

# Balancing Innovation and Protection under the AI Act

Boccardelli Paolo

---

RELATORE

Pauciolo Domenico

---

CORRELATORE

Francesco Pisani - 775251

---

CANDIDATO

# Table of Content

<b>INTRODUCTION</b> .....	<b>p. 4</b>
 CHAPTER 1	
<b>Institutions as Strategic Compass of the Firm</b> .....	<b>p. 6</b>
1.1 Institutions as architects.....	p. 6
1.2 Institution-based View Theory.....	p. 9
1.3 Role of Supranational Institutions.....	p. 12
1.4 Institutional Theory and Technological Innovation.....	p. 16
1.4.1 The “TITAN” Model.....	p. 17
1.5 Comparing Tech Regulatory Models: Europe, the United States, and China.....	p. 21
1.5.1 The US Model: “Market-Driven Innovation and a Decentralised Approach”.....	p. 21
1.5.2 The Chinese Model: “State-Guided but Enterprise-Led”.....	p. 23
1.5.3 Europe: The “Rule-Based” Paradigm in Technological Governance.....	p. 25
1.6 Flexible vs Rigid Regulation.....	p. 26
1.6.1 Regulatory Rigidity.....	p. 27
 CHAPTER 2	
<b>The Raise of AI</b> .....	<b>p. 30</b>
2.1 Definitions of AI.....	p. 30
2.2 AI and Its Impact on Business.....	p. 34
2.3 Generative AI: The New Frontier.....	p. 37
2.4 GenAI as a Driver of Business Disruption.....	p. 39
2.5 Beyond Business: The Socio-Economic Dimensions of AI.....	p. 44
 CHAPTER 3	
<b>The European AI Act: Between Regulatory Ambition and Technological Realities</b> .....	<b>p. 46</b>
3.1 From Normative Ideal to Final Text: Genesis and Structure of the AI Act.....	p. 46
3.2 Innovation Dynamics under the AI Act.....	p. 51
3.3 Criticism of the EU Regulation: Ambiguities, Costs and Risks for Businesses.....	p. 54
3.4 Philosophical and Ethical Critiques of the AI Act.....	p. 58
3.5 Transparency, Accountability, Sustainability: Principles for a European AI.....	p. 60

## CHAPTER 4

<b>Voices from the Ecosystem: Perception and Impacts of the AI Act</b> .....	<b>p. 64</b>
4.1 Research Question.....	p. 64
4.2 Research Design and Methodology.....	p. 64
4.2.1 Sample Selection.....	p. 64
4.2.2 Data Collection and Analysis.....	p. 65
4.3 Results.....	p. 67
4.3.1 Regulatory Design and Implementation.....	p. 68
4.3.2 EU Market and Investment.....	p. 69
4.3.3 Ethics.....	p. 70
4.3.4 Benchmarking the US Model.....	p. 71
4.4 Considerations.....	p. 72

## CHAPTER 5

<b>Graduated Compliance Model</b> .....	<b>p. 74</b>
5.1 The Costs of Regulation.....	p. 74
5.2 <i>Graduated Compliance Model</i> .....	p. 78
5.2.1 <i>Early Phase and Maturity Phase</i> .....	p. 80
5.2.2 Regulation and Investment: Risk Realignment.....	p. 81
5.2.3 The Third Way.....	p. 82
5.3 Ethics as a competitive advantage?.....	p. 83

<b>CONCLUSION</b> .....	<b>p. 85</b>
-------------------------	--------------

<b>BIBLIOGRAPHY</b> .....	<b>p. 87</b>
---------------------------	--------------

# INTRODUCTION

The current global macroeconomic and geopolitical scenario is witnessing a transformation comparable in depth to the industrial revolutions of the past, but characterised by unprecedented speed of diffusion and technological complexity: the advent of Artificial Intelligence (AI). This technology is emerging as a force capable of redefining governance structures, competitive logic and the very foundations of the social fabric. In this context, this thesis aims to navigate the complex architecture of regulation, with a particular analytical and critical focus on the European Artificial Intelligence Act; the first global attempt to provide a legal framework for the development and use of intelligent systems.

The research question aims to investigate the search for a crucial balance for the future of the continent: the possibility of reconciling the protection of fundamental rights and the safety of citizens with the equally vital need to preserve and stimulate the innovative capacity and competitiveness of the European economic ecosystem. To answer this question, the thesis starts from a solid theoretical premise rooted in neo-institutionalist literature, arguing that economic organisations do not operate in contexts that are neutral or indifferent to their strategic choices, but are immersed in environments that are strongly structured by external elements.

Institutions, understood according to Douglass North's classic definition as the "rules of the game", influence the perception of opportunities and the scope for legitimate action. Through the lens of the *Institution-based View* (IBV), the analysis highlights how institutions today must be considered much more than mere conditions; they directly determine the allocation of resources and the growth trajectories of companies, constituting, together with the *Resource-based View* and *Industry-based View*, a "tripod" of theoretical approaches to strategic management. Furthermore, in an era characterised by multi-level governance, the role of supranational institutions, capable of influencing the conduct of companies beyond their own borders of sovereignty, emerges strongly. This is where the phenomenon of the "Brussels Effect" comes in: the European Union's unilateral ability to export its regulatory standards on a global scale, making its internal market a powerful *de facto* regulatory tool for international actors.

The transition from the institutional framework to the specificity of the technological object requires deep reflection on the nature of AI, whose definition remains subject to ongoing academic and legislative debate. Although the term often evokes visions rooted in science fiction, today's applications are characterised by so-called "narrow AI", which is extremely effective in specific tasks but lacks true cognitive ability. However, the rise of Generative AI has marked a point of no return, introducing content creation capabilities that challenge the centrality of human agency in creative

processes. This evolution represents the modern embodiment of Schumpeter’s concept of “creative destruction”, a process through which innovation demolishes old economic structures and replaces them with entirely new value creation paradigms. The impact of this technology on the business world invites, or rather compels, organisations to rethink the very essence of strategic rationality, transforming data into fundamental assets for competitiveness.

The centrality of the document is evident from the complexity of the EU AI Act, analysing its origins and structure based on risk categorisation. The European approach is distinguished by a “rule-based” philosophy, which subordinates innovation to compliance with ethical and legal principles, aiming to create an “ecosystem of trust”. However, this ambition is not without its critics, often linked to regulatory rigidity that risks not keeping pace with the rapid evolution of intelligent systems, creating potential obstacles to productivity and speed to market.

To demonstrate the theoretical reflection on reality, this work devotes a chapter to listening to voices from the innovation ecosystem. Through a qualitative survey inspired by the “Gioia” methodology, the research highlights the perceptions of the impact of the new legislation by investors, legal advisors and institutional representatives of Big Tech companies. In this regard, benchmarking with the US model, which has historically been more flexible and market-oriented, provides an essential point of comparison for assessing the economic “sustainability” of the Brussels approach.

Ultimately, the evidence gathered leads to the formulation of the *Graduated Compliance Model*, a framework that aims to suggest an alternative path. In particular, the idea is to promote a “third way” based on the principle of *co-regulation* between authorities and ecosystem players and a different compliance approach.

# CHAPTER 1

## INSTITUTIONS AS THE STRATEGIC COMPASS OF THE FIRM

### 1.1 Institutions as architects

In every complex economic system, companies operate within contexts that are never completely neutral or indifferent to their choices. In addition to market or resource variables, there are a series of external and persistent elements that contribute to defining the ways in which organisations formulate, articulate and implement their strategies. Elements such as regulatory, cultural or cognitive act over time as structuring factors, influencing the perception of available opportunities and the legitimate margins for action that regulate the interactions between economic, political and social actors.

In this sense, the institutional environment represents a dimension that is not immediately visible in the decision-making process, yet it still exerts a significant influence on how the process unfolds. The environment – including regulatory bodies, state entities, governmental agencies – influences a firm's strategic decision by creating and defining entry and exit barriers, market rules and regulations (Felsmann 2016, 1).

One of the central issues is the very definition of institution. As the most commonly used definition of Douglass North clarifies: “Institutions are the rules of the game in a society or, more formally, are the humanly devised constraints that shape human interaction” (North 1990, 3). They are what regulate the structure of exchanges, be they political, social or economic. In the theoretical framework proposed by North, it is fundamental to distinguish between institutions and organisations, two concepts that are often overlapping but analytically distinct.

Institutions constitute “a stable (but not necessarily efficient) structure for human interaction”, providing a set of rules – both formal and informal – that define the behaviour expected within a social and economic context (North 1990). Organisations, on the other hand, represent the subjects – political, economic, social or educational – that operate within these rules, with the aim of achieving specific results. Using North's well-known metaphor from the world of sport, institutions are the rules of the game, while organisations are the players: the former establish how the game must be played, the latter try to win it, using their skills and strategies within the defined parameters.

However, the relationship between institutions and organisations is not unidirectional. Although institutions profoundly influence organisational behaviour, determining the conditions under which companies can be created, operate and compete, organisations also contribute to shaping

the evolution of the institutional framework over time. As Nye observes, “institutions influence the economy and are themselves shaped by the behaviour of economic actors” (Nye 2004, 76). This results in a dynamic of reciprocal interaction: organisations are formed and transformed in response to the institutional context, but at the same time, through their practices, strategies and interactions, they can generate pressures that induce institutional change. From this perspective, the strategic behaviour of companies cannot be fully understood without considering the institutional context in which they are immersed, nor can the evolution of institutions be analysed without considering the actions of organisations.

Institutions have the capacity to influence human actions, albeit not always in an efficient manner (Boccardelli & Brunetta 2014, 16). They encompass a wide array of constraints, encompassing both formal and informal dimensions.

Informal constraints, also known as unwritten codes, function as an interpretative framework for formal rules, pervading everyday life. The human condition has always been subject to constraints imposed by individuals on their own relationships, with informal rules representing the earliest forms of institutionalisation. These constraints emerge in circumstances where information is limited, and formal constraints are absent, indeed informal rules serve a complementary function, reinforcing the effectiveness and applicability of formal rules, particularly in contexts where the latter are incomplete or challenging to implement (Boccardelli & Brunetta 2014, 16). Relationships are characterised by a set of behavioural norms and beliefs that serve as guidelines for interaction and conduct. As this information is socially transmitted, it becomes part of traditions and takes on different meanings depending on the social context (North 1990).

On the other hand, formal constraints are comprised of the provisions that define the proscriptions, as well as, in certain cases, the specific conditions under which activities can be legitimately carried out. Consequently, they constitute the regulatory framework that delimits individual and collective action, comprising explicit, codified and documented rules. As social and economic complexity has increased over time, there has been a corresponding necessity to formalise these constraints with greater rigour, to facilitate the systematic regulation of increasingly complex relationships. This progression has given rise to the emergence of institutions such as legal systems, organisational hierarchies, contracts and other formal coordination mechanisms.

In such context, companies operate in environments that are strongly structured by shared social expectations and consolidated regulatory standards and tend to adopt similar behaviours in an attempt to obtain legitimacy (Meyer & Rowan 1977). This tendency towards strategic and organisational convergence has been conceptualised within the framework of neo-institutionalist theory through the notion of *isomorphism*. The objective of neo-institutionalist analysis is to examine

the factors that institutions exert on organisations, with a view to facilitating adaptation to the prevailing criteria of rationality. The analysis seeks to understand how *isomorphism* processes actually develop.

DiMaggio and Powell were the first to introduce the concept of “organisational field”, defined as a group of organisations and subjects that influence each other (Boccardelli & Brunetta 2014, 17). This concept encompasses a group of organisations and subjects that exert a significant influence on one another, with all the organisations constituting a recognised sector of institutional life playing a leading role. This group includes suppliers, buyers, producers, regulators, and organisations that produce similar products or services and are defined through empirical analysis rather than a priori assumptions (DiMaggio & Powell 1983). It is noteworthy that neo-institutionalists, in contrast to classical institutionalists, place significant emphasis on this concept.

The defining characteristics of the “organisational field” encompass both the object and the subject of the pressures that are exerted, with the process of *isomorphism* emerging as a consequence of reciprocal pressures. This tendency, in response to institutional pressures, leads to a gradual resemblance between organisations: this process of homogenisation can be attributed to the competitive pursuit of legitimacy among organisations. The adoption of “institutionalised” behaviours by organisations leads to their becoming “isomorphic”, a phenomenon that confers certain advantages related to their survival. Institutionalisation signifies that behaviour models accrue value and are elevated to the status of legitimate rules within the organisational field, thereby becoming archetypal and distinguishing themselves from other behaviour models. Consequently, as the “organisational field” evolves, a process of homogenisation occurs within the organisations, resulting in an increasing similarity between them (Boccardelli & Brunetta 2014, 17).

DiMaggio and Powell introduce three mechanisms of diffusion: coercive, normative and mimetic (DiMaggio & Powell 1983). The first of these occurs when external forces, whether formal or informal, compel organisations to adopt organisational elements, leading to homogenisation and the avoidance of potential sanctions. The second mechanism, termed normative, occurs when organisations are motivated to respect social obligations, basing their strategic decisions on common cognitive frameworks, and conforming to the established rules of behaviour. Finally, mimetic diffusion is derived from conditions of uncertainty, causing imitation to other organisations.

The context in which organisations operate is made up of established and shared elements, which represent a frame of reference for defining this whole as a template. Changes to the latter give rise to organisational changes to which companies in the field conform (Camuffo & Cappellari 1996). It is precisely the neo-institutionalist perspective that argues that organisational survival is determined by the degree of alignment with the regulatory environment; in an attempt to achieve stability,

companies conform to the rules of the institutional context, with the aim of improving their survival and obtaining greater legitimacy and social status (Boccardelli & Brunetta 2014, 23).

As previously stated, all organisations in the organisational field are both the object and the subject of the pressures that are produced in the field, and *isomorphism* is the result of these reciprocal pressures. The interconnections between organisations can be seen as a form of institutional embeddedness (Boccardelli & Brunetta 2014, 23).

Academic literature uses different words for the institutional approach. This sometimes causes confusion in common understanding. This situation can be traced back to the symbiotic roots of the theory in sociological, economic, legal and business fields. Sociologists tend to use the label “institutional theory”, while economists have obtained the label “institutional economics”. Mike W. Peng suggest a new term, “institutional based view” (IBV), for the management-oriented research field to differentiate management-oriented research from existing work in economics and sociology.

## 1.2 *Institution-based View Theory*

In recent decades, there has been growing recognition among strategic management researchers that institutions should be considered more than just background conditions; “institutions directly determine what arrows a firm has in its quiver as it struggles to formulate and implement strategy” (Ingram & Silverman 2002, 20). This realisation forms part of a broad intellectual movement centred on new institutionalism across the social sciences. Notably, this view has been argued to be one of the three leading perspectives in strategic management with *industry-based* and *resource-based* views.

The rise of the institutional vision in strategy is being driven by two sets of forces: external and internal (Peng et. 2009, 64). The former is represented by the broader institutionalist movement currently sweeping through the social sciences, which appears under various labels across different strands of the literature. Regardless of disciplinary roots, there has been a growing consensus around a fundamental proposition: “institution matter” (Peng & Pleggenkuhle-Miles 2009, 3).

In addition to external forces that spill over into the field of strategy, IBV has also grown in response to internal dynamics – specifically, the long-standing criticisms directed to *industry-based* and *resource-based* views for their lack of attention to contextual factors. Considering the first of Porter’s five forces: interfirm rivalry and its prescription for a cost leadership strategy. *The industry-based view* seldom interrogates the institutional foundations underpinning such rivalry. In reality, formal government policies, as well as informal elements such as media discourse and consumer

sentiments regarding acceptable corporate behaviour, play a significant role in shaping the nature and intensity of competition (Peng et. 2009, 65). Under certain institutional conditions, a cost leadership strategy may be perceived as unethical or even deemed illegal. What constitutes legitimate competitive behaviour is not universally defined but is shaped by the formal and informal rules embedded in specific institutional environments. Likewise, *the resource-based view* has been criticized for its “little effort to establish appropriate contexts” (Richard & Butler 2001, 32). Valuable, rare, and hard-to-imitate resources and capabilities in one context may become nonvaluable, plentiful, and easy to imitate in other contexts; “the value of a firm’s resources must be understood in the specific market context within which a firm is operating...” (Jay 2001, 52).

IBV posits that strategic choices are determined not only by industry-specific factors or by the available resources and capabilities but also constrained by the institutional framework; this theoretical framework provides a third perspective on strategic analysis, contributing to a comprehensive “tripod” of theoretical approaches to strategic management (Boccardelli & Brunetta 2014, 137). Building on this foundational role, the theory advances by treating institutions as independent variables and highlighting the dynamic interaction between institutions and organizations.

In this context, Peng argues that a full understanding of strategic decision-making – viewed as a reflection of both formal and informal institutional constraints – requires the articulation of two core propositions that lie at the heart of the institution-based perspective.

One of the primary objectives of institutions is to reduce the degree of uncertainty for various actors by influencing regulatory norms of behaviour, while at the same time defining the question of legitimacy (Peng et. 2009, 66). From a rational choice perspective, it is proposed, as first *proposition*: “managers and firms rationally pursue their interests and make strategic choices within the formal and informal constraints in a given institutional framework” (Peng et. 2009, 67). Here, rationality is understood as bounded rather than perfect. Managers and firms are seen as economic actors who aim to act rationally, but do so within the limits of available information, cognitive capacity, and time. This reflects the idea that actors are intendedly rational, but their decision-making is limited by institutional constraints, a concept that Williamson describes as a “simultaneous reference to both intended and limited rationality” (Williamson 1985, 45).

The second *proposition* is based by the observation posed into the constraints emerged by formal and informal institutions: “while formal and informal institutions combine to govern firm behaviour, in situations where formal constraints are unclear or fail, informal constraints will play a larger role in reducing uncertainty providing guidance and conferring legitimacy and rewards to managers and firms” (Peng et. 2009, 68).

These two propositions are not merely abstract theoretical claims – they are empirically observable across a variety of institutional contexts. A widely studied example is the case of China, where rapid economic growth has occurred despite historically underdeveloped formal institutions, such as weak judicial systems and limited regulatory enforcement (North 1990). In such settings, managers and firms have responded by cultivating interpersonal networks – commonly referred to as *guanxi* – which serve as informal mechanisms to access resources, reduce uncertainty, and facilitate transactions.

This behaviour exemplifies *Proposition 1*, as firms act rationally within the boundaries of their institutional environment, making use of the tools that are available and effective within that context. At the same time, the prevalence of *guanxi* also reinforces *Proposition 2*, highlighting how informal institutions can become dominant in governing strategic behaviour when formal structures are either absent or ineffective.

This case underscores the broader insight of the *institution-based view*: firm strategies are shaped not only by market forces or internal capabilities, but by the specific constellation of formal and informal rules that structure economic interaction.

A clear illustration of the IBV's distinct contribution emerges when considering cross-national differences within the same industry. A striking example is provided by the Japanese pharmaceutical industry. While Japanese firms are globally recognized for their innovation in sectors such as automotive and electronics, their pharmaceutical companies have not reached comparable levels of global competitiveness. This divergence is not due to a lack of technological potential, but rather to institutional arrangements that discourage investment in innovation. In Japan, drug prices are fixed by the Ministry of Health and cannot be increased during a drug's prespecified life cycle. As production costs decline over time due to economies of scale, the highest profit margins are realized not from the newest drugs, but from the oldest. As a result, under such institutional constraints, managers behave rationally by deprioritizing R&D-intensive strategies.

This case exemplifies *Proposition 1* of the institution-based view: firms make strategic decisions based on bounded rationality within the parameters defined by their institutional environment. In contrast, Western pharmaceutical firms operate in systems that strongly incentivize breakthrough innovations, reinforcing their commitment to R&D.

