economic sectors, but it is certain and considerable beyond doubt, as well as its weaknesses and the threats it poses.** The current low usage of AI presents an opportunity for a mindful and responsible implementation of such technologies. **Research in all four countries has steadily increased over the last two decades, while there has been a significant surge of investments in AI startups in the last three years.** In addition, the EU funding for digital technologies in the 2021-2027 Multiannual Financial Framework, especially through the national Recovery and Resilience Plans, is an outstanding complement to national funding of the digital transition of each country. **All these constitute a fertile environment for the upscaling of AI deployment so as to take advantage of its strengths and sieze its opportunities while alleviating the possible negative consequences.**