A further topic discussed from the three perspectives that form the “tripod” is the analysis of what determines the success or failure of companies around the globe. According to *the industry-based view*, a company's performance is mainly determined by its degree of competitiveness. *The resource-based view*, on the other hand, emphasises the capabilities of the company itself. As Peng argues, from the moment a company begins to undertake venture activities abroad, it is advisable for

it to analyse the “rules of the game” of foreign markets – both formal and informal rules (Peng et. 2009). This can be considered one of the main factors for success in doing business internationally.

Such an approach underscores the broader relevance of IBV, which offers new perspectives on fundamental issues in management strategy. These perspectives should not necessarily be seen as contradictory to those introduced by *the resource-based* or *industry-based* views, but rather as complementary. As Peng suggests, its full explanatory power emerges when considered as one of the three pillars of the so-called “strategy tripod”, alongside industry structure and firm-specific capabilities. By emphasizing the importance of formal and informal institutional frameworks, the IBV provides critical insight into how organizations behave within diverse regulatory, normative, and cultural environments.

This becomes increasingly relevant in today’s globalized and multilayered governance landscape, where supranational institutions, such as the European Union, are shaping the strategic context through emerging regulatory frameworks. Incorporating this institutional dimension is essential to understand how firms adapt not only to markets and internal resources, but also to evolving external rules and constraints that transcend national boundaries.

### 1.3 Role of Supranational Institutions

In the contemporary economic landscape, supranational institutions are acquiring an increasingly important role in influencing the operating conditions within which companies formulate their strategies. While national institutions operate within their own legal boundaries, supranational institutions differ in their ability to transcend national sovereignty defining them as “supranational rules, norms, and beliefs, which structure interactions among individuals and organizations” (Hartmann et. 2022, 1285).

The shift from a national regulatory sphere to one beyond established legal boundaries entails a profound transformation of the concept of governance, which from a system based on nation-states has progressively evolved into a multilevel ecosystem with multiple actors. This recognition does not neglect the importance given to national institutions, but rather, recognises that in a globally connected world, they co-exist in a global institutional ecosystem where supranational structures govern and moderate relations between them and multinational (MNCs) and national corporations (Hartmann et. 2022, 1285).

Supranational institutions play a central role in shaping companies’ strategic decision-making processes. Traditional institutional literature has mainly focused on the analysis of national

institutions, but today it is necessary to highlight the emergence of an institutional layer that overlaps and, in some cases, prevails over local institutional dynamics.

In this context, it is important to understand the role of these institutions by analysing their conceptual properties, as well as the ways in which they influence and interact with the surrounding actors. In particular, there are six properties that characterise supranational institutions: supraterritoriality, coexistence, contextuality, selectivity, consensuality and co-evolution. These not only enrich our understanding of the institutional framework but also highlight how companies must navigate highly complex environments (Hartmann et. 2022, 1299).

Firstly, the concept of supraterritoriality is linked to the ability to influence even outside the national environment of origin of the company, without direct political legitimacy in the country. This implies the ability of regulations to have a cross-border impact beyond traditional jurisdictional boundaries. The ways in which this influence is exercised differ depending on the nature of the institution.

In the case of formal institutions, it is determined by the coercive application of agreements and contracts with extraterritorial enforcement (Scherer & Palazzo 2011). On the other hand, with regard to informal institutions, this supraterritoriality acts as a form of social and reputational pressure, conveyed by actors such as global media, NGOs and transnational advocacy networks.

This supraterritoriality is clearly represented by the “Brussels Effect”, which describes the European Union’s unilateral regulatory capacity to export its regulatory standards beyond its borders, radically influencing the conduct of global economic actors (Bradford 2020, 14). This phenomenon shows how the effectiveness of the EU does not depend directly on direct political legitimacy in certain national contexts, as it has the ability to shape the global economic environment without resorting to international institutions. This is due to its large domestic market, which represents both an opportunity and a risk for global companies: market size is a prerequisite for exercising unilateral regulatory authority, making large economies sources of global standards (Bradford 2020, 25). For this reason, companies are influenced to voluntarily comply with these standards in order to consolidate their access to the market.

Furthermore, the European Commission enjoys robust regulatory capacity, fuelled by an institutional architecture that has translated the size of the European market into tangible regulatory influence. This capacity is the result of a highly professionalised Commission and the presence of strong regulatory resources, born out of a desire to build a regulatory machine capable of affirming the integration of member countries with a single market.

A central element of the “Brussels Effect” is represented by the non-divisibility to the practice of standardising production or business practices across different jurisdictions (Bradford 2020, 54).

In many cases, the adoption of a European regulatory standard leads to an adaptation of practices at a global level, given that differentiating in multiple markets would be costly. In fact, the General Data Protection Regulation (GDPR) requires compliance not only from actors residing in the EU, but also from all entities that process EU citizens' data. The principle of “privacy by design”, as set out in the GDPR, obliges companies to incorporate these requirements into the design phase of their products and services, ensuring that these standards, once adopted, may become *de facto* benchmarks for companies operating in multiple jurisdictions.

The second property, coexistence, demonstrates how supranational institutions do not operate in a vacuum, but rather intertwine with national institutions, generating a complex regulatory environment. This scenario requires companies to simultaneously manage constraints and incentives from different institutional levels, which are sometimes in conflict with each other, leading to ambiguity and regulatory uncertainty. Given this “institutional triality” – supranational, national and subnational – firms need to adopt institutional relationships that go beyond mere adaptability to a single regulatory framework and instead require a dynamic capacity for multi-level integration (Hartmann et. 2022, 1293). When a company operates in markets subject to European Union, national and sub-state authorities, it has a duty to develop consistent lobbying practices.

Thirdly, contextuality is important in demonstrating that the effects of supranational institutions on companies are not entirely universal but depend on specific national circumstances. Many of these formal institutions do not have true sovereign status or direct executive power, therefore they require national actors for effective implementation (Hartmann et. 2022, 1293). This means that “top-down” decisions by supranational institutions may not be applicable to some countries. Some national governments may protect their own companies by hindering the application of potentially conflicting rules or values. However, from a “bottom-up” perspective, such contextuality on the part of inefficient national institutions is less relevant, given that a company voluntarily chooses to adhere to such rules.

The fourth property, selectivity, highlights how supranational institutions can influence companies unequally within a national context, causing advantages or disadvantages for some companies (Hartmann et. 2022, 1293). It manifests itself when supranational regulations – established by bilateral or multilateral treaties – introduce protection mechanisms. One example is investment protection agreements, which feature ISDS (Investor-State Dispute Settlement) clauses that offer foreign companies guarantees of protection for their investments vis-à-vis other private economic actors. On the other hand, however, selectivity can also be considered a competitive disadvantage; when the most “visible” companies are subject to a high degree of media exposure and strong pressure

from civil society, causing to be more vulnerable to supranational regulatory action. For this reason, many companies are forced to voluntarily adopt informal standards to preserve their global reputation.

Consensuality, which represents the fifth property, is one of the most important aspects for the application of supranational institutions, given that, by definition, they lack national sovereignty. This property, in its highest degree, can manifest itself through agreement between national governments on the establishment of a supranational institution ratified and applied at the national level, which has a direct influence on MNCs and their transactions. However, in informal institutions consensus can emerge among various actors, often from multiple countries, who are able to exert pressure on companies. A clear example of this was the proliferation of ESG (Environmental, Social, and Governance) principles in the financial sector, which was driven by consensual adoption by actors in various countries rather than by national financial regulators (Contreras et. 2019).

Finally, as a last property, co-evolution describes the dynamic and reciprocal interaction in which businesses, supranational and national institutions influence each other endogenously. Companies are not simply constrained by institutions in their operations, whether national or international, but can also contribute to their development (Kostova & Zaheer 1999). They can take on the role of “institutional entrepreneurs”, directly shaping supranational institutions through a “bottom-up, one-tiered” approach, such as the International Financial Reporting Standards (IFRS), developed by a business-led body (Hartmann et. 2022, 1294). Furthermore, they can also contribute to the emergence of supranational institutions from local institutions, with a “bottom-up, two-tiered” approach, such as the integration of national standards by the European Committee for Standardisation (CEN) (Hartmann et. 2022, 1294). Moreover, companies can also organise themselves collectively, under the guidance of supranational actors, to establish supranational institutions in the form of global public policy networks. One example is the use of non-market strategies by American MNCs to strengthen intellectual property protection in the US trade agenda, influencing the content of the WTO’s TRIPs (Trade-Related Aspects of Intellectual Property Rights) agreement (Lawton & McGuire 2001, 228).

In conclusion, the six properties just analysed paint an analytical picture whose objective is to better understand the growing role of supranational institutions in defining corporate strategies. In this scenario, the European Union, through its regulatory capacity, with a preference for strict standards and a willingness to regulate sectors, offers an exemplary case of the “Brussels Effect”, demonstrating how supranational institutions can translate into global regulatory influence.

Furthermore, it is also worth considering the effects that such influences may have on corporate innovation processes. While a certain degree of regulation may lead to forms of responsible innovation, excessive regulation could represent an obstacle to innovation itself. In fact,

understanding the role of institutions in modulating the relationship between regulatory constraints and innovation is an essential step.

#### 1.4 Institutional Theory and Technological Innovation

Technological innovation has long been recognised as a real driver of economic growth in society, with many scholars analysing the factors that influence it positively and negatively. In this context, among the many players involved, the institutional environment has emerged as a priority area of research (Du 2018).

As seen during the writing of the thesis, the institutional environment has a significant impact on strategic behaviour and business performance; its dynamics can profoundly influence business strategies.

The importance of institutional theory in innovation studies, although underestimated, is immense. It provides a theoretical lens through which it is possible to examine the effects of the social context on innovation, going beyond the innovation process itself and examining the institutional environment in which it occurs. Thus, it is able to provide crucial insights into the factors that motivate innovation in specific contexts. In the contemporary context, characterised by complex challenges such as AI regulation, institutional theory offers a framework capable of aiding policy formulation by ensuring that innovation serves the collective good.

The broadening of the field of interest in the dissemination of tools and services necessary to create new technology represents a move towards the more comprehensive idea of an “innovation community”, recognising how the institutional environment influences the development and propagation of technology (Lynn et. 1996).

In particular, the use of the concept of innovation communities recognises that the process of commercialising new technology is the result of interaction between actors within an organisational structure, which is divided into a superstructure, consisting of organisations with coordination functions, and a substructure, consisting of entities responsible for producing the technology. The links between these two actors play a fundamental role in supporting the innovative process of societies. Here, the environment is seen as a co-evolutionary structure that goes beyond technology itself, showing a dynamic relationship between innovation and the organisational structure.

The fundamental characteristic that must be highlighted in this concept is the central role played by technology within societies, identifying the subjects involved in the commercialisation of technological innovation.

In this regard, the distinction between substructure and superstructure is crucial. Organisations belonging to the former are the protagonists in the production of innovation and its technological complementarities. Superstructure organisations, on the other hand, provide collective goods to their members, often specialising in coordinating information flows and the activities of substructure organisations.

A crucial variable is the relative power attributed to organisations with a distribution of roles that varies between innovation communities.

For instance, In Japan, in the 1950s and 1960s, the process of adopting the basic oxygen furnace (BOF) for steel production saw strong involvement from national institutions (Lynn et. 1996, 95). The government with metallurgical companies played a key role in sharing technical data, studying the process and funding research. Moreover, industry associations – Japan Iron and Steel Federation – acted as central nodes in the network, facilitating a high degree of communication between various actors such as universities and authorities. This superstructural approach allowed for the efficient and rapid spread of BOF technology.

On the other hand, in the United States, during the same period, the model was the opposite. The steelworks – U.S. Steel, Bethlehem Steel, and Kaiser Steel – acted as experimental laboratories, where each company developed its own BOF plant (Lynn et. 1996, 95). The lack of coordination by institutions led to a sort of duplication of efforts with parallel projects characterised by individual patents. This model made the adoption of the technology slower and more fragmented than in Japan.

The coordination function is a vital component in the process of commercialising innovation, often carried out through organisations or relationships that resemble a complex field of many actors rather than a bilateral one.

#### 1.4.1 The “TITAN” Model

The relationship between technology and institutions is a central issue in public policy. Emerging technologies are generating complex challenges that require new approaches for effective regulation. In this regard, the “Theory of Institutional Technological Actors and Norms (TITAN)” emerges as a useful regulatory framework for the regulation of modern technology.

“TITAN” focuses on the moral and legal rights and responsibilities of relevant actors, with the aim of informing the development of regulations that are fit for purpose, confirm rights and are fair to all stakeholders (Smith & Miller 2025). Four key categories of groups in society are considered:

(1) Technology producers; (2) Technology users; (3) Government regulators of technology; (4) Normative policy shapers.

The theory is based on two interconnected theoretical concepts. The first is represented by “normative institutional theory”. According to this perspective, technology is intrinsically embedded within institutions, which play a fundamental role in the dissemination of collective goods for the well-being of society. Normative theory has the task of identifying these goods by offering normative guidance.

Today, the interaction between institutions and society is mediated by technology, which is not merely a tool at the service of institutions; it is an integral part of them, influencing the structure of society. One example is how the internet determines private and public communication among the many actors in modern society.

In light of this, it is precisely the regulatory instrument that must ensure the fair integration of technology into society. An adequacy criterion is needed that provides technology with the capacity to promote the production and maintenance of collective goods in an ethical and effective manner. In order to comply with this criterion, it is important that the legal legitimacy of technologies reflects ethical principles, individual autonomy and the protection of privacy, without compromising technological innovation itself.

The second theoretical concept is related to collective responsibility. Research, development, regulation and use of technology involve many actors, who have a duty to ensure collective responsibility. There are two conceptually separable but interconnected aspects: (1) the collective responsibility to ensure that technology is produced and used in an effective, lawful and ethically responsible manner; and (2) the collective responsibility to ensure that an important means to this end, namely regulation itself, is effective, lawful and ethically sustainable (Smith & Miller 2025).

Although members of the government, such as legislators and regulators, have greater responsibility, all groups within the four categories mentioned above have responsibilities in relation to regulation, even if only in terms of assessing compliance with effectiveness. The actors belonging to the other three categories – (1), (2) and (4) – have the task of assuming different responsibilities in relation to other instruments for these purposes. In fact, scientists and technology companies that develop new biotechnologies for public health must necessarily comply with regulations, but they should also provide useful input to the government to contribute to the formulation of regulatory policy.

In order to formulate appropriate technological regulations, it is important to consider the risks and benefits for each of the four categories: producers, users, regulators, and policy shapers.

Considering technology companies – which provide significant collective goods – their devices greatly increase business efficiency, making personal activities more convenient through greater communication opportunities. Moreover, with a rapid increase in access to information, there is a profound impact on business productivity and the standard of living of individual workers. Despite this, it is also important to note the possibility of risks and harm, such as the loss of privacy rights with opportunities for outsiders to access sensitive personal and financial data. Furthermore, with the advent of new technologies, multinationals have taken advantage of them to shift their value chains to other countries, resulting in tax avoidance and lower revenues for governments, demonstrating how some economic actors today have greater bargaining power than others.

In light of this, it is legitimate to ask what the purpose of a new technology is, or rather what actual benefits it brings to society. Considering normative institutional theory, such benefit is conceptualised as a collective good. These coincide with public goods in the definition given by economic theory, as goods characterised by non-rivalry in consumption and non-excludability in access.

However, considering “TITAN” theory, not all collective goods fall within this classic economic definition. What really distinguishes collective goods conceptually is the way in which they are produced and preserved; they originate from a co-created process, the result of which is cooperation between a plurality of economic, social and institutional actors (Smith & Miller 2025). Technology-supported goods are collective goods, insofar as their functionality depends essentially on the various actors involved during the phases of the technological cycle – from design to implementation – necessarily contributing to the production of social value.

Given that technology provides important collective goods, it takes on moral significance, giving rise to moral rights and responsibilities, some of which are necessarily institutional responsibilities.

The advent of artificial intelligence represents one of the most complicated challenges for contemporary institutions. Its effects are reflected in crucial sectors of society, such as work, education and medicine. It can be the source of significant benefits but also a source of risks. For this reason, it is imperative that institutions – regulatory – play an important mediating role, so that AI strengthens collective goods, and not the opposite.

Assuming collective responsibility is a necessary condition, ethicists and jurists have the right to offer regulatory advice to national and EU policy makers, and on the other hand, they have the duty to promote a culture of responsibility in the scientific and business sectors. This contribution is not only legitimate from a regulatory point of view, but also and above all from an ethical point of view.

This scenario indicates the need to adopt a collaborative regulatory approach, with shared responsibility among all stakeholders.

Technology producers, in addition to being responsible for the design of the technology, they are equally responsible for sharing crucial knowledge to ensure appropriate regulation.

At the same time, technology users – ordinary citizens and/or consumers – play an active role in the market, influencing innovation priorities and electing governments responsible for regulating new technology.

Conversely, government regulators of technology have the power to develop regulatory policies that guide technological development. In EU, this task is carried out by the AI Office, a regulatory body established within the European Commission, whose objective is to supervise, implement and coordinate the application of the AI Act – the world's first regulatory framework dedicated to artificial intelligence.

Finally, normative policy shapers must provide relevant ethical concepts, including moral rights and principles, for institutional action aimed to structure a coherent regulatory framework. A significant picture is the European ethics committees in drafting the guiding principles for a trustworthy artificial intelligence, published by the European Commission through the High-Level Expert Group on Artificial Intelligence (European Commission 2019).

Collective goods, in this sense, are produced jointly, and institutions include not only those who participate in the production or maintenance of collective goods, but also those who benefit from or consume them, such as citizens who vote for democratically elected governments or consumers in markets (Smith & Miller 2025).

In conclusion, in order to address the challenges posed by artificial intelligence, effective regulation of technology is vitally important for modern society. By analysing the innovation community framework using the Theory of Institutional Technological Actors and Norms (TITAN), it has been possible to offer useful regulatory approaches for understanding the role of institutions.

Today more than ever, it is necessary to illustrate how organisations and inter-organisational relationships influence the commercialisation of technology, and how technology influences the structure and functioning of society. It is clear that the vitality and efficiency of innovation depend on the institutional environment, where a favourable environment facilitates the efficient allocation of resources, fostering both the aspiration toward innovation and its practical implementation.

## 1.5 Comparing Tech Regulatory Models: Europe, the United States, and China

New technologies, particularly artificial intelligence, are proving to be a transformative force in the 21st century, redefining industries, governance structures and social interactions in unprecedented ways. However, this rapid advancement is also introducing many complex regulatory and geopolitical challenges. The vision for AI varies significantly across global regions, radically influencing the trajectory of innovation. This variation affects how different regions approach AI development, balancing technological progress with ethical and regulatory considerations.

In this regard, it is necessary to compare the institutional approaches to regulating new technology in the three main global contexts: the US, China and the European Union. The aim is to provide information on the differences, which are almost philosophical in nature, that shape the relationship between the state, business and technology in each of these jurisdictions, highlighting how different legal and political cultures give rise to distinct regulatory frameworks.

The absence of cohesive global governance on the development and use of AI is having significant consequences around the world; and without clear governance mechanisms, these circumstances could threaten democratic processes and economic stability.

Today, there is a clear challenge of geopolitical competition in AI development; and this race between these three major regional players has led to fragmented regulatory landscapes.

### 1.5.1 The US Model: “Market-Driven Innovation and a Decentralised Approach”

The United States stands out for its AI governance model, which favours market-driven innovation with minimal regulatory constraints (Kulothungan & Gupta 2025). This approach, which we can define as “light-touch” and decentralised, aims to promote entrepreneurial freedom and rapid innovation, accepting a certain margin of risk to avoid possible obstacles to technological development.

One of the most significant features of this approach is that leadership in AI development is primarily entrusted to the private sector, with a regulatory methodology characterised by self-regulation standards. It is believed that excessive regulation could hinder technological progress; commercial entities are the drivers of AI generative development, which is characterised by very rapid implementation cycles. Notable examples include large language models from technology companies such as OpenAI and Google.

In the United States, with no comprehensive federal law regulating AI, there is a fragmented regulatory framework and a sectoral approach. In this context, AI is addressed through a “mosaic of individual agency approaches and specific legislation” (Davtyan 2004, 6). Regulation is spread across various federal agencies and state initiatives, each with specific jurisdiction; the Consumer Financial Protection Bureau regulates AI in financial services to prevent harm, while the National Highway Traffic Safety Administration focuses on AI applications in autonomous vehicles to ensure safety. This model allows each federal agency to use its authority to address AI-related issues.

A key strategy of the US approach is to encourage self-regulation by individual industry players. In July 2023, the Biden-Harris Administration secured voluntary commitments from major AI companies, including Amazon, Google, Meta, and Microsoft, to improve the safety, security, and transparency of AI (Davtyan 2004, 6). This involved companies agreeing to rigorous testing, sharing security protocols, reporting vulnerabilities, and developing tools to identify AI-generated content.

In relation to risk management, this task is entrusted to the National Institute of Standards and Technology (NIST) through the “NIST Artificial Intelligence Risk Management Framework” (AI RMF). This is a completely voluntary and non-binding framework that provides a roadmap for mitigating AI risks, offering companies a tool for designing and managing reliable AI. The AI RMF provides a socio-technical approach to risk, considering the impact of AI on society and human behaviour, as well as highlighting the characteristics of reliable AI: security, transparency and fairness (Davtyan 2024).

Over the course of recent presidencies, executive orders have been used to guide policy on artificial intelligence. One of the most recent and significant is the “Executive Order 14110” from the Biden Presidency, which came into effect in October 2023. It is the most comprehensive piece of AI regulation in the United States to date, delegating responsibility for AI to over 50 existing federal agencies. These tasks are grouped into eight main policy areas: safety and security, innovation and competition, support for workers, AI biases and civil rights, consumer protection, privacy, federal use of AI, and international leadership (Chun et. 2024, 15). This Executive Order focuses on basic dual-use models and large computing clusters, requiring companies to provide details on these models, information on cybersecurity measures and security testing. It also aims to attract talent to the world of artificial intelligence, provide useful tools and environments to help startups grow, and develop transparency standards.

The US approach is characterised by expanding the capabilities of agencies rather than establishing centralised regulation. This approach is defined as “permissive”, encouraging innovation through competition and relying on an extensive network of existing laws and regulations against abusive, illegal or negligent practices (Chun et. 2024, 16).

In recent years, Congress has been actively involved in regulatory initiatives related to new technology. Although it has not introduced direct legislation on the use of AI in the private sector, it has invested heavily in R&D and expanded the federal government's capacity to manage AI development. It has helped to formulate acts such as the John S. McCain National Defense Authorization Act (which directed the Department of Defense towards AI initiatives), the AI in Government Act of 2020 (which created an AI Center of Excellence), the Advancing American AI Act (further guidelines for the use of AI by the federal government), the AI Training Act, the CHIPS and Science Act (promoting AI R&D and creating standards for trustworthy AI), and specific acts to combat human trafficking and detect AI-manipulated content (Davtyan 2024).

From the perspective of state-level initiatives, California Senate Bill 1047 (SB 1047) has been the most significant example in recent years of the regulatory debate on AI. Although it did not come into force, it pointed out the need to adopt proactive safety measures and accountability in AI development, raising awareness of the potential risks that can arise from advanced systems. It contributed significantly to the serious consideration of safety and governance issues.

### 1.5.2 The Chinese Model: “State-Guided but Enterprise-Led”

The Asian giant adopts a governance model that is described as state-guided but enterprise-led. It is a hybrid approach characterised by strategic and ideological control by the state, with close collaboration with large companies in defining technological standards. The primary objective is to balance rapid technological deployment with strong government oversight, particularly in areas such as surveillance and public safety, while aiming to become a global regulator in emerging sectors (Kulothungan & Gupta 2025).

The model features centralised regulatory control, but emphasises centralised innovation, bringing economic development to the local level. China focuses on developing AI with the fundamental values of socialism. The regulations introduced are the product of a process involving various stakeholders, including bureaucrats, academics, companies, start-ups and think tanks (Chun et. 2024, 10).

The central government relies on local experts and officials to ensure that objectives and results are aligned with the country's ideology.

The peculiarity lies in a state-driven model that balances the rapid spread of technology with strong government oversight. In fact, regulations impose content moderation and controls on misinformation, resulting in a unique balance between innovation and control.

From a regulatory point of view, China has adopted a sectoral approach, with ad hoc laws on specific cases, ranging from data privacy to generative AI recommendation algorithms.

The country was the first to introduce safety regulations for generative AI: the system imposes a certain degree of content moderation. The aim is to define the main security risks of training data and generated content, including the violation of fundamental socialist values, discrimination, violation of commercial laws, violation of the legitimate rights and interests of third parties, and failure to meet the security requirements of specific types of services (e.g. medical information or psychological counselling services) (Zhu 2024).

Through the policy adopted with “Made in China 2025”, the national industrial strategy aims to make the country a global leader in advanced technologies, strongly supporting innovation. The plan provides support for 10,000 “Little Giants”, small and medium-sized enterprises (SMEs) recognised as a key source of innovation (Chun et. 2024, 11).

In addition, the state has created a “national AI team” composed of 23 private companies that are leaders in the field of AI applications (Zhu 2024,6). These companies are responsible for developing a National Open Innovation Platform for AI (AIOIP), which provides access to data, frameworks and computing power to start-ups and SMEs. The process is a form of standardisation, as SMEs use the resources of the leaders, making them dependent and strengthening the leaders’ power to set standards.

However, China has overly burdensome registration requirements. As of March 2024, only 546 AI models had been registered, of which only 70 were large language models (LLMs), a number significantly lower than the millions of commercial and open-source LLMs available elsewhere (Chun et. 2024). This highlights how strict regulation can slow down innovation and economic growth. One of the main problems arises when “destabilising models” or “excesses” emerge that challenge the values or authority of the Chinese Communist Party, leading to a proactive regulatory response in terms of punitive measures (Chen & Liu 2023).

In this context, the state plays different roles in the AI standardisation process.

In the early stages, the state acts as a catalyst, issuing national policies and plans for standardisation (Zhu 2024, 7).

During the standard-setting process, support for businesses is evident. This includes the establishment of technical committees, access to critical research and development resources (e.g., testing spaces for autonomous vehicles for Baidu, public medical data for Tencent, city infrastructure data for Alibaba) and approval for the creation of AIOIPs (Zhu 2024, 7).

In the next phase, the state takes on the role of supervisor, setting limits on business-led standardisation work and safeguarding market order with macroeconomic countermeasures and antitrust laws (Zhu 2024, 7).

Finally, the state is responsible for the adoption and publication of national standards (Zhu 2024, 7).

The approach analysed is based on a relationship between the state and businesses that is neither purely top-down nor bottom-up. China has established a “community of practice” in which the state, businesses and other public sector stakeholders – academics and the media – develop a joint problem-solving mindset to position China at the forefront of AI technologies (Qiao-Franco & Zhu 2024).

### 1.5.3 Europe: The “Rule-Based” Paradigm in Technological Governance

Although a more detailed analysis of the European regulatory approach will be discussed in subsequent chapters, with reference to the AI Act, it is worth anticipating the general coordinates in order to obtain an excellent comparison with the other models just analysed.

The European Union adopts an approach commonly described as “Rule-Based” or “Regulation First”, focused on the protection of fundamental rights, transparency and corporate accountability (Kulothungan and Gupta 2025). It is a regulatory philosophy that, rather than focusing on regulatory agility or strategic state intervention, aims to formalise a coherent legal framework capable of ensuring a balance between innovation and the protection of individuals. The strategy is centralised and horizontal, establishing broad standards applicable to different industries and ensuring consistency among Member States.

Unlike the US model, which opts for a certain degree of flexibility and self-regulation, and the Chinese model, where the state guides and coordinates technological development, the European Union only “accepts” and supports innovation if it complies with a series of ethical and legal constraints. The precautionary principle, respect for human dignity, protection of privacy and prevention of harm are cornerstones of this approach, which is reflected in all regulatory initiatives, from the GDPR to the AI Act (European Commission 2019).

Here, there is a strong focus on the lawfulness and legitimacy of technological systems; innovations must be compatible with European treaties, the Charter of Fundamental Rights and the entire body of EU law. In addition, great emphasis is placed on values such as fairness, non-

discrimination and transparency. The approach aims to promote regulatory harmonisation at the supranational level to avoid inconsistencies in the application of rules.

The EU is seen as a global rule-maker in the field of new technology. It is expected that the AI Act, like the GDPR, will influence global policies, pushing all companies to comply with European regulations in order to access the market, confirming the so-called “Brussels Effect”.

However, the approach is not without its critics; the debate between precaution and promotion of innovation is one of the central issues in European institutional thinking.

## 1.6 Flexible vs Rigid Regulation

As anticipated, the United States adopts a decentralised and flexible approach to the governance of new technology, clearly distinguishing itself from more regulated and top-down models, such as those of the EU and China. This strategy aims to promote market-driven innovation with minimal regulatory constraints, emphasising the importance of self-regulation in the private sector. Flexibility is considered an advantage in promoting rapid technological innovation.

This model facilitates rapid development cycles and the accelerated adoption of new technologies, such as AI. Companies have the opportunity to innovate quickly and bring their products to market without delays that could result from a more centralised and rigid regulatory framework. In the context of autonomous vehicles – with minimal federal oversight – this has allowed companies such as Alphabet’s Waymo to reach millions of miles driven, demonstrating rapid implementation. This vision contrasts with the “regulation-first” approach, which involves delays in bringing technology to market due to much longer approval processes.

The vision on regulation supports the idea that “dynamic” regulations can have a positive impact on productivity. Flexibility gives companies latitude on how to achieve their goals, allowing them time to develop and validate; this encourages entrepreneurship and risk-taking (Majumdar & Marcus 2001).

The United States, with a sectoral strategy, assigns AI responsibilities to over 50 existing federal agencies, each with its own specialised expertise in its respective sector; this allows regulations to be tailored to the specific needs of different sectors and use cases (Davtyan 2024). This decentralised and sectoral nature allows for greater adaptability; policies are constantly interconnected with the emerging needs of new technology, which are characterised by rapid development and “radical uncertainty” regarding their future effects. It is evident a more resilient and

“future-proof” regulatory framework in the face of unpredictable developments in artificial intelligence. The importance of a bottom-up approach is acknowledged.

The adaptability is represented by the adoption of “regulatory sandboxes”: safe spaces that allow companies to test innovative products, services and business models in a real but controlled environment, without having to deal with the current regulatory framework (Fenwick et. 2021). This implies a reduction in barriers and normative costs for innovation. The main objective of sandboxes is to generate data and information useful for implementing regulation in the new digital world, facilitating greater collaboration between regulators, companies, start-ups, investors and end consumers.

An empirical study on Fintech technologies has shown that jurisdictions that adopt a more proactive approach, particularly through “regulatory experimentation” such as sandboxes, tend to attract more investment and be more attractive for starting Fintech operations (Fenwick et. 2021).

In the US, this approach is considered a preferable alternative to rigid *ex-ante* control, which, if applied excessively, can stifle innovation.

The United States’ perspective to AI regulation is characterised by a markedly flexible and market-oriented philosophy that prioritises innovation and economic growth through private sector self-regulation and a decentralised, sector-specific regulatory framework. The main benefits include rapid acceleration of innovation, a significant competitive advantage, and high adaptability to technological uncertainties. However, this flexibility may also entail risks such as regulatory fragmentation and potential accountability gaps. The preference for a regulatory framework that promotes experimentation and collaboration among stakeholders is seen as an effective way to address the radical uncertainties inherent in AI technologies, while fostering a thriving innovation ecosystem. The future path of AI governance in the United States will require a continuous balancing act between the need to stimulate technological progress and the need to ensure the responsible and safe deployment of AI.

### 1.6.1 Regulatory Rigidity

Contrary to what was just established, regulatory rigidity refers to a regulatory approach characterised by strict constraints, little or no discretion for regulated entities, and often an emphasis on means rather than desired outcomes. This type of regulation has significant impacts in terms of productivity, innovation, and the adaptability of businesses.

One of the most distinctive elements of this regulatory framework is the limited discretion of private actors, with a strong centralisation of power in the hands of public authorities. Companies do not have sufficient leeway to define customised approaches, but are strictly required to comply with the provisions laid down by the legislator. There could be a risk of standardising dynamic and technologically complex environments.

This is compounded by the imposition of specific processes on companies, rather than focusing on performance objectives. For example, early versions of US air and water pollution laws imposed standards such as “Best Available Control Technology (BACT)” or “Lowest Achievable Emission Rate (LAER)” without allowing economic considerations to be considered in the calculation (Majumdar & Marcus 2001, 172). Adopting these provisions without assessing operational sustainability reflects a so-called “end-of-pipe” approach, which leaves little or no autonomy to companies in choosing how to comply (Majumdar & Marcus 2001, 172).

When regulations are based on a “precautionary principle”, whereby anything that is not explicitly authorised is considered prohibited, there is a risk of making the system more rigid, entrenching the *status quo* and consequently compromising innovation.

The effects of this context are tangibly evident in terms of productivity. Businesses suffer a competitive disadvantage in complying with inflexible regulatory constraints. The obligation to adopt processes imposed from above, regardless of their economic efficiency, can lead to an inefficient allocation of resources with an increase in operating costs, without any direct performance benefits.

This could lead one to think that this context stifles innovation and entrepreneurial initiative, discouraging risk-taking; the so-called “brain drain” should come as no surprise. Moreover, startups in highly regulated sectors face greater difficulties in obtaining venture capital financing. Venture Capital (VCs) firms tend not to invest in contexts with many regulatory obstacles; slower go-to-market.

In the AI sector, these critical issues translate into delays in the commercialisation of technologies. Although the solutions may be the best in terms of product, they are at a competitive disadvantage due to their lack of speed.

The European Union’s approach to AI governance, exemplified by the AI Act, is often described as precautionary and risk-based, imposing stricter oversight on high-risk applications. This regulatory framework, while aiming to establish rigorous ethical and privacy standards, could result in slower implementation cycles and significant delays in commercialisation (Kulothungan and Gupta 2025). For example, for autonomous vehicles, the EU’s “regulation-first” approach has led to delays in commercialisation compared to the United States.

On paper, China has very burdensome regulatory requirements, such as the registration of AI models, which can significantly slow down innovation. However, China often adopts a flexible approach to enforcement for small and medium-sized enterprises (SMEs), selectively promoting innovation and economic growth, while large companies (“National Champions”) must comply fully (Chun et. 2024). This hybrid approach aims to combine regulatory consistency with the practical benefits of promoting innovation.

Regulatory rigidity in a context of high technological innovation undoubtedly has its limitations. Stringent constraints and a precautionary principle could represent a competitive disadvantage on a global scale. However, it should also be recognised that this approach stems from a desire to protect the public interest. It can provide a clear and predictable framework, which is useful for ensuring trust and preventing unknown side effects.

Personally, I believe that the real issue is not the presence of regulation itself, but its ability to combine ethical rigour and flexibility. The goal should be to build a regulatory environment that can adapt to unpredictable technological developments. For this reason, the rest of this thesis will focus on the details of the new emerging technology – artificial intelligence – and how the European Union is attempting to establish a clear regulatory framework on a global scale through the famous AI ACT.

## CHAPTER 2

### THE RAISE OF AI

#### 2.1 Definitions of AI

Formulating a definition of artificial intelligence is a complex task; there is no universal definition of the concept. There are various definitions, which can easily lead to confusion. The variety of definitions is not due to negligence, but is inherent in the phenomenon itself.

AI is commonly equated with algorithms, but this is not sufficient for an accurate analysis of the term. The term “algorithm” derives from the 9th-century Persian mathematician Mohammed Ibn Musa al-Kharizmi and refers to a specific instruction for solving a problem or performing a calculation. Consequently, if we defined AI as the use of algorithms, we would include many other activities.

In its most stringent definition, AI is the representation by computers of the ability to imitate intelligence, which is unique to humans.

A current definition of AI is that it is a technology that allows machines to imitate complex abilities common to humans. However, this definition does not offer much. On the contrary, it simply expresses the term “artificial intelligence” in a different way, nothing more.

Other definitions go further in explaining these skills and tasks. For example, the computer scientist Nils John Nilsson describes a technology that “functions appropriately and with foresight in its environment” (Nilsson 2009, 13).

The difficulty in formulating a valid definition is evident and unsurprising. AI is a simulation of something that we ourselves do not yet fully understand: human intelligence.

Since we do not have a complete understanding of what intelligence is, we cannot even define precisely how it can be reproduced artificially.

There is an interconnection between the study of human and artificial intelligence: advances in one field influence the other. One example is the game of chess, which AI has been successfully mastering for several decades: “If one could devise a successful chess machine, one would seem to have penetrated to the core of human intellectual endeavour” (Bostrom 2016, 14). In 1965, Russian mathematician Alexander Kronrod called chess “the fruit fly of intelligence” – meaning, the key to understanding it (Floridi 2014, 139). That is why people were amazed when a computer – Deep Blue – defeated world champion Garry Kasparov. Chess was considered the pinnacle of human intelligence, and Deep Blue’s victory was seen as a turning point.

However, today we know that chess is not the ultimate representation of human ingenuity; it is “simply” a problem that can be solved through precise rules and a multitude of alternatives. In this sense, a chess programme is not very different from a pocket calculator, which is also capable of performing operations that are too complex even for the most intelligent people (Sheikh, Prins, and Schrijvers 2023). But this does not make it an artificial form of human intelligence.

Although AI is capable of competing in one of the most complex games for humans, this does not necessarily demonstrate its ability to simulate all human behaviour. Something that is seemingly simple for humans, such as visually recognising a subject, is a very difficult task for artificial intelligence.

This phenomenon is known as Moravec’s paradox: some activities that are very difficult for humans, such as chess or advanced calculation, are relatively easy for computers (Moravec 1988). As Hans Moravec states: “It is relatively easy to get computers to perform like adults in intelligence tests or in the game of checkers, but it is difficult or even impossible to equip them with the abilities of a one-year-old child in terms of perception and mobility” (Moravec 1988, 15).

What was once considered an advanced example of AI is now simplified to a normal calculation; a phenomenon we can define as the “AI Effect”. According to Nick Bostrom, director of the Oxford Institute for Internet Governance, AI encompasses everything that impresses us at a given moment. When it ceases to amaze us, we simply call it “software” (Bostrom 2016).

The difficulty in formulating a true definition of AI does not stem from inaccuracy or superficiality, but from the fact that for a long time we have been unable to clearly establish what type of intelligence we wanted to imitate artificially.

Perhaps it is a mistake to use the term intelligence. The fact that a machine acts in ways that would be defined as intelligent if a human being behaved in the same way, does not necessarily mean that the machine is thinking, or that it can be considered intelligent. It is a view related to behaviour, not thought.

Alan Turing, as we all know, devised an experiment – Turing Test – to assess a machine’s ability to answer questions in such a way that the result is indistinguishable, in terms of its source, from the result of a human agent performing the same task (Turing 1950). Although the test was very successful, the experiment does not justify the term “intelligent”. The important thing is to perform a task successfully, in such a way that the result is as good as, or better than, what human intelligence would have been able to achieve (Floridi 2022, 45). Here, the how is not in question, only the result.

It is no wonder that, someone claim that modern technology is characterised by only one component: prediction, which is why they talk about “predictive machines” (Agrawal, Gans, and Goldfarb 2018). On the contrary, what is truly surprising is the view of philosopher Daniel Dennett,

who considers AI to be a new kind of entity, comparing it to oracles: beings that make predictions but, unlike humans, possess no personality, consciousness, or emotions (Dennett, 2019).

The reality is that AI performs actions similar to those performed by a human agent, but it does so in a completely different way, in my opinion, with zero intelligence. Despite this, the term “artificial intelligence” has become so widespread that it would not make sense to replace it now.

Today, artificial intelligence is often equated with the most advanced technologies and has experienced a remarkable acceleration in recent years. One of the main factors driving growth has been progress in *Machine Learning* (ML), which has led to the innovation now known as *Deep Learning* (DL). This technology has enabled computers to recognise the faces of different individuals. Unlike traditional approaches, such as “Expert Systems”, in which computer systems follow predefined rules, ML and DL algorithms, also known as *self-learning*, are able to recognise patterns in data.

*Machine Learning* and *Deep Learning* are both sub-sets of AI. The first refers to the study of computer system that learn and adapt automatically from experience without being explicitly programmed. With these models, computer scientists can “train” a machine by feeding it large amounts of data. In fact, the more data the machine parses, the better it can become at performing a task. On the other hand, *Deep Learning* must be seen as an evolution of ML. Where ML algorithms generally need human correction when they get something wrong, the second one can improve their outcomes through repetition, with any intervention. If one were to explain it from a technical standpoint, DL is a ML technique that layers algorithms and computing units – or neurons – into what is called an artificial neural network, allowing the system to identify and learn relevant features automatically without human guidance (Coursera 2023).

Although it is important to classify the main subcategories of AI, it does not help in formulating a definition. It would be wrong to limit the definition to a technological component, as this could lead to other applications of AI being overlooked.

The nature of this scientific discipline implies that the definition is bound to change over time. Within this context, the definition adopted by the European Commission’s High-Level Expert Group on Artificial Intelligence (AI HLEG) is considered one the most appropriate: “*Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition [...] and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model [...] and adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several*

*approaches and techniques, such as machine learning [...], machine reasoning [...], and robotics”* (European Commission, Joint Research Centre 2020).

It is a sufficiently broad definition to encompass all applications currently classified as AI, from advanced *Machine Learning* and *Deep Learning* technologies to more traditional approaches, while providing scope for future developments. Furthermore, it is sufficiently rigorous to distinguish AI from algorithms and digital technology in general.

It is important to note, however, that the current applications classified as AI under this definition fall into the category known as “Narrow AI” or “Weak AI”. Artificial Narrow Intelligence (ANI) systems can perform one or a few specific tasks and operate within a predefined environment e.g., those exploited by personal assistants Siri, Alexa, language translations, recommendation systems, image recognition systems and face identification (European Commission, Joint Research Centre 2020). It is very efficient at processing data, improving productivity in many practical applications; it is capable of translating more than 100 languages in real time and accurately recognising billions of faces and objects. ANI is also very useful for performing routine, repetitive and tedious tasks that humans do not enjoy.

However, Artificial Narrow Intelligence it is incapable of generalization, i.e. to re-use learned knowledge across domains, e.g., the ANI capable of image recognition cannot transfer its knowledge in the domain of speech recognition.

The inability to generalise distinguishes ANI from Artificial General Intelligence (AGI), which refers to machines capable of performing any intellectual activity that a human being is capable of performing. AGI aims to replicate human cognitive abilities, including reasoning, learning and adaptation to different domains. It is the subject of representation in science fiction; machines equipped with consciousness, self-awareness and emotions – currently not possible today.

Going beyond this concept, what truly belongs to the realm of science fiction – for the time being – is Artificial Superintelligence (ASI), which refers to a hypothetical future intelligence capable of surpassing human capabilities in all areas: creativity, emotional intelligence, and strategic thinking (Bostrom, 2016).

Although I enjoyed tackling the definition of “Artificial Intelligence” in an academic and philosophical manner, with a dash of technological science, it is necessary to immerse oneself in the definition from a regulatory point of view. The definition proposed by the European Commission’s High-Level Expert Group on AI in 2018 was indeed sufficiently broad to encompass both traditional approaches, such as symbolic reasoning, and more advanced techniques like machine learning and deep learning, thereby leaving scope for future developments. From a scientific and technical standpoint, it remains a valuable and inclusive reference.

But, for legislative purposes it is considered inadequate. Its detailed reference to specific methodologies made it vulnerable to obsolescence, its cognitive framing risked excluding simpler predictive systems, and its level of technicality rendered it difficult to apply in legal practice. For these reasons, policymakers eventually preferred a more abstract and technologically neutral definition – one that could provide ensure clarity and enforceability in regulatory contexts.

This shift does not imply abandoning the value of technical or philosophical perspectives, but rather reflects the need, in a regulatory environment, for a stable concept that can serve as a reference point for governance.

In particular, the European Union, in coordination with the OECD, has recently adopted a regulatory definition of “artificial intelligence system” within the AI Act (EU Regulation 2024/1689). This definition reads as follows: *“AI system means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments”* (European Union 2025).

This definition, with the notion of “AI” and “machine-based” system – referring to the fact that AI systems run on machines – aims to ensure legal certainty, facilitate international convergence and broad acceptance, providing the flexibility to accommodate the rapid technological developments in this field (European Union 2025). Its function goes beyond the philosophical or scientific ambiguities surrounding the term “intelligence”.

The AI Act’s definition should therefore be understood not as the final word in the conceptual debate on artificial intelligence, but as a pragmatic legal framework designed to guide its regulation.

## 2.2 AI and Its Impact on Business

Following the analysis of the debate on the definition of AI, it is right, and above all necessary, to analyse how this technology is impacting the business world. By now, it is well known that it is establishing itself as a disruptive force within the global corporate landscape, significantly determining strategy, operations, and consequently, competitive dynamics. This revolutionary technology – along with all its nuances – is ushering in a new era of data-driven decision-making, where data is no longer considered a mere by-product of business operations, but is strategically positioned as a key asset.

Considering the automotive sector, one type of artificial intelligence that is redefining business models and corporate strategies is *computer vision*: a technology that involves the automation of the processes of observation and interpretation of visual information, which can take the form of photographs, videos or real-time streams from the physical world. It is no coincidence that, Autonomous or semi-autonomous cars currently being developed by companies such as Tesla, BMW, Audi and Volvo incorporate a sensor system consisting of a multitude of cameras designed to perceive the surrounding environment: vehicles, pedestrians, road signs, obstacles and traffic lights. These systems use *computer vision* techniques to transform visual data into useful information for autonomous decision-making (Dong and Cappuccio 2023). The technique used by Tesla it is called “*Tesla Vision system*”, which utilises eight cameras and a complex neural network to manage environmental perception tasks. Today it is being redefining the very concept of the vehicle itself – not only transforming how manufacturers conceive, design, and produce it, but also reshaping consumers’ expectations and interactions with mobility as a whole.

From a strategic point of view, integrating AI into business processes is no longer optional: it has become an essential requirement for companies that desire to remain competitive in the digital age. This is hardly surprising; its distinctive strength lies in its ability to process enormous amounts of data, identifying patterns and insights that are often difficult or impossible for the human eye to recognise. Indeed, it allows to anticipate market trends and make decisions based on solid quantitative analysis to an unprecedented degree.

However, it is also important to note that in order to adopt successfully AI, companies must adapt their business models and align technology implementation with overall strategic objectives. This concept echoes the views of many modern philosophers: “it is the environment that must adapt to artificial intelligence, not the other way around” (Floridi 2019).

That being said, at the heart of this strategic transformation lies the aforementioned machine *learning*. Its importance lies in the capability in enabling systems to learn from data and identify recurring patterns as a basis for predictions. The ability to use data to make more accurate predictions about the future is hugely valuable in many contexts. Considering energy supply organisations, Google’s *DeepMind* has developed an AI system that uses weather forecasts and turbine data to predict the inflow of energy from wind farms 36 hours in advance (DeepMind 2019). By doing so, it is possible to optimise the use of wind energy, despite the variability of natural elements.

While, in the financial industry, which is closely linked to risk management, *machine learning* is applied on a large scale for credit rating systems and fraud prevention. A case point is represented by Mastercard which uses a system called *Decision Intelligence* to detect potentially fraudulent activity by analysing transaction patterns (Mastercard n.d.).

This predictive value is also evident in other sectors. Some supermarket chains are experimenting the *dynamic pricing* technique to reduce waste and maximise profits. Using algorithms, it is possible to optimise product assortment or automatically manage discounts based on data regarding food shelf life, store location and historical sales trends (Heijn 2019).

Some strategic applications of AI also include enhancing product development processes, which can be achieved by analysing customer feedback. In particular, *Natural Language Processing* (NLP), the technology that enables systems to read and interpret human language, plays a key role in this field. Through techniques, such as *sentiment analysis*, companies can extract valuable information from reviews, social media posts and market surveys, anticipating consumers' preferences and needs.

Beyond its influence on directional planning visions, this technology is increasingly shaping the executive-operational structure of contemporary enterprises. Once again, this is hardly surprising, since the automation process enables organisations to streamline operations and speed up task execution: significantly impacting cost savings and overall operational efficiency (Shrinivas and Shetty 2020).

In this regard, it is difficult not to mention *Robotic Process Automation* (RPA). In simple terms, it consists of bots that automate manual and repetitive tasks. Bots can interact with different systems and applications, just as a human being would, performing tasks such as automatic data entry, email management and form filling. This is vitally important, as it involves the reallocation of human resources to more strategic activities that are considered to have greater added value. A crucial application lies in predictive maintenance, where AI systems examine sensor data and past maintenance records to anticipate equipment failures before they occur, allowing companies to plan maintenance proactively, maximising asset utilisation (Shrinivas and Shetty 2020).

The influence of technology, at an operational level, is equally significant in the media and entertainment industries, where AI is widely used to personalise products and services based on user preferences. Platforms such as Netflix, YouTube and Spotify employ *recommendation systems* to suggest relevant content based on previous choices (Sheikh, Prins, and Schrijvers 2023, 51).

Similarly, the personalisation process is also a fundamental pillar for e-commerce giants such as Amazon, Alibaba e Zalando which analyse user profiles and adapt their marketing strategies accordingly. These latter applications are not merely aimed at enhancing customer satisfaction and loyalty; they also represent a broader shift towards predictive and adaptive business models, in which continuous learning from user behaviour becomes a key source of competitive advantage (Sheikh, Prins, and Schrijvers 2023, 51).

A further influence within organisations stems from AI's ability to reshape the operational dimension of the Human Resource department, also generating broad organisational implications. Through advanced tools – mainly based on *machine learning* and *natural language processing* – AI support talent management, recruitment, and employee engagement (Shrinivas and Shetty 2020). Today, in fact, it is very common to use recruitment algorithms that can quickly analyse candidates' qualifications and, subsequently, match them with the most suitable job positions. One example is the American start-up Juicebox, which on 15<sup>th</sup> October 2025 raised €36 million in Series Round A from Sequoia Capital – one of the world's largest venture capital funds with more than 50 billion dollars in assets under management (Paffenholz 2025). By automating these processes, AI enhances the efficiency and objectivity of recruitment while allowing HR professionals to devote greater attention to strategic tasks such as talent development and organisational culture.

In essence, the diffusion of artificial intelligence within contemporary enterprises signals a profound epistemic shift rather than a mere technological evolution. It invites organisations to rethink the locus of intelligence itself. What emerges is a new form of strategic rationality, one that privileges anticipation over reaction and transforms data into a medium of organisational sense-making. Yet this phase of transformation is only an intermediate stage in a broader continuum. As artificial systems evolve from interpreting to generating information, a new paradigm unfolds – that of Generative AI – where the boundary between analysis and creation becomes increasingly blurred, redefining both the nature of innovation and the role of human agency within it.

### 2.3 Generative AI: The New Frontier

Generative Artificial Intelligence (GenAI) is an emerging technology that is spreading rapidly in modern society, with such a profound impact that it could represent a new industrial revolution. The advantages are now well known, as they are for AI, but with one particular feature: GenAI has the ability “to create”. Put this way, it seems to suggest the emergence of a new kind of entity – a notion that recalls the view of philosopher Daniel Dennett, previously criticised in the opening paragraph of this chapter. However, without mentioning to its creative capacity, it is impossible to fully explain what Generative AI truly is.

GenAI systems fall within the broad category of *Narrow Artificial Intelligence* (ANI) and are powered by *Machine Learning* techniques. According to McKinsey's definition, Generative Artificial Intelligence describes algorithms that can be used to create new content, including audio, code, images, text, simulations, and video (McKinsey & Company 2024).

Understanding its functioning is far from difficult. The software starts from requests or descriptions (prompts) formulated in natural language by a user (human or software) and generates texts from texts (Text-to-Text), images from texts (Text-to-Image) or even images from images (Image-to-Image). The outputs are combinations of the data used to train the algorithms. Since the amount of data output for the training process is massive, the results are considered creative. But in reality, the generation is only due to a mix of sources, where the enormous amount of data makes the output new. Perhaps this creativity is merely reworking.

Generative AI originates from advances in *Deep Learning*. Over time, different architectures have driven its evolution, progressively expanding the creative potential of machines.

The first major step was the introduction of *Generative Adversarial Networks* (GANs), made up of two neural networks that challenge each other: a “generator” that produces synthetic images or data, and a “discriminator” that evaluates their authenticity. This rivalry enables the system to generate increasingly realistic results, paving the way for the first creative uses of AI.

A second turning point came with *Transformer* models, which made it possible for AI to understand and generate sequences such as text or code. This design, based on a mechanism called *self-attention*, gave rise to *Large Language Models* (LLMs) like ChatGPT, capable of producing coherent text, writing code, or summarizing information.

Today, the field is dominated by *foundation models* – very large neural networks, often built on the *Transformer* design, trained on enormous and diverse datasets. Unlike earlier algorithms, they are not limited to one task: the same model can write, analyse images, generate sound, or combine them all. In this sense, LLMs are simply the language-focused branch of this broader family.

Considering that its primary characteristic lies in its ability to generate a wide variety of content, it can be applied across numerous domains. It can generate textual content, such as poems, political statements, and academic papers, which are difficult to distinguish from content generated by humans (Nah et al. 2023, 279). Examples of AI-generated images include artworks, synthetic faces, and magnetograms of the Sun, which range from humanities to sciences (Nah et al. 2023, 279).

These applications will give rise to new ways of working enabled by technology. Over time, technology will move from performing ordinary tasks to more sensitive ones. These will involve decisions of greater importance, which may even involve moral judgment (Feuerriegel et al. 2023).

The sector’s expansion is remarkable: at the beginning of 2023 alone, global investments in generative AI exceeded \$12 billion, with an average annual growth rate of 74 percent between 2017 and 2022 (Chui et al. 2023). This surge reflects the extraordinary speed of innovation, as demonstrated by the rapid succession of OpenAI new models such as and GPT-5 and Sora. Notably, this wave of innovation is largely concentrated in North America, which accounts for approximately 75% of global

private investments in GenAI between 2020 and 2022 (Chui et al. 2023). Such geographic asymmetry highlights both the technological leadership of the United States and the competitive imbalance that currently defines the global AI landscape.

The concern is that in a rapidly changing global landscape, the EU is increasingly alarmed by the prospect of remaining marginalised in the generative AI “revolution”, having already lost momentum in the previous digital phase. According to the 2024 Mario Draghi report, European competitiveness is at risk of “slow agony” if accelerated industrial policies and massive investments are not implemented (Draghi 2024).

The gap is evident in the AI ecosystem: for example, a report highlights that over 70% of foundation models since 2017 are American, compared to only around 15% from China and an even smaller share from Europe (Simion 2024). In particular, the 2025 European Commission-JRC report highlights how the EU, while recognising the potential of Gen AI to increase European productivity by up to 3% per annum by 2030, is held back by investment, infrastructure, scalability and fragmentation of the internal market (Sukharevsky et al. 2024).

This leads to the realisation that it is not just a question of adopting new technologies, but of redefining Europe’s industrial and strategic paradigm: the challenge is not only technological, but also geopolitical and systemic. The EU must ensure that it does not once again miss the boat when it comes to the digital revolution, by harnessing its talent and innovative culture.

## 2.4 GenAI as a Driver of Business Disruption

In contemporary management theory, the term *disruption* refers to a process that radically alters the structure and logic of economic activity. It does not merely describe an innovative advance, but indicates a quantitative and qualitative transformation in the way value is created.

The phenomenon recalls what the economist Joseph Schumpeter, back in 1942, defined as “*creative destruction*”: a process through which capitalism renews itself, demolishing existing structures and replacing them with new ones. In the digital age, this dynamic goes beyond the concept of material production and extends to the cognitive sphere. Today, technologies also amplify cognitive abilities, ushering in a “second machine age”.

In this evolution, Gen AI is the ultimate representation of *disruption*. Unlike previous forms of automation, it intervenes directly in creative and communicative processes, representing a paradigm shift in the evolution of artificial intelligence.

The impact of GenAI on business is unprecedented. According to a McKinsey analysis of 63 business use cases, estimates indicate that generative AI could add between \$2.6 trillion and \$4.4 trillion in value annually (Chui et al. 2023).

GenAI excels as an empowerment tool for “knowledge workers”, who spend much of their time searching for and gathering information. With this technology, this time can be drastically reduced, as the systems can act as a “virtual expert.”

However, continued use and adoption by end users and organizations necessarily involves a series of socio-technical considerations. This is due to the fact that, in recent years, AI capabilities were mostly understood as analytical, suitable for decision-making tasks. Today, on the contrary, AI is also gaining deeper consideration due to its ability to perform generative tasks. Although the process of content generation could still, in some respects, be interpreted as analytical, its outcomes may nonetheless be considered creative or even artistic, as Generative AI combines existing elements in novel ways.

With the advent of this technology, the assumption of the “primacy of human agency” could be compromised, with important implications for how we conceive of the relationship between humans and AI, based on their respective capabilities. In particular, on the topic of interaction between humans and AI, so far, with analytical AI, the concept of *delegation* has often been discussed in order to establish a hierarchy in decision-making processes (Baird and Maruping, 2021). Whereas, with generative AI, humans use prompts with the aim of generating content; AI interprets the user’s intentions and provides responses that, in turn, stimulate new prompts. This may seem like a *delegation model*. However, the next phase deviates from this pattern, as the output generated can be suggestive for the user, directly or indirectly influencing further decisions. Consequently, rather than a *delegation model*, it is more of a *co-creation model*: a collaborative practice in which the parties share diverse perspectives and contributions to guide the design process.

In light of this, the interactions described above give rise to the concept of *hybrid intelligence*, understood as the integration between humans and AI, which must aim to combine the strengths of both. This approach supports the need to overcome the limitations of the two types of intelligence, combining human intuition, creativity, and empathy with the computational power, precision, and scalability of AI systems in order to improve decision-making and problem-solving capabilities (Dellermann et al., 2019).

To maximize this potential, companies must develop new skills. In this regard, *Prompt Engineering*, the art of designing inputs (prompts) to guide the model toward accurate and desired outputs, becomes a crucial form of digital literacy. The ability to formulate meaningful questions and

judge the quality of the answers provided by the machine is essential, as GenAI is based on linguistic probability rather than logical reasoning.

Entrepreneurs and small business owners have much to gain from using *Prompt Engineering*. Since entrepreneurs need to know how to use market information both during launch and growth processes, GenAI, with LLMs, provides contextual information that facilitates entrepreneurial decision-making. Historically, entrepreneurship has always been marked by a high degree of uncertainty due to information asymmetry, and the best entrepreneurs are those who are able to reduce it. Now, with GenAI, the landscape is changing, especially when it is used properly.

Indeed, it is no coincidence that many modern business schools are integrating GenAI into their curricula – not defining it as a simple “answering machine,” but as a partner, capable of enhancing entrepreneurial thinking and assisting in the creation of business models. The entrepreneur in training does not merely learn to use GenAI, but to think with it, developing a critical awareness of the limits and potential of artificial intelligence as a co-creator of knowledge.

Considering instead the corporate dimension, GenAI has a wide range of applications given its characteristics, but most of its potential value is concentrated in five main functional areas: Customer Operations, Marketing, Sales, Software Engineering, and Research and Development. What these functions have in common is that Gen AI generates value by taking on the role of a virtual collaborator.

In Customer Operations, the technology is revolutionising interaction through digital self-service and increasing operator skills. Chatbots, with generative technology, can provide immediate and personalised responses even to complex requests, reflecting the brand’s identity. This means that adopting this technology could reduce the volume of contacts handled by human agents. Meanwhile, for operators, GenAI acts as an “assistant”, providing response suggestions and call scenarios, as well as quickly accessing customer data for fully personalised assistance. AI-based assistance could improve service quality, especially for less experienced agents. In the future, we should not be surprised to see hybrid cognitive ecosystems, where interaction between humans and machines will be completely normal.

In Marketing, GenAI is essential for large-scale personalisation. The technology allows for the creation of tailored messages that reflect the interests of individual customers. It is also very useful for generating initial drafts of creative content such as slogans or product descriptions. The content creation process is the main operational benefit, as it significantly reduces the time needed for drafting. Another transformative effect concerns the ability to use data more effectively. GenAI helps teams overcome the difficulties associated with unstructured and disconnected data – from different databases – by interpreting even abstract sources. This translates into better use of information such

as regional performance and purchasing behaviour, generating data-driven marketing strategies, such as targeted customer profiles.

Generative artificial intelligence could also change the way B2B and B2C companies approach Sales. It could identify and prioritize potential customers by creating comprehensive consumer profiles using a combination of data, while also suggesting actions to take to improve the engagement process. In addition, the technology could help sales representatives cultivate leads by synthesizing product sales information and creating conversation scripts to facilitate dialogue with customers – including ideas for up-selling and cross-selling activities. However, integrating these technologies requires careful management, as models trained on public data may reproduce copyrighted or biased content; hence, human oversight remains vital to preserve creative integrity and brand credibility.

In software engineering, the use of Generative Artificial Intelligence is opening up new opportunities. Today, engineers can use it for programming, coding, or training language models capable of generating code from simple prompts. Its relevance is set to grow as large companies, technological or otherwise, increasingly integrate software into their products and services. For example, the value of today's new vehicles is characterised by digital features such as IoT connectivity and adaptive cruise control. The impact of GenAI can have a significant effect on business productivity. The value comes from reducing the time spent on certain activities. One study found that developers using Microsoft's GitHub Copilot complete tasks 56% faster than those who do not use it (Chui et al. 2023). With faster development speeds, a shift in skills towards code and architecture design is expected.

One business function that is greatly underestimated in terms of GenAI's influence is R&D. It is estimated that adopting this technology could generate a productivity increase of between 10% and 15% of total R&D costs (Chui et al. 2023). The reason lies in its ability to support every stage of the development process, from data collection to the physical validation of prototypes. In life sciences, *foundation models* are crucial for *generative design*, identifying candidate molecules and dramatically accelerating the drug development process.

The distinctive feature lies in the fact that, unlike traditional machine learning techniques – often constrained by high costs or the need for vast amounts of data – generative AI models offer greater versatility and adaptability. Such innovations, beyond generating real leaps in performance for individual companies, could have disruptive effects on global economic growth.

However, it's important to note that the adoption of GenAI is a double-edged sword for companies. It promises to increase efficiency and stimulate creativity, but simultaneously it also brings with it fundamental challenges.

One of the most immediate concerns is the issue of data privacy and security. There is a need to implement measures to ensure that information remains confidential, as there could be a risk that sensitive data could be released to the public or used to train the models themselves.

Companies must be extremely cautious when interacting with these systems, considering the risk of disclosure of confidential information about their organizations. A significant example is JPMorgan, which in 2023 restricted the use of ChatGPT precisely because of this concern (Horowitz 2023).

From an operational standpoint, the main challenge lies in the reliability of AI-generated outputs. GenAI models are known for the phenomenon of *hallucination*, whereby the generated content appears plausible but is actually meaningless and not faithful to the input provided, thus generating misinformation. Relying on incorrect answers is dangerous, especially in contexts that require a certain degree of accuracy.

A related risk is over-reliance on these tools. There is no need for data to explain how beneficial it is to harness the power of GenAI. The problem, however, is that many workers may become accustomed to receiving answers from systems without careful verification. Rather than bringing benefits in terms of productivity, this dependence could have the opposite effect: it could hinder the development of essential skills such as critical thinking or creativity itself.

Ethical concerns are also central to this issue. In reality, they are central to all dimensions: social, economic, and legal. Remaining in the corporate world, the problem arises in the biases inherent in training data. When GenAI, or AI in general, assists in decision-making processes, the existence of biases, such as gender or racial stereotypes, can have harmful consequences for the surrounding environment. For this reason, it is vitally important that the developers of such technologies, particularly tech companies, ensure that training data is representative in order to guarantee fairness.

Ultimately, returning once again to Schumpeter's vision, GenAI is the ultimate representation of "*creative destruction*" as well as a challenge to the very paradigm itself. It embodies its purest essence, given its central role in the process of replacing old forms with new ones; yet, through its generative capacity, it also calls into question the figure of the Schumpeterian entrepreneur, the traditional engine of innovation. Whereas "*creative destruction*" necessarily implied a human act, GenAI shifts the balance toward a hybrid plane in which initiative is balanced between humans and machines.

We are witnessing the emergence of a new "immaterial" protagonist that marks a point of no return: innovation is no longer merely a force transforming capitalism, but a phenomenon that compels us to rethink its very foundational categories – creativity and labour.

## 2.5 Beyond Business: The Socio-Economic Dimensions of Artificial Intelligence

Analysing Artificial Intelligence solely from a business perspective is a necessary but insufficient condition for understanding its true significance. It is necessary to extend the analysis beyond the corporate dimension and consider its implications on a broader scale.

According to a study conducted by the consulting firm Accenture, which surveyed 12 developed economies, AI could double annual global economic growth rates by 2035 (Szczepański 2019). This growth is due to three main reasons: a sharp increase in labour productivity, estimated at up to 40%, made possible by technologies capable of maximising workforce time management; the creation of a new virtual workforce, defined as “intelligent automation”, capable of self-learning and problem solving; and the spread of innovation itself, which will generate new revenue streams in various sectors.

Moreover, a study by PricewaterhouseCoopers (PwC) estimates that global GDP could increase by up to 14%, equivalent to US \$15.7 trillion, by 2030, thanks to the acceleration in the development and adoption of AI (PwC 2024). Specifically, the consulting firm identifies two main channels through which AI will impact the global economy. The first, in the short term, focuses on productivity gains from the automation of routine tasks, an effect that will expand, above all, in capital-intensive sectors such as manufacturing and transport, with the use of robots and autonomous vehicles. Furthermore, as mentioned in the previous paragraphs, productivity will also increase because companies will supplement their workforce with technologies based on assisted and autonomous intelligence, leaving time for activities considered to be of greater value. The second channel, with long-term prospects, is the current and future availability of personalised, superior-quality products and services. The offering will stimulate consumer demand, triggering a “virtuous circle” of more data *touchpoints*, better insights, superior products and, consequently, higher consumption.

Structurally speaking, AI can have highly disruptive effects on the economy, creating the risk of “super-firms” that act as centres of wealth and knowledge. McKinsey suggests that AI and automation could both facilitate the rise of large-scale organisations and enable small players and individuals to take on jobs currently performed by larger companies (Szczepański 2019, 5). This could lead to a *barbell-shaped* economy, where medium-sized companies would be penalised.

Moreover, there is a risk of increased competition and a widening gap between technology leaders and laggards: companies that are “early adopters” of AI will benefit disproportionately,

contributing to a “winner-takes-all” phenomenon, which could even slow aggregate productivity growth if the benefits do not spread to laggards (Szczepański 2019, 6).

On the labour market front, the consequences are complex and uncertain. Nevertheless, a significant shift in the workforce across many sectors is foreseeable, as well as a transformation in the very nature of work itself, making the process of reskilling essential. Occupational polarisation is a tangible risk: lower-paid jobs that typically require routine manual and cognitive skills stand the highest risk of being replaced by AI and automation, while well-paid skilled jobs that typically require non-routine cognitive skills will be in higher demand (Szczepański 2019, 6).

The growing demand for highly skilled jobs capable of using AI could lead to higher wages, but at the same time others could suffer wage compression or unemployment. The changes could therefore worsen the overall distribution of income. Obviously, this depends greatly on how quickly this phenomenon occurs. This may lead to a “paradox of plenty”: society would be far richer overall, but for many individuals, communities and regions, technological change would only reinforce inequalities (Szczepański 2019, 7).

In light of these transformations, we have reached a point where technological progress cannot be separated from legislative issues. What emerged from the analysis conducted is that AI is no longer merely a productive force, but a phenomenon capable of redefining markets, labour structures and even cognitive hierarchies. Just as previous industrial revolutions required the codification of labour and company law, the Industrial Revolution 4.0 requires a regulatory architecture capable of guiding non-human agency at every level.

The challenge lies not in controlling this technology, but in integrating it into the principles of democratic accountability. In this perspective, regulation must take a proactive approach, capable of recognising the dual nature of AI: on the one hand, an engine of economic growth; on the other, a source of ethical and social externalities.

This is where the European Union’s objective with the “Artificial Intelligence Act” comes in: the first global attempt to build a regulatory model that strikes a balance between innovation and the protection of fundamental rights.

However, fear and haste in regulating something that we are not sure what it is and what it will be could jeopardise innovation itself, creating a braking effect, compromising European competitiveness.

## **CHAPTER 3**

# **THE EUROPEAN AI ACT: BETWEEN REGULATORY AMBITION AND TECHNOLOGICAL REALITIES**

### 3.1 From Normative Ideal to Final Text: Genesis and Structure of the AI Act

Many have questioned why the European Union has decided to embark on such a far-reaching legislative path. These considerations are entirely legitimate. After all, this technology did not originate here, nor is this where the world's largest digital companies are based. It is also clear that in the race towards the new industrial revolution, the United States and China are in pole position, while Europe is really struggling.

Explanations for such questions are not in short supply.

One of the most common is on the subject of “nobility” (Viola & De Biase 2024). Indeed, Europe is somewhat the home of the welfare state; it is the only continent where there is a functioning free healthcare system that is accessible on a large scale. Furthermore, despite the current crises and problems in the states that now make up the EU, a high level of social protection and defence of fundamental rights has been achieved. In light of this “nobility”, it may seem logical to want to take on such responsibility, not only for one's own citizens, but also for the rest of the world.

In addition to explanations of goodness, other very curious ones emerged. One in particular that caught my attention is on the issue of population aging (Viola & De Biase 2024). According to this view, given that there is a certain degree of maturity in the EU, there is a greater tendency to establish rules. This is partly true; in fact, dreaming of something new and different is something that belongs to young people.

With such principles, it is not surprising that such a civilization addresses the risks that may arise with the development of Artificial Intelligence with a regulatory approach.

Although these explanations are entirely plausible, or at least partially so, someone is probably missing something. Don't get me wrong, the core values of the European Union are the most virtuous, but part of me is inclined to believe and even hope that there is also a geopolitical strategy at play here.

Fortunately, the European Union is traditionally recognized as the main guardian of fundamental rights, democratic pluralism and the protection of individuals. Alongside this value-based dimension, it is worth emphasizing that the EU is also perhaps the only jurisdiction in the world capable of influencing the regulatory systems of the rest of the world. The “Brussels Effect” discussed in the first part of the thesis does exist.

As was the case with the GDPR, the EU is likely aiming to leverage the combination of the size of its market and the high costs of maintaining different compliance regimes: global companies are thus obliged to comply with European requirements in order to operate in the EU. The AI Act should be interpreted as an expression of European values, but given today's circumstances, it probably also has a clear strategic dimension. Regulating first means occupying the global regulatory space, establishing the EU as a rule-maker in AI governance and in the race for innovation. Through regulation, Europe not only aspires to safeguard its values, but also to exercise a position of power at the international level.

Beyond thoughts on the subject – which will be explored in greater depth later – what is certain is that this legislative act represents fundamental values, both in terms of content and procedure.

The European legislative process is based on a defined inter-institutional balance, where no decision-making body should prevail over another. Obviously, this balance is the goal of all constitutional orders, but in the European context, the principle of separation of powers and roles is more important than elsewhere (Viola & De Biase 2024). The body with legislative initiative is the Commission, as in the case of artificial intelligence, but the legislative bodies are the Council and the Parliament. It is a system characterized by allowing full representation of all interests. If one tries to analyse the functioning of the institutions, it becomes clear that the decision-making process, both in theory and in practice, is a mechanism oriented toward a balanced consensus among the three actors involved. In fact, the European Commission's proposals pass through the legislative branches, which amend and discuss them in trilogue talks before they become law. This process can take time and, although it does not require formal unanimity, it cannot be concluded without agreement between all the actors involved. In the case of the AI Act, negotiations lasted a full 56 hours.

These constitutional principles are similar in the United States, but decisions are made more quickly there. In Europe, on the other hand, the process is slower and decisions are probably better, in order to protect the interests of its citizens in a balanced way when dealing with very complex issues.

The journey began, informally, back in 2019, when Ursula von der Leyen stated her intention to start the process of regulating AI (Viola & De Biase 2024). The Commission President's position was not taken very seriously, which is not surprising, given that GenAI did not yet exist at the time.

The international community began to take this position seriously in 2020, when the Commission published its first reflection, the so-called "White Paper" on AI, dealing with two ecosystems that must coexist.

An ecosystem of excellence, whose goal is to create the conditions necessary for AI to develop in the EU as a technology that is useful for its citizens, a true collective good, in which the benefits are distributed equally between those who invent and those who use it (Viola & De Biase 2024, 88).

The ecosystem of trust, on the other hand, refers to the creation of safeguards that allow citizens to use technology safely, without fear of being exposed to risks (Viola & De Biase 2024, 89).

The Commission's objective was to affirm the need for regulation capable of containing potential risks and maximizing benefits.

The AI Act classifies the risks to be considered as a pyramid. At the top of the pyramid are unacceptable risks (banned), just below are high risks, further down are limited risks, and finally, minimum risks.

It is a regulatory framework based on the principle of risk. This approach may ensure that the legislation remains adaptable even in the face of future and as yet unknown iterations of AI technologies, placing ethics at the heart of the regulatory process.

At the top of the pyramid are applications for which the regulation provides an absolute prohibition on commercialization and use. The reason is that they risk creating significant and irreparable social harm.

Prohibited applications include the use of AI-enabled manipulative techniques aimed at persuading individuals to engage in undesirable behaviour, thereby compromising their freedom of choice.

The use of AI to evaluate and classify people based on their social behaviour, known as *social scoring*, is also banned. If this practice were not banned, we would risk living in an Orwellian society, somewhat like China, a completely surveilled society, where the evaluation of a simple driving licence application could be based on how many bottles of wine or whisky one buys on average.

Furthermore, the identification of emotions or intentions of natural persons through Biometric data in workplaces or educational institutions is considered an unacceptable risk, except in cases related to medical reasons. It would not be ideal if the manager or teacher were aware of any critical feelings towards them.

All applications prohibited under the AI Act share the potential risk of threatening fundamental values such as human dignity, freedom, equality, democracy, data protection and the rule of law.

Slightly below the top of the pyramid are the high-risk systems. The core focus of the AI Act revolves these systems, as most obligations and protective measures outlined in this act refer to high-risk AI applications. High-risk AI systems are those that are deemed to negatively impact safety or

fundamental rights of EU citizens and given that presumed risk need to be assessed before being put on the market and also throughout their life cycle (Deloitte 2024).

Identifying an AI system as “high risk” occurs through two distinct regulatory channels. The first route involves inclusion in the specific applications listed in Annex III of the AI Act. The second occurs when the AI system is incorporated into products already covered by harmonised EU safety legislation, as specified in Annex I. In both cases, the operator using or supplying the AI system is responsible for determining whether their product falls into these categories. Given the rapid evolution of systems, it is expected that such assessments will be made on an ad hoc basis.

To ensure that the classification is applied proportionately, covering only AI that poses a demonstrably high risk, the EU has introduced significant *opt-out* exceptions for suppliers. An AI system that would at first glance fall into the high-risk category but does not pose significant risks to health, safety or fundamental rights may be excluded from this classification, provided that operators provide the necessary documentation and register in advance (Deloitte 2024).

Annex III aims to designate specific contexts in which the use of technology is most likely to pose high risks to consumers (Deloitte 2024). The categories of application defined in this annex cover various sectors. They include Critical Infrastructure, such as traffic management or water and gas supply. One sensitive sector is Education and Professional Training, where use cases include admission to institutions at all levels and assessment of educational attainment. The Administration of Justice and Democratic Processes are also included, with reference to the use of AI in researching and interpreting facts and the law.

The AI Act places specific importance on products that incorporate AI. In fact, Annex I specifies the sectors considered “high risk” given the importance of human health and safety when AI is used (Deloitte 2024). The sectors covered by Annex I are broad, some of which include Civil Aviation, Agricultural and Forestry vehicles, two or three-wheel vehicles, and Marine Equipment. The list also extends to medical devices, machinery (such as robots) and toys.

High-risk classification is essential for the protection of individuals. However, they are not considered unacceptable for two reasons. The first is that the application of AI in such areas could be socially useful, and the second is the concern that excessive prohibition could hinder the development of the protagonist of the new industrial revolution (Viola & De Biase 2024).

The risk classification system established by the AI Act also includes categories with less stringent requirements. For AI applications with “Limited Risk”, the main requirement is compliance with certain transparency rules. This requirement may be underestimated, but it is one of the most important. Citizens have the right to know whether they are interacting with a human or an AI, and they have the right to be informed whether text, audio or video content has been generated by AI.

The pyramid also includes systems classified as “Other Risks”, but since they are not subject to any specific obligations, there is no need to dwell on them.

This risk-pyramid model seeks to provide a balance between protection and proportionality: imposing stringent requirements only where the level of risk is high, while allowing greater flexibility in contexts where the risk is comparatively limited.

The risk-based approach just analysed was originally designed for Single-Purpose AI systems. However, applying this framework to General-Purpose AI Models (GPAIMs) is complex: their versatile nature and the fact that they can be integrated into very different applications makes it difficult to determine in advance what risk they pose.

In this regard, the AI Act introduces a regime – Chapter V – directed at providers of the so-called *foundation models*, rather than at General-Purpose AI systems as a whole (Deloitte 2024). The reason is that these models constitute the technical core that enables the system to generate outputs from the inputs provided.

General-Purpose AI models are not classified as “high-risk” systems because the model parameters remain fixed after the training phase, making the associated risks more predictable. However, a system built on this model may fall into one of the risk categories established by the regulation.

Here, European legislators have opted for a two-tier classification, distinguishing between models with and without systemic impacts. A General-Purpose AI model is classified as a “high impact” model when it demonstrates a systemic risk through specific technical criteria. In particular, this occurs if the computing power used in its training exceeds the threshold of  $10^{25}$  FLOPs, or if a scientific assessment indicates its significant potential impact (Deloitte 2024).

Moving away from the categorisation of Artificial Intelligence systems, the AI Act delves into the division of responsibilities incumbent upon the various actors involved in the life cycle of an artificial intelligence system. These obligations vary significantly depending on the specific role of the operator and the use case of the system.

The first fundamental step for any entity is to identify its risk category and specific role. The AI Act organises operators into four main categories, each with distinct obligations: the Provider, the Deployer (or User), the Distributor and the Importer (Deloitte 2024). The obligations imposed by law are differentiated according to the role of the stakeholder.

A Provider is a natural or legal person who develops an AI system or a model, or has it developed, and places it on the market or puts it into service under their own name or trademark, whether for payment or free of charge (Deloitte 2024). Providers of “high-risk” AI systems are required to establish a Risk Management System (RMS) covering all stages of the system's life cycle.

This includes “safety by design”, which aims to eliminate or reduce identified risks, and the implementation of appropriate mitigation measures for risks that cannot be eliminated (Deloitte 2024). Furthermore, another fundamental requirement is the establishment of a Quality Management System (QMS) that extends throughout the entire life cycle, covering the pre-market phase, the post-market phase and the ongoing phase (Deloitte 2024).

A Deployer is an individual or entity that uses the AI system under their own authority, excluding use in the context of a non-professional personal activity (Deloitte 2024). They are obliged to apply the instructions for use provided by the Providers, ensure human supervision of the system and continuously monitor the activity of the AI system. For certain categories of systems, they are also responsible for conducting Fundamental Rights Impact Assessments.

The Importer, instead, is the natural or legal person established in the Union who places an AI system bearing the name or trademark of a person established outside the Union on the market (Deloitte 2024).

Finally, the Distributor is any person in the supply chain, other than the Supplier or Importer, who makes the AI system available on the Union market (Deloitte 2024).

The latter two are required to verify that the AI system complies with the requirements and formalities set out in the AI Act, retain compliance certificates for ten years, and take actions such as withdrawing, recalling, or refraining from placing non-compliant systems on the market, as well as cooperating with the competent authorities.

With the drafting of this regulation, more than just a representation of a technical intervention in the functioning of artificial intelligence systems: in some ways, it embodies fundamental European values. With its entry into force on 1<sup>st</sup> August 2024 and its gradual implementation, it can be considered a starting point for a debate that must be addressed in the near future in Europe and the rest of the world.

Many already consider it the first of its kind to strike a balance between protecting fundamental rights and innovation, but it is best to wait before drawing such conclusions. Regulating such a powerful tool is entirely “right” and necessary, both for reasons of development and use, but at the same time, this does not make it immune to possible criticism.

### 3.2 Innovation Dynamics under the AI Act

The introduction of the AI Act represents a moment of regulatory discontinuity with global implications. It shapes the dynamics of innovation and market opportunities for companies operating

in the artificial intelligence ecosystem, both within and beyond European borders, through the establishment of a regulatory framework that aims to promote the adoption of trustworthy AI.

In this context, the discipline dedicated to high-risk systems assumes particular significance. It does not merely introduce compliance obligations, but for companies capable of adapting quickly, it may also turn into a competitive advantage.

One of the most important points is the rules on data governance, outlined in Article 10 (Regulation (EU) 2024/1689). These rules oblige providers of high-risk systems to ensure that the datasets used for training, validation and testing are of high quality and free from bias. The aim is twofold: to ensure the accuracy of the system and to minimise the risk of discrimination. The Act places a strong emphasis on detecting and correcting bias (also provided for in Article 10(5)), obliging companies to invest in dataset audits and the use of mitigation techniques. To encourage such responsible innovation, providers can obtain a *presumption of compliance* with Article 10 if they demonstrate that the data used for training accurately reflects the specific geographical, behavioural, contextual or functional contexts in which the systems are intended to be used (Deloitte 2024). This *presumption of compliance* allows companies to avoid rigorous processes, accelerating time-to-market.

Another element that proactively drives innovation is the transparency requirement – Article 13. Companies are encouraged to develop “explainable AI systems”, as deployers (users) of high-risk systems must provide clear and understandable information on the functionality and limitations of the system (Regulation (EU) 2024/1689, Article 13). This obligation requires companies to provide detailed documentation, enabling users to understand how the models work. For example, in sectors such as financial services, AI-based credit scoring systems must now disclose their decision-making criteria, ensuring fairness and protecting consumer rights.

With the Article 14, high-risk AI systems must be designed and developed in such a way, that they can be effectively overseen by natural persons during the period in which they are in use (Regulation (EU) 2024/1689). This provision obliges suppliers to integrate human-in-the-loop mechanisms into their systems, allowing operators to override erroneous outputs (Okoro 2025). This requirement is vital to prevent possible system failures, as illustrated by several incidents involving Tesla vehicles in the US (Sky News 2020).

Such regulatory burdens provided for high-risk systems are particularly demanding, especially for organisations with limited resources. Precisely to prevent these obligations from becoming an obstacle to innovation, the European legislator has introduced “regulatory sandboxes”, already discussed in the first chapter of the thesis (Regulation (EU) 2024/1689). For those who may have missed it, these are controlled environments that allow start-ups and SMEs to develop and test their

AI systems before they are placed on the market, while reducing the weight of compliance and supporting them in the path toward the requirements, in this case, of the AI Act. The ability to demonstrate compliance with regulatory requirements in this controlled environment translates into a significant competitive advantage, as the exit report, which documents the activities carried out and the results achieved, serves as a *presumption of compliance* (Deloitte 2024).

From a governance perspective, Member States are required to establish “regulatory sandboxes” at national level within 24 months of the Regulation’s entry into force, with a target date of the third quarter of 2026 (Regulation (EU) 2024/1689). Moreover, to ensure that innovation is widely supported, Member States must ensure equal access and equal coverage for participants, and larger states are expected to establish several centres to ensure support at regional or local level.

Each sandbox will have to submit annual reports on the activities, such as best practices, incidents, lessons learned and the set-up of the sandbox to the EU AI Office (Deloitte 2024). This mechanism establishes sandboxes not only as testing grounds, but as an integral part of the EU’s strategy to drive innovation towards the safety and reliability standards required by the regulatory framework.

As regards foreign companies, the AI Act has a significant impact through its extraterritorial scope, giving substance to the “Brussels Effect”. The act applies to suppliers and users of AI systems outside the EU if the output of those systems is used within the Union. This means that any company or organisation offering AI systems that impact individuals in the EU is subject to the regulation, regardless of its geographical location. This jurisdictional extension imposes complex compliance obligations on global technology companies, especially those operating in jurisdictions with divergent regulatory frameworks (Okoro 2025).

In the United States, for example, where the approach has historically been sectoral and fragmented, companies such as Microsoft, Google, and OpenAI must now comply with the more stringent requirements of the AI Act while navigating US federal and state regulations, creating a double compliance burden.

Similarly, companies based in China, whose governance model is state-centric and requires AI models to be aligned with “socialist values” and security reviews, face requirements that could conflict with the EU’s emphasis on transparency and human rights, potentially forcing them to maintain separate AI models for different regions (Okoro 2025).

Even Canada, despite adopting the risk-based Artificial Intelligence and Data Act (AIDA), differs significantly from the EU’s AI Act; AIDA has a narrower scope, focusing on harm prevention and trade, while the AI Act takes a broader horizontal approach.

Not to be underestimated for non-EU companies is the overlap between the AI Act and the General Data Protection Regulation (GDPR). Today, companies that process large amounts of personal data for their AI models must now navigate two layers of European compliance: AI governance and data protection.

The AI Act also requires non-EU suppliers to appoint an authorised representative established in the EU, who is responsible for ensuring that the providers complies with regulatory obligations and maintains the required documentation.

These compliance responsibilities also extend to importers and distributors of AI systems in the EU, requiring them to ensure that their global suppliers meet regulatory standards, adding a layer of due diligence to the supply chain.

AI ACT is redefining the terms of success in innovation: both European and foreign companies must now proactively invest in governance, transparency and risk mitigation, necessarily transforming regulatory requirements into strategic resources.

The declared intention to support innovation – expressed through regulatory sandboxes and *presumptions of compliance* based on harmonised standards and codes of conduct – offers pathways to market access. At the same time, its broad applicability means that European standards on ethics and safety are becoming, in effect, the global benchmark for the next phase of AI development.

In this new European landscape, innovation is inextricably linked to responsibility.

### 3.3 Criticisms of the European Regulation: Ambiguities, Costs, and Risks for Businesses

The European Union’s Artificial Intelligence Regulation, although adopted with the aim of establishing a harmonised regulatory framework and promoting “trustworthy AI”, has been the subject of substantial criticism, which threaten regulatory clarity, impose excessive costs and jeopardise the competitiveness and innovation of European businesses.

The process of adopting the text itself has been described as an “exhausting and protracted affair”, culminating in the approval of the text under pressure from the European Parliament’s June 2024 election deadline (Financial Times 2025). This alleged political urgency has led to compromises that have sacrificed quality in the name of speed, fuelling a critical perception that the European continent favours regulation at the expense of innovation.

One of the main points of criticism lies in the ambiguity and conceptual complexity present in the structure of the Regulation, mainly resulting from the attempt to subject a dynamic and rapidly evolving technology to a regulatory framework designed for static products.

Firstly, the AI Act is based on the New Legislative Framework (NLF), a regulatory system historically applied to static products, such as a simple elevator. However, the dynamic nature of AI, which is characterised by being a constantly evolving process, could create obstacles to the application of hard-coded requirements (Financial Times 2025). Consequently, the *ex-ante* conformity assessment, required for high-risk systems before their placing on the market, clashes with the fundamental uncertainty surrounding the evolution of the system and its potential future risks. In this sense, the attempt to regulate could not only fail to anticipate the dangers, but also lead to regulatory uncertainty within the legal framework.

Secondly, the definition of “AI system” adopted by the AI Act into Article 3(1) (Regulation (EU) 2024/1689). In its wording, which describes a system capable of operating with varying degrees of autonomy and “inferring” from an input to generate an output, it risks equating AI with generic software. Recital 6 attempts to narrow this scope by clarifying that the definition should not cover simple rule-based software and that “inference” must go beyond basic data processing. Thus written, it ensures that a simple software, such as the auto-sum function in Excel, does not qualify as AI. As can be seen, this clarification reduces uncertainty partially: the decisive criterion ends up being reduced to the single term “inference” (Hacker 2023).

Publicly, the issue of ambiguity exploded with the advent of General-Purpose AI Models, such as ChatGPT, whose public introduction in late 2022, forced legislators to “rush to rewrite the text” (Financial Times 2025). Originally, the Commission’s draft contained no reference to large language models, causing a complex process of forced adaptation, defined as “shoehorning”, which led to the inclusion of rules for GPAIMs, which, as we have seen in the previous paragraphs, represent a new category distinct from traditional high-risk AI systems (Financial Times 2025).

The introduction of the “systemic risk” category for these models is one of the focal points of academic criticism, which argues that the concept creates more uncertainty than it resolves. “Systemic risk”, in Art. 3 (65), is described as a specific risk linked to the “high-impact capabilities” of general-purpose AI models, which manifests itself through negative effects that can spread widely throughout the entire AI value chain, affecting health, safety, fundamental rights and society as a whole (Regulation (EU) 2024/1689).

Sociological risk analysis suggests that the category of systemic risk does not operate as a strict technical threshold, but as a proxy indicator of the epistemic uncertainty surrounding the capabilities and effects of general-purpose AI models. This means that legislators use systemic risk to fill a knowledge gap, rather than to define an actually measurable risk. According to several scholars, this choice ultimately institutionalises regulatory and scientific uncertainty rather than reducing it, as it incorporates uncertainty into the very structure of regulation (Carey 2025).

This hasty approach is evident in the adoption of a fixed computational threshold of  $10^{25}$  FLOPs for General-Purpose AI Models as a criterion for classifying “systemic risk” (Regulation (EU) 2024/1689, Article 51 (2)). This threshold has been criticised for being too high. Models trained with less computational effort can also pose systemic risks. For example, a model trained using less than  $10^{25}$  FLOPs could still pose a systemic risk if integrated into critical infrastructure, while other models with higher computational effort may pose no systemic risk at all (Carey 2025).

The legislative function, in itself, embraces the principles of legal certainty and normative predictability: it must be able to allow the addressees of the legal provision to clearly foresee which behaviours, if carried out, will result in a violation of the provision. Based on this premise, by analysing the relevant regulation, it appears that such regulation reflects the technical knowledge available at the time when it was drafted, based on studies which, although current at that time, today prove not to correspond to the developments that the field has undergone. This therefore results in a complete instability of the normative source: the continuous variation of these thresholds deprives the law of its primary functions, thus producing a de-temporalisation of the normative effect, which, rather than – as the general rule requires – providing for the future (since the law regulates for the future), ends up corresponding only to past situations.

This instability operates on two levels. First, the thresholds themselves remain provisional and dependent on future observations. Second, any future observation will inevitably be retrospective, since the law cannot evolve at the same pace as technological change. The law thus remains anchored in the past. Moreover, the deferral of the provisions to future ad hoc assessments does nothing but reinforce this instability and uncertainty, being a provision that is too broad and general – not in line with the principle of certainty.

With such an open provision, which represents a “regulatory void”, there is a risk of giving excessive room to the arbitrariness of the authorities responsible for making assessments in the matter, thereby eliminating the stable point of reference that an effective normative source should provide.

Even Gabriele Mazzini, an architect of the act, now concedes that it is too broad, complex and “doesn’t provide the legal certainty that is needed” (Financial Times 2025).

The conceptual weaknesses mentioned above are compounded by the second major criticism: compliance costs. The costs of complying with the AI Act are expected to be substantial, especially for small and medium-sized enterprises and start-ups (Hacker 2023). While well-funded companies may navigate these financial demands more easily, smaller companies, which often lack sufficient financial and human resources, will encounter serious difficulties in meeting the requirements. There are concerns that these burdens will delay or slow down the adoption of AI and its development.

The European Union’s response, in defence of start-up innovation, lies in the introduction of sandboxes. But even these are not immune to criticism. According to Article 57(7) of the AI ACT, the maximum result that a participant can achieve is a sandbox exit report, which summarises the activities carried out and the results (Regulation (EU) 2024/1689). Although this report must be taken into positive consideration by the supervisory authorities, this does not equate to granting a *presumption of compliance*. The competent authority managing the sandbox is not authorised to issue a certification guaranteeing a *presumption of compliance*. The inability to authorise this creates a real risk of duplication of effort and delays in go-to-market, as innovators must still undergo the required assessment procedure (Ahern 2025).

Furthermore, innovators are motivated by considerable confidentiality concerns, tending to jealously guard their intellectual property. They may be reluctant to participate in an environment where they may be required to disclose details about algorithms, trade secrets and other confidential information to the relevant authorities, especially when the latter are also potentially involved in supervising their competitors, although the authorities are subject to confidentiality obligations under Art.78 (Ahern 2025).

Sandboxes will probably not be decisive. Successful start-ups do not need them at present and will probably not really need them later either (Hacker 2023).

The intersection between ambiguity and costs necessarily leads to the third critical point: the risk to the competitiveness and innovation of the European ecosystem. Alexandru Voica, head of corporate affairs at London-based AI start-up Synthesia, says: “A lot of these scale-ups will never be able to reach the same sort of size and impact as American or Chinese companies, because they are buried under all of this regulation” (Financial Times 2025).

Even large companies are critical of such regulation. Tech companies such as Meta have argued that the bloc is cutting itself off from accessing cutting-edge services, because companies delay or avoid rolling out certain AI features out of fear of non-compliance with the AI Act (Financial Times 2025). In July 2025, large companies including Airbus, BNP Paribas, Mercedes-Benz, and TotalEnergies, urged the European Commission to halt the AI Act’s timeline for two years to simplify the rules and allow time for companies to implement them (Financial Times 2025).

Taken together, these critical issues represent a structural tension at the heart of the AI Act. A regulation that aims to protect EU citizens and promote “trustworthy AI” paradoxically risks hindering the conditions that make technological progress possible. The combination of ambiguity, regulatory instability and compliance burdens risks creating a regulatory environment that discourages investment and innovation.

Although the AI Act, according to the Commission, aims to position Europe as a global leader in ethical and trustworthy AI, its architecture is considered by many to be ill-suited to the speed and unpredictability of AI.

### 3.4 Philosophical and Ethical Critiques of the AI Act

The European AI Act has been the subject of some philosophical criticism.

One criticism in particular, which is worth highlighting, focuses on the arbitrary incorporation of what is referred to as “the speed paradigm”, a concept that manifests itself both in the regulatory process and in the very vision of AI technology. The theory highlights how the Regulation assumes a priori that an “intrinsically fast” phenomenon such as AI requires synchronised governance.

The presumed need for speed does not arise from ethical reflection, but is the result of market and political forces, causing potential regulatory inefficiency. The paradigm establishes a self-sustaining cycle in which the speed of technology serves to justify the speed of regulation, and, conversely, regulation facilitates the speed of AI diffusion in the market, closing a loop that is deeply problematic from an ethical perspective (Anderson 2022).

The claim that AI development is rapid in itself, and therefore rapid regulation is justified, is philosophically debatable. The very idea of “intelligence” is not universally equated with speed; although intelligence can be associated with quick-wittedness, it is equally, and often, linked to a slower, more thoughtful process of gathering and processing information. In fact, the concept of wisdom – which can be considered an “acceptable synonym” for intelligence – is typically associated with the term reflection, without any connotation of speed.

Considering technological development, it is incorrect to view AI as a rapid process. The work involved in achieving modern AI took around 60 to 70 years, which is even slower than some significant human inventions. It is wrong to confuse the recent rapid spread of this technology with its intrinsic evolutionary speed. Even its capacity for self-evolution is not rapid in itself, since it depends on data collection, although this may give an illusion of speed. AI self-evolves, but “it does not self-evolve swiftly, all things considered” (Anderson 2022).

Assuming that these assumptions are correct, it follows that the speed with which legislators drafted the regulations was a choice.

According to Padre Paolo Benanti, in a meeting of the AI Revolution series in 2024 with Sole24, this regulatory text strikes a balance between innovation and the protection of fundamental rights, integrating, albeit not explicitly, the concept of algorithmic ethics (Maestri 2024).

However, although ethics is understood as a set of “right” values and norms that regulate human behaviour or protect humans, ethics is also understood as the practice of reflective interpretations that emerge from the elaboration of examples – ethical or unethical – revealed in lived experience. Ethics as a practice requires slow time, patience, reflection and retrospection.

This haste contrasts with various ethical currents. For instance, it conflicts with the *ethics of discourse*, which requires a “slow process of continuous public questioning and answering” (Anderson 2022). Only through patient dialogue is it possible to build coherent ethical foundations, something that regulatory haste prevents, as we saw in the endless negotiations between 6<sup>th</sup> and 9<sup>th</sup> December 2023.

Paradoxically, Big Techs began to suddenly promote initiatives related to AI ethics. This phenomenon has often been described by critics as a mere “marketing exercise” (Anderson 2022). The aim of this strategy, in fact, was not so much to address moral dilemmas as to demonstrate to the public and policy makers that they were taking ethics “seriously”, in an attempt to persuade them that external legislative intervention could be avoided.

In this context, it is not surprising that there has been a desire to adopt the logic of the New Legislative Framework (NLF) in the AI Act. Treating the protection of concepts such as fundamental rights with the same compliance mechanisms used for products risks reducing ethics to a mere bureaucratic checklist (Veale & Borgesius 2021).

Recognising the “paradigm of speed” within the AI Act implies acknowledging that such regulation is not based on preparatory ethics (Anderson 2022). Ethics shares the temperament of philosophy in being patient and cannot be distorted to suit the temperament of technological rapidity.

The regulatory rush to regulate something we do not yet fully understand risks degrading ethics to a mere facade of respectability.

Luciano Floridi (2021), in his article “The European Legislation on AI: a Brief Analysis of its Philosophical Approach”, also points out some limitations of European legislation.

Firstly, he criticises the adoption of an anachronistic terminology, in particular the definition of the approach as “human-centred”. Although it is conceptually obvious that technology should be useful to humanity, this label is considered both trivial and ambiguous. The term “human-centred” seems synonymous with “anthropocentric”. It is obvious that any technology, AI included, must be at the service of humanity, but one must also consider the environment as crucially important (Floridi 2021, 218). Fortunately, despite the unfortunate and obsolete terminology, the underlying vision is sound: the AIA at that time – referred to the Proposal for a Regulation – and subsequently the AI Act emphasises the value of AI as a technology that can be very “green” and provide extraordinary support

against pollution and climate change and for the sustainable development of information societies (Floridi 2021; Cowls et al., 2021).

From a regulatory point of view, the text is considered vague. A concept such as the definition of high-risk systems leaves room for uncertainty or possible loopholes. There is conceptual confusion in distinguishing between high-risk systems because they must function correctly and high-risk systems, considered as such because their unethical use can cause problems. Failure to distinguish between these two meanings confuses the resilience required for “good” AI with the resistance that must be opposed to “bad” AI (Floridi 2021, 219). Similarly, the text is vague in prohibiting the use of AI systems designed to distort human behaviour with probable physical or psychological harm; although the intent is commendable, such an application could risk prohibiting even non-problematic systems (Floridi 2021, 219).

Furthermore, some expectations are considered overly idealistic; for example, the requirement that the datasets used for training must be “sufficiently relevant, representative and error-free” describes desirable characteristics that are almost impossible to achieve in practice (Floridi 2021, 219).

The convergence of the ethical issues discussed above, and the regulatory and economic challenges analysed previously gives rise to an indirect risk: far from benefiting the European ecosystem, the new legislation could inadvertently encourage the relocation of AI research and development to jurisdictions outside the EU.

### 3.5 Transparency, Accountability, Sustainability: Principles for a European AI

The European Union has outlined its own strategy for the future of Artificial Intelligence. The primary objective, as has been emphasised many times, is not merely technological, but value-based: to promote the development of AI that is ethical and secure, and, if possible, cutting-edge.

The instrument representing this vision was the *High-Level Expert Group on AI* (AI HLEG), an independent body that published the Ethics Guidelines for Trustworthy AI in 2019. At the heart of the project is the concept of “Trustworthy AI”, considered a necessary condition for society to be able to live confidently with new technological systems. Trustworthiness is considered to be the result of the convergence of three essential components that must exist throughout the entire life cycle of the technology: lawfulness, as regulatory compliance; ethics, understood as shared values; and robustness, both technical and social, necessary to prevent unintended harm (AI HLEG 2019, 4).

As mentioned above, this approach stands out for being human-centric. It does not put technology first, but rather the fundamental rights enshrined in the EU Charter. From this ethical foundation arise four cardinal principles that every AI system is required to respect: respect for human autonomy, prevention of harm, fairness and explicability.

However, putting these principles into practice is not without its complexities. In this context, the requirements of Transparency, Accountability and Sustainability are essential tools for ensuring socially “fair” innovation.

In the context of “Trustworthy AI”, Transparency is a necessary condition for building trust between humans and machines. It encompasses everything from the origin of the data to the final interaction (output), through three interdependent dimensions: traceability, explainability and communication.

The first level is understood as the ability to report the “history” of the system. For AI to be verifiable, it is necessary to document, in detail, the data sets and algorithms used. This documentation is required to ensure auditability; in fact, only by tracing the decision-making process is it possible to identify any present or future errors or biases.

The second dimension is explainability, which responds to the need to make AI system processes understandable. On a technical level, this means that decisions must be clear and understandable to humans, while at the same time guaranteeing the right to receive information appropriate to the level of competence of the interlocutor. However, the use of complex models, such as *Deep Learning*, raises the issue of opacity, also known as “Black Box”, where a causal explanation of an output is not always possible. In such circumstances, if total interpretability is not achievable, it must be compensated for by greater traceability and communication, while ensuring that fundamental rights are always respected (AI HLEG 2019).

Finally, Transparency is achieved through communication. It must be seen as a principle of honesty; AI systems must be identified as such, ensuring that humans are always aware of who they are interacting with. This is crucial to prevent manipulation and deception. Furthermore, it is necessary to openly communicate the capabilities and limitations of the system to its stakeholders.

The requirement of Accountability, in some ways, complements the other elements of “European” AI. Closely linked to the principle of fairness, it requires a guarantee of the right to redress. Accountability requires the establishment of clear mechanisms to ensure that responsibility for AI systems and their outcomes is assigned throughout all stages of the life cycle (AI HLEG 2019).

A key element, previously spoiled, is auditability. It is somewhat like the natural continuation of transparency: if traceability documents the “history”, auditability allows it to be evaluated. For specific high-risk applications, systems must be subject to independent audits.

Accountability also includes the obligation to minimise and report negative impacts, which requires identifying, assessing, documenting and minimising the potential negative impacts of AI systems, with particular attention to those who are directly or indirectly affected (AI HLEG 2019). The Guidelines encouraged the use of impact assessments such as red teaming – an independent group simulates an adversary role to find security vulnerabilities and ethical flaws – or forms of Algorithmic Impact Assessment, to be carried out before and during the use of the system. In fact, according to Article 27, before putting a high-risk system into operation, certain deployers must specifically assess the impact of that system on fundamental rights.

In this context, the most important aspect is the ability to provide redress. There can be no talk of responsibility unless there are accessible mechanisms in place to ensure adequate compensation for damage or negative impact. The knowledge that compensation is possible is a necessary condition for ensuring trust. This translates into the New Product Liability Directive (PLD), which Member States must transpose by 9 December 2026, and which will apply to products placed on the market or put into service after that date (Giangiacomo & Lorenzi 2024).

However, while the AI Act introduces *ex-ante* protection systems, the withdrawal of the proposed “AI Liability Directive” leaves a new regulatory framework, such as human negligence in the use of the system. Without this directive, the “presumption of causality” is lacking. Consequently, in order to obtain compensation for the negligent use of AI, European citizens must rely on national laws, which are often not in step with new technologies.

Finally, the last requirement for analysing “Trustworthy AI” is Sustainability. This is based on the idea that within society, stakeholders, such as the environment or individuals, must be considered an active part throughout the entire life cycle of an AI system. From this perspective, “Trustworthy AI” is attributed a role of collective prosperity and well-being, facilitating the achievement of Sustainable Development Goals (SDGs) (AI HLEG 2019).

Sustainability manifests itself through the need for AI that is sustainable and environmentally friendly. It is no coincidence that the training process is assessed for its environmental impact, through an examination of resource and energy consumption.

Moreover, the social dimension is also addressed. With the present of this requirement, it has been represented the awareness that exposure to AI systems can alter social relationships and collective well-being.

The impact on society and democracy is another critical component, requiring AI to promote democratic processes and respect for pluralism of values. The Guidelines has raised critical concerns, including: the efficient use for the identification and tracking of individuals (such as facial recognition), which raises serious legal and ethical concerns and requires that the application of such

technologies be clearly justified in existing law, distinguishing between identification and mass tracking/surveillance; Covert AI systems, where the user is unaware that they are interacting with a machine, a practice that should be avoided; and the practice of citizen scoring (large-scale normative assessment of moral character), which is dangerous as it can lead to a loss of autonomy and threaten the principle of non-discrimination (AI HLEG 2019, 34).

The joint implementation of the requirements of Transparency, Accountability and Sustainability is characterised by a continuous approach, assessed throughout the entire AI life cycle, as represented in Article 9, which states: “The risk management system shall be understood as a continuous iterative process planned and run throughout the entire lifecycle of a high-risk AI system, requiring regular systematic review and updating” (Regulation (EU) 2024/1689).

Although the European regulatory framework is not without its critics, it should be recognised that it is based on the ambition to build truly trustworthy systems within its own market. This is to its credit. Trust is considered a fundamental condition for enabling individuals to fully reap the benefits of new technology.

The European regulatory approach, supported by a robust framework that inherits the ethical foundation of the HLEG, aims to ensure that the development of AI is aligned with the fundamental values of respect for human rights, democracy and the rule of law.

The emphasis placed on these principles is not accidental. They represent the mechanism by which Europe, drawing on its tradition of placing people at the centre of its initiatives, wants to position itself as a global leader. The vision is to promote AI that serves the common good and human flourishing, ensuring collective as well as individual well-being (AI HLEG 2019).

# **CHAPTER 4**

## **VOICES FROM THE ECOSYSTEM: PERCEPTIONS AND IMPACTS OF THE AI ACT.**

### 4.1 Research Question

This research analyses the perceptions of key players in the European ecosystem on the regulation of artificial intelligence, with a particular focus on the AI Act, in order to understand whether and to what extent regulatory initiatives can balance protection and innovation. The research question can be formulated as follows: does the AI Act succeed in balancing the need for protection with the need to preserve and stimulate innovative capacity in Europe? Based on this question, the study aims to (i) reconstruct the main perceived critical issues and (ii) identify areas for possible improvement.

### 4.2 Research Design and Methodology

The approach is inspired by a qualitative research approach based on grounded theory (Glaser & Strauss, 1967). The interviews were conducted following a common outline, with room for adaptation and further exploration depending on the profile of the interviewee. This type of approach was considered the best in relation to the research objectives and the nature of the topic, which is still relatively unexplored in the literature.

This approach made it possible to generate novel theory, based on the reconstruction of the experience of individual directly involved in the phenomenon under analysis and to draw interpretations and assessments that different actors attribute to the AI Act and its effects. The analysis followed a rigorous research protocol, which is summarised below.

#### 4.2.1 Sample selection

The sample selection was based on a key informant approach, to include professionals from complementary fields, such as investment and innovation, institutional relations in multinational tech companies, and legal expertise on AI regulation. These individuals are considered capable of contributing new knowledge to the topic under analysis. The choice of this hybrid sample is due to

the desire to bring out different perspectives. Moreover, these individuals were identified through professional and personal connections. As shows Table 1, the sample has been categorized by using code, indicating role and sector:

Code	Role	Sector / context
INT1	Venture Capital Fund Manager	Italian VC operating across Italian and European markets; Deep-Tech focus
INT2	Head of Institutional/Government Affairs in EU	U.S. based multinational tech company
INT3	Professor and Legal Consultant	AI law and regulation in Europe
INT4	Head of Institutional/Government Affairs in EU	U.S. based multinational tech company (U.S. interviewee)

Table 1 – Sample

#### 4.2.2 Data Collection and Analysis

The interviews were conducted between November 2025 and January 2026 via video call, with an average duration of 45 minutes per interview. The conversations were collected in transcripts and research notes; one was conducted in bilingual mode (Italian/English). As indicated in Table 2, the interview guide combined a core set of research-question-driven prompts with role-specific probes. This flexibility is consistent with a semi-structured design and allowed the interviewer to explore issues emerging from each interviewee’s domain expertise (VC, public affairs, AI law/regulation):

Topic	Core questions (asked across interviews)	Tailored probes (examples, depending on interviewee)
Role of the EU / positioning	How do you interpret the EU’s decision to regulate AI now – does it imply a “regulator/observer” role in the global AI race?	<b>PA/US tech:</b> How is the AI Act perceived in Washington (global standard vs competitiveness drag)? <b>Legal:</b> Is it a geopolitical positioning strategy?
Balance innovation vs protection (RQ)	What are the main risks for Europe if the AI Act fails to balance innovation and protection? Is such a balance achievable under the current framework?	<b>VC:</b> Risk of talent/startup flight; slower go-to-market. <b>PA:</b> Provider relocation decisions due to guidelines/compliance burden.

Impact on competition / Big Tech	Could the AI Act unintentionally consolidate Big Tech dominance rather than foster EU competition and innovation?	<b>Legal:</b> How to avoid effects that penalize SMEs (e.g., “Brussels Effect” concerns)? <b>PA:</b> Perceived lobbying pressure / strategic effects.
Timing and regulatory uncertainty	Is regulating “early” a prudent choice, or does it risk regulating an uncertain future?	<b>Legal:</b> Pacing problem – can legislative techniques address the innovation/regulation speed gap? <b>PA:</b> Is regulation driven by fear/public trust dynamics?
High-risk domain governance	Is the AI Act overly restrictive in high-risk areas where AI is needed (e.g., healthcare, education)?	<b>Legal/PA:</b> How should high-risk obligations be calibrated to avoid blocking beneficial use-cases?
Tools for responsible innovation	Are regulatory sandboxes effective in enabling responsible innovation?	<b>PA/Legal:</b> Are sandboxes clear enough to be usable? Too many instruments?
Market outlook and EU capacity	How do you see the EU AI market evolving toward 2030?	<b>VC:</b> Investment trends; execution gap; deep-tech scaling. <b>PA:</b> Market attractiveness vs bureaucracy.
Policy and ecosystem measures	What instruments should the EU/Member States adopt to better support the innovation ecosystem under the AI Act?	<b>VC:</b> Changes in AI due diligence; resources for startup compliance. <b>Legal:</b> Co-regulation/public-private partnership routes.
Ethics and legitimacy	Can “ethics” realistically be a competitive advantage for Europe, or is it disconnected from market realities?	<b>PA:</b> Narrative critique; <b>Legal:</b> operationalization of principles into enforceable rules.
Final improvements	If you could change one element of the AI Act, what would you change?	<b>Legal:</b> Future-proofing vs adaptability; technical thresholds (e.g., compute-based) sustainability. <b>PA:</b> competitiveness and feasibility.

Table 2 – Questions and Topics

The material was analysed using manual thematic coding, informed by the Gioia Methodology (Gioia et al. 2013). In the first phase, the transcripts and notes were read systematically, assigning descriptive labels close to the language used by the interviewees. In the second phase, these labels were progressively grouped into interpretative categories, until a coherent set of themes was formulated. The results are presented in section 4.3 through a thematic structure, accompanied by selected quotations to support the interpretations.

To strengthen the robustness of the analysis, the coding was conducted on the basis of transcripts and notes, maintaining textual traceability. The hybrid composition allowed for the formulation of interpretations from different perspectives on the same object (AI Act), reducing the risk of single-perspective readings. As shows Table 3, first-order concepts were manually coded from transcripts/notes and progressively grouped into intermediate categories (second-order themes), which were consolidated into aggregate dimensions:

<b>First-order concepts (illustrative)</b>	<b>Intermediate categories (Second-order themes)</b>	<b>Aggregate dimensions</b>
“Regulation as a fundamental protection tool”; “responsibility principles”; “risk principles”; “product safety”; “prevent rather than cure”	Protective rationale and risk governance	<i>Regulatory design and implementation (processes and modalities)</i>
“Slow regulation / chasing innovation”; “too many interpretations”; “vague definitions”; “too many regulatory layers”; “unclear division of competences across Member States”; “need for co-regulation (multi-stakeholder)”	Implementation frictions and interpretive uncertainty	<i>Regulatory design and implementation (processes and modalities)</i>
“Strong in R&D but weak in execution”; “turn technology into products and services where Europe has an advantage”; “data as an asset to leverage”	Execution gap and sectoral commercialization	<i>EU market conditions and investment capacity (competitiveness and scaling)</i>
“Unrealistic to expect an EU AI big player”; “compliance costs hit SMEs and startups”; “regulatory sandboxes are unclear / too many”	Scaling constraints and asymmetric compliance burden	<i>EU market conditions and investment capacity (competitiveness and scaling)</i>
“Insufficient investment / loss of innovation”; “need more investment in the ecosystem”; “greater flexibility to attract foreign VC”	Capital scarcity and investment attraction	<i>EU market conditions and investment capacity (competitiveness and scaling)</i>
“People are always at the core”; “protecting citizens”; “ethical principles should guide AI”; “revolutionary technologies require ethical safeguards”	Anthropocentric legitimacy and ethical justification	<i>Ethics (anthropocentric justification and contested framing)</i>
“Translating principles into rules is delicate”; “ethics as a (potential) advantage”; “ethics used as a crutch / weakness narrative”; “timing concerns (too early)”	Operationalization tensions and contested meaning	<i>Ethics (anthropocentric justification and contested framing)</i>
“US approach is more pragmatic”; “the state focuses on national security; the rest is left to the market”; “business comes first”	Market primacy and regulatory pragmatism	<i>Benchmarking the US model (pragmatism, business primacy, geopolitical framing)</i>
“High trust between public and private actors”; “deregulation to support innovation”; “technological leadership shapes geopolitics”; “Europe regulates while others innovate”	Public–private relations and geopolitical narrative	<i>Benchmarking the US model (pragmatism, business primacy, geopolitical framing)</i>

Table 3 – Gioia data structure

### 4.3 Results

This section presents the results of the qualitative analysis of the interviews, with the aim of answering the research question regarding the AI Act’s ability to balance innovation and protection.

In particular, the results highlight how respondents assess this balance, not only in terms of the guiding principles of European regulation, but also in terms of the institutional and economic conditions that determine its applicability and effects on the internal European AI ecosystem.

As will be seen, the analysis reveals a structure divided into four themes, selected because (i) they reflect the sample and (ii) they are useful for reconstructing different dimensions of the trade-off between protection and innovation. The first concerns the regulatory architecture of the AI Act, with a particular focus on processes, methods and the *ex-ante* approach. The second theme integrates the dimension of the EU market and investment, two elements that are closely interdependent in determining competitive sustainability. The third theme analyses the role of ethics, both as an anthropocentric foundation of regulation and as an area of interpretative tension. The fourth and final theme analyses the comparison with the United States as a benchmark through which respondents evaluate the concepts of priorities, regulatory approach and strategic effects.

#### 4.3.1 Regulatory Design and Implementation

In the dataset, AI regulation is first and foremost recognised as a necessary protective measure. This is an important legitimising framework, clarifying that the opposition is not to regulation itself, but rather to the way in which it has been designed and implemented in the European context.

A first critical issue concerns Europe's trajectory of "always regulating before innovating", which respondents consider to be a distinctive feature of the approach. The timing issue is not only chronological but also strategic point: if regulation systematically precedes the maturation of the industrial ecosystem, there is an increased risk that the regulation will become a constraint that stifles the very conditions for innovation: "*Regulation is necessary, but I don't know if this is the right time or the right way...*" (INT1).

Paradoxically, this rush to regulate has led to the drafting of regulations that are unable to keep pace with technological progress: "*Regulation is perceived as slow and trying to catch up with innovation*" (INT2).

The most conceptually relevant aspect that emerged from theme 1 concerns the *ex-ante* approach adopted. The principle of protection is considered consistent, but it becomes problematic when it translates into an overly ambitious attempt to "anticipate the future". The "good intentions" are recognised, but it is noted that an approach has been attempted that is "*somewhat unrealistic in terms of predicting the future, trying to predict what is actually unpredictable*" (INT3).

This has two implications: (i) *ex-ante* “limits concrete cases”, thus risking making adaptation to real contexts of use more rigid; (ii) future-proofing, if built on broad definitions, risks becoming vague and unclear.

This ambition has led to the drafting of broad and unclear definitions, increasing interpretative discretion and making their application uncertain. When regulations leave too much room for interpretation, there is a risk that protection will become uneven and that compliance will depend on interpretation: “*You can always innovate as long as the rules are clear*” (INT2).

Added to this vagueness is the perception of stratification (too many levels) and the risk of confusion of competences in Member State offices. Here, the critical issue concerns the transition from text to implementation: “*a multi-level framework, if excessively stratified and unclear, can reduce the effectiveness of protection and increase indirect costs for regulated actors*” (INT2).

For some respondents, this combination of critical issues leads to a clear assessment of the balance: the approach is considered unbalanced between innovation and protection and would require a methodological correction. In particular, there is a proposal for *co-regulation* approach, i.e. structured collaboration between the public and private sectors: “*the approach is considered unbalanced in terms of innovation and protection... [there is a need for] co-regulation (multi-stakeholder)*” (INT3).

Finally, the research reveals a more radical perspective, shifting the focus to the relationship between regulation and industrial capacity. One of the factors that led to the “failure” of the European Union stems from having embarked on the legislative process before generative technology existed and without adequate technical expertise, with competitive effects that penalise start-ups most, and therefore the possibility of a new European tech big player. From this perspective, regulation becomes a brake and may even be functional for incumbents, because a complex and uncertain framework raises barriers for small players, consolidating dominant positions. This gives rise to the (highly political) idea that regulatory pressure can become a substitute for market competition: “*Those who have nothing to sell are the ones who regulate*” (INT4).

#### 4.3.2 EU Market Conditions and Investment Capacity

The European AI market is described as a context in which regulatory issues are intertwined with a structural capacity to “ground” innovation. In particular, Europe is recognised as relatively strong in terms of R&D, but weaker in terms of implementing and transforming technology into

competitive products and services: *“In Italy, as in the EU... we are very good at R&D... However, in the EU, as in Italy, there is a lack of execution capacity”* (INT1).

The proposed trajectory is a strategy to enhance sectors in which Europe has competitive advantages: *“We need to know how to exploit technology in products and services where we still have a competitive advantage”* (INT1); *“We can recover by investing in sectors where we have advantages”* (INT2).

At the same time, there is a very clear opinion about the likelihood of a major European AI player emerging. Here, the argument is not about regulation but market structure: the ecosystem is perceived as insufficient to generate a European technological champion comparable to the major global players: *“It is illusory to imagine a major AI player in the EU at this point”* (INT2).

In this context, respondents report a distributional problem linked to compliance: in an already fragile market, compliance costs are perceived as particularly burdensome for start-ups and SMEs, while they are more sustainable for large companies. This leads to the conclusion that regulation, while legitimate in its aims, can produce selective effects that penalise those who attempt to innovate, pushing some players to move towards other markets: *“With this regulation, compliance costs are high, especially for those who want to innovate; SMEs and start-ups are the most affected”* (INT2).

In light of this, investments take on an enabling role. The respondent argues that the EU ecosystem should be supported with investment policies: *“Instead of focusing too much on regulation... [we need to] help the ecosystem through greater investment in the sector. The lack of investment... [is] representative of a loss of innovation in the AI ecosystem”* (INT1).

The sample shows a link between investment and institutional attractiveness: Europe is described as a market where little investment is made and where bureaucracy and rigidity reduce its attractiveness to foreign capital; greater flexibility is cited as a prerequisite for attracting venture capital, in a context perceived as lacking in domestic VCs: *“There is little investment in Europe and bureaucracy makes the market less attractive for foreign investment... there is a need for flexibility in the EU”* (INT4).

#### 4.3.3 Ethics

In the sample, ethics is configured as the anthropocentric foundation of regulation; the legitimacy of recourse to ethical principles derives from the centrality of the person and from the assumption that the ultimate goal of regulatory intervention is to protect citizens from the risks potentially associated with the adoption and use of artificial intelligence. In this interpretation, ethics

is not a rhetorical addition, but the framework that justifies the very existence of constraints and responsibilities: *“Ethics is right because people are always at the heart of it; it is good to address the issue in order to best protect citizens”* (INT1).

Secondly, ethics is also interpreted as necessary because AI is perceived as a “revolutionary” technology. For some respondents – INT1, INT2 and INT3 – the integration of ethical principles into high-impact systems appears to be a necessary step precisely because of the transformative scope of the technology: *“It is right to incorporate ethical principles into such revolutionary technologies”* (INT2).

However, the dataset also reveals a tension in terms of implementation; ethics are considered acceptable at the level of principles, but translating them into operational rules is described as a delicate process. The argument here is typically institutional: the problem is not whether ethics are desirable, but how to transform general principles into applicable criteria without producing ambiguity or distorting effects. In line with this, there is also the idea that ethics can be a positional resource, but only on condition that the transition from principles to rules is managed carefully: *“Ethics [is] right, but the transition from principles to rules [is] very delicate”* (INT3).

A further point concerns the (more controversial) issue of ethics as a competitive advantage. Two interpretations coexist in the sample: on the one hand, the idea that ethics can become a reputational/strategic asset (especially in the European context); on the other, a criticism that challenges the very association between ethics and competitive terminology, interpreting it as ex post rationalisation or as a “moral anchor” in the absence of market power. This divide is useful empirically because it shows that ethics, in the interviewees’ discourse, is not just a value, it is also a terrain of interpretative conflict over Europe’s identity and positioning: *“Ethics can be an advantage”* (INT3); *“Ethics? ...people cling to ethics out of weakness...”* (INT4).

#### 4.3.4 Benchmarking the US model

The reference to the United States serves mainly as a comparative term for assessing the sustainability of the European approach: rather than providing an exhaustive description of the US model, respondents refer to it to highlight differences in priorities between protection, innovation and industrial strategy. From this perspective, the transatlantic comparison becomes a useful interpretative tool for understanding the balance of the AI Act: *“In the US... priority is given to business...”* (INT4).

A particular difference concerns the perception of a more pragmatic US approach that relies more heavily on market dynamics. The US is described as a context in which public intervention

focuses on areas considered essential (national security), while market logic prevails in other areas; this is interpreted as an element that favours rapid innovation, while potentially exposing it to greater risks in terms of protection: *“In the US, the state mainly takes care of national security, the rest is left to the market... [a more pragmatic approach]”*(INT3).

Furthermore, returning to the concept of co-regulation, it is emphasised that in the US context there is a greater perception of trust and cooperation with the private sector: *“In the US, there is great trust between the private and public sectors”* (INT3).

Finally, a reading oriented towards competitive and geopolitical factors emerges, highlighting how the US adopts deregulation approaches to support innovation and consolidate global tech leadership, unlike Europeans, whose goal – for some – is to consolidate regulation as a global standard: *“In the US, they deregulate to help innovation... today, those who are ahead determine innovation and its geopolitical implications”* (INT4); *“In the end, they regulate, we (the US) innovate, and the Chinese copy”* (INT4); *“The EU wants to be a global standard because it has neither products nor investments”* (INT4).

#### 4.4 Considerations

All of these observations point to a clear conclusion: although the AI Act was created with the laudable intention of promoting a reliable and responsible ecosystem, it risks undermining the very technological competitiveness that the European Union claims to want to protect. The concerns raised show how excessive regulation, far from guaranteeing certainty and trust, can translate into a structural brake on the spread of innovation. The absence of advanced models in the European market, the continent’s declining attractiveness to investors and the growing distance from the United States are not marginal side effects, but indicators of a trend already underway: where regulation takes precedence over incentives for innovation, the rational choice for economic operators is to move their research, development and implementation elsewhere.

In this sense, the most profound risk of the AI Act is not that it regulates too much – regulation is, in itself, necessary and inevitable – but that it regulates in such a way as to make Europe an increasingly unfavourable environment for the adoption of emerging technologies. A continent that declares its ambition for technological sovereignty risks creating, with its own hands, the conditions for its own dependence. Unless a balance is struck between protection, flexibility and growth – not blind deregulation, but regulation that accompanies innovation rather than precedes it – the AI Act

could represent not Europe's turning point in the global race for artificial intelligence, but confirmation of its competitive marginality.

In this race and geopolitical paradigm, even if regulation can become a global standard, merely being referees is not enough.

# CHAPTER 5

## *Graduated Compliance Model*

### 5.1 The Costs of Regulation

The research conducted in drafting the work on artificial intelligence regulation, with a particular focus on the AI Act, has yielded predictable results given the current atmosphere in Brussels. The research question – *Does the AI Act succeed in balancing protection and innovation?* – starts from a fundamental assumption of neo-institutionalist literature: organisations are the result of adaptation to the institutional environment and the constraints imposed by it (Boccardelli & Brunetta 2014, 14). Consequently, it is impossible not to think that regulation of this calibre does not influence the innovative environment of the European ecosystem. As Boccardelli and Brunetta point out (2014), the economic performance of companies depends on how they interpret and react to the institutional environment, in this case the AI Act.

In light of the results of Chapter 4, the issues that emerged represent the critical points of the current AI regulation in the EU, as well as starting points for the development of possible solutions: (i) *Regulatory Design and Implementation*, (ii) *EU Market Conditions and Investment Capacity*, (iii) *Ethics* and (iiii) *Benchmarking the US model*.

The desire for regulation in itself should not be criticised; it is considered an important and commendable initiative, but above all unsurprising, for two reasons.

First, the European Union always moves forward with rules first, and making rules often does not lead to innovation, unless they are clear and enforceable. Unfortunately, as emerged, this is not the case. As expressed in IBV, when institutional constraints, which should provide certainty, impose uncertainty and excessive rigidity, resources and skills may be needed to deal with the regulatory environment (Cook et al. 1983). Neo-institutionalism recognises that institutions could guide human action inefficiently, with the trade-off of providing certainty, but with the AI Act, certainty, except for compromises in innovation, is difficult to identify. The uncertainty is due to the textual ambiguity of the legislative act: (i) it is considered difficult to interpret and (ii) enforcement is entrusted to a complex network of different authorities. Rigidity, on the other hand, is most of all represented by the classification of systems at systemic risk, in particular has been criticised the measure risk of General-Purpose models based on the amount of computing power required to train them, using a threshold of  $10^{25}$  FLOPS. Although this measure represents the characteristics of modern models such as GPT-4, the use of this static threshold presents problems of foresight. In fact, one need only imagine

that if improvements in algorithmic approaches reduced the need for such a quantity of FLOPS, the metric would become obsolete.

Secondly, regulation is for those who do not have a product to sell. Today technology is key to understanding geopolitical competition, with innovation determining the nature and implications of the global scenario, and given that the EU is neither a geopolitical nor a geo-economic power, what remains is the possibility of being considered a regulatory power; a path that could prove counterproductive. Perhaps, rather than a positioning strategy, it is an acknowledgement of an inability to innovate, a kind of surrender.

When, in 2019, the President – Ursula von der Leyen – of the commission informally announced her intention to introduce regulation, it was not entirely rejected by the big foreign players. Regulation favours incumbents, who have the necessary resources to meet compliance costs and could have seen it as a way of safeguarding their competitive advantage by blocking the entry of new players. As pointed out by one of the interviewees - INT2 – “for large companies, this regulation is a sustainable burden, both from an economic and relational point of view”. Indeed, they often already have institutional relations and public affairs departments that maintain ongoing relations with European institutions, in this case the AI Office. Furthermore, the possible recruitment of new staff dedicated to compliance and regulatory management is a sustainable cost. From this perspective, adaptation to the new regulatory framework takes place through a strengthening of existing resources. It should be noted that the sustainability of the regulatory burden depends greatly on the position of companies in the artificial intelligence value chain. In fact, the situation changes for system developers, even when it comes to large players. The AI Act directly affects the products and services offered, requiring the adoption of engineering measures. In particular, for high-risk applications, compliance necessarily requires testing practices that must be incorporated into development processes. Despite this, for large players, this can always be considered a “light” burden.

On the contrary, it appears that those who pay the highest price for such regulation are start-ups and SMEs, entities traditionally associated with innovation initiatives, which often operate with limited budgets and personnel. These entities cannot afford to set up internal regulatory offices or technical compliance support structures and are therefore forced to resort to external consultants. Unlike large operators, they do not have direct access to institutional channels or the relational tools that enable informal and ongoing dialogue with the competent authorities, such as the AI Office. Furthermore, the costs of compliance – technical upgrades, personnel training, and legal consultations – are not easily dilutable: they are mainly CAPEX investments, the value of which is uncertain *ex-ante*, as companies are unable to predict how often and to what extent these resources will actually be used.

Confirming these considerations, as clearly expressed by Boccardelli and Brunetta (2014), in environments with high institutional constraints, network structures can positively influence the ability of actors to manage the constraints and opportunities of the environment.

From an institutionalist perspective, relationships between actors are considered interdependent rather than independent; links therefore take on the role of resources, both tangible and intangible, as well as skills, generating potential competitive advantages (Boccardelli & Brunetta 2014). For this reason, collaboration between the various actors in the ecosystem – institutions and businesses – is essential, through the building of a relationship of mutual trust. However, the European Union still has a predominantly *top-down* regulatory process, characterised by limited consultation and participation of businesses, particularly innovative ones. This dynamic is clearly evident from the results of empirical research, in which one interviewee (INT3) highlights the presence of regulatory initiatives developed without any real consultation with the private sector. A public policy framework integrated with private sector know-how – particularly that of providers and deployers – could lead to regulation that is more consistent with the modern paradigm. This would require a compromise between hard law and content sharing with those most affected by regulation, creating a paradigm that could be defined as “co-regulation”, largely inspired by the US model, characterised by a high level of trust between the public and private sectors. On the contrary, the ex-ante approach taken by the European Union, aimed at preventing high-risk models, is perceived as a form of regulatory “petrification”.

This gives the impression that the European Union is chasing technological developments, making it difficult to imagine that this approach can withstand future advances. This concept can be traced back to literature on market pathologies, in a sort of “technological inadequacy”, which occurs when technological developments create difficulties that prevent institutions from maintaining regulatory order (Boccardelli & Brunetta 2014, 147).

In light of this paradigm, it is not surprising that the research conducted shows that the compliance costs associated with the AI Act disproportionately affect start-ups and SMEs, compromising innovation processes and creating barriers to market entry. As widely emphasised in the literature, entrepreneurial choices depend to a large extent on the perception of the institutional quality of a given context (Boccardelli & Brunetta 2014). This quality derives from the certainty and clarity of the institutional framework, which tend to favour entrepreneurial activity. However, it should be emphasised that certainty and clarity do not necessarily coincide: mere formal existence does not guarantee clarity.

In line with one of the central concepts in institutionalist literature – organisational isomorphism theorised by Powell & DiMaggio (1983) – such a regulatory architecture can trigger

two specific mechanisms: coercive and mimetic. The first refers to formal and informal pressures – in this case mainly formal – exerted on organisations, i.e. companies, by other organisations, particularly regulators. This leads actors to converge towards joint homogenisation processes in order to avoid possible sanctions. The second mechanism, on the other hand, is characterised by the imitation of behaviours considered successful by other organisations. This dynamic is particularly evident in contexts characterised by high uncertainty, where the strategy perceived as most effective is to adapt to the practices adopted by other actors. The European Union should necessarily hope for the prevalence of the first mechanism of isomorphism, considering that the most successful organisations in the field of innovation today are, in most cases, located overseas.

The problem underestimated stems from a legitimate assumption: the organisational field tends to survive according to its degree of alignment with the regulatory environment and, in an attempt to gain legitimacy, companies comply with the rules in order to survive. However, little is said about cases where the pursuit of regulatory alignment itself leads to the failure of companies to survive; when regulations do not favour such survival, the only alternative is to move to other contexts. An alternative that is interpreted as a form of avoidance strategy is *jurisdiction shopping*, i.e. the selection of an institutional environment considered to be more favourable (Boccardelli & Brunetta 2014, 148).

In this context, the issue of investment emerges as one of the key elements in understanding the effects of regulation on the European innovation ecosystem. The results show that the European market is characterised by a structural tension between strong research and development (R&D) capacity and a weak capacity to transform technological innovation into competitive products and services. This structural tension, combined with the compliance costs introduced tends to amplify existing fragilities. This results in a problem of *institutional attractiveness*: where the regulatory framework is perceived as rigid – particularly by overseas players, where the availability of risk capital is greater – and not offset by adequate investment policies, there is a risk of further compromising the possibility of a global technology player with a European signature emerging.

The regulations currently being introduced are not fully aligned with the subject under discussion. Although the intentions are certainly laudable, the attempt to regulate and predict a phenomenon that is inherently difficult to anticipate is also affected by the partial absence of the subject itself in the European landscape: AI. Those who interpret this process as a possible institutional failure are not entirely wrong. As noted by Boccardelli and Brunetta (2014, 146), “institutional failure can occur when the institution is absent from a given context, creating situations of advantage for certain actors or favouring market opportunities”; but, in the case in question, such opportunities and economic advantages appear difficult to identify.

It's not a coincidence that it is precisely those major players who were not initially entirely opposed to regulating the new technology who are adopting strategies that can be traced back to influence-based rents. The first is *delaying*, which falls within the group of avoidance strategies and aims to evade the institution (Boccardelli & Brunetta 2014). In the current context, delaying manifests itself through continuous pressure exerted by large operators aimed at slowing down the implementation of the AI Act, as demonstrated by repeated requests for clarification and interpretation of its application. The second method is *perception management*, which constitutes a manipulation strategy (Boccardelli & Brunetta 2014). Its objective is to manipulate the reputation of regulation, not by directly challenging the rule, but by constructing a narrative that emphasises its rigidity, ambiguity and potential negative effects on innovation. In this way, regulation is gradually delegitimised.

Confirming these tensions that emerged in the results, on 19 November 2025, the European Commission presented the long-awaited Digital Omnibus, which aims to reduce technological bureaucracy in the European Union.

Among the simplifications envisaged, the package provides for an extension of up to 16 months to comply with certain obligations of the AI Act, particularly those concerning high-risk systems. This simplification is by no means bad news; it is recognition that the burden of such bureaucracy increases uncertainty and transaction costs, hindering the innovation process. However, this same dynamic creates a regulatory paradox. Developers are faced with a strategic dilemma: on the one hand, they must comply with a regulatory regime that is formally in force but already perceived as likely to be revised; on the other hand, they must wait for the announced changes, exposing themselves to the risk of temporary non-compliance. In both cases, regulatory uncertainty ultimately translates into additional costs, which affect development, investment and time-to-market decisions.

This idea of simplification is the recognition of an unbalanced approach between protection and innovation, the recognition that the European Union cannot afford to protect every possible aspect in this race and be left out of the game by being only the referees. In addition to being a partial step backwards, it is the search for an alternative.

## 5.2. *Graduated Compliance Model*

The aim of *Graduated Compliance Model* is to contribute to the debate proposing a regulatory approach that seeks to strike a more effective balance between the protection of fundamental rights and technological innovation. In the current European paradigm, regulation tends to favour protection

with an *ex-ante* approach, assuming that anticipating risk is the most efficient path: while this is the most correct in “protective” terms, it is entirely counterproductive in “innovative” terms.

The recommendations developed in this work stem from the need for a partial rethinking of the regulatory approach, especially with regard to the methods and timing of its application. The peculiarity of the model is based on the idea that effective regulation of artificial intelligence cannot be the result of unilateral intervention by public institutions alone, but must instead be based on a principle of *co-regulation* between the public and private sectors, understood respectively as institutions and actors in the technological ecosystem. Flexibility is a necessary condition for a regulatory system to be able to regulate but, above all, to adapt to a technology that evolves over time at an unpredictable speed. Ignoring this dynamic, known in the literature as the “pacing problem”, means not being fully connected to the object being regulated.

Although the best solution would be to deregulate the system – or at least partially deregulate it with mainly *ex-post* enforcement – it is recognised that the most realistic “politically” is the following. For this reason, the model proposed here does not aim to subvert the AI Act, but rather to identify more sustainable ways of introducing elements of flexibility without sacrificing the protection objectives pursued by the European legislator.

The core of the model lies in the distinction between two phases of the business lifecycle: early and maturity. The differentiation is inspired by regulatory experiences already present in the Italian legal system, in particular Decree-Law 179/2012 regarding “innovative start-ups”. In that context, the status of start-up derives from the recognition of the specific nature of the entity engaged in the development of high-tech innovations, characterised by uncertainty and limited resources. Similarly, in the proposed model, the *innovative provider* in the initial phase (*Phase I*) is defined by its functional role in the technological development process. In particular, the model adopts the identification criteria of Decree-Law 179/2012 focusing on requirements relating to incorporation and duration (5 years), residence (European space) and, above all, the corporate purpose, relating to the development and marketing of innovative products and/or services with high technological value. To further qualify the innovative profile, as also established by Italian law, the entity must meet one of the three requirements set out in Italian legislation: a minimum share of investment in Research and Development, the employment of highly qualified personnel or the ownership of intellectual property. It is advisable not to include limits relating to turnover or profit distribution, focusing instead on the intrinsic nature of the activity carried out. It is therefore recognised that a provider operates in a regime of technological uncertainty and economic fragility, placing it at a disadvantage compared to big players. In light of this, it would be rational for the European Union to grant such entities a

certain degree of regulatory flexibility: a proportionate adaptation of the burdens of the AI Act to the actual organisational capacity and level of maturity of the provider in the early development phase.

Of course, this distinction is not intended as an automatic rule. In line with the nature of the model and with the aim of avoiding regulatory arbitrariness, it is recommended that the identification of the two phases and the transition from one to the other should take place through a process of consultation with the actors in the ecosystem, enhancing the principle of *co-regulation*. No fixed thresholds or technical criteria are proposed, but a central role is given to dialogue.

### 5.2.1 *Early Phase and Maturity Phase*

As previously mentioned, the model applies with a distinction between two phases, the objective of which is to try to modulate the application of regulatory obligations over time in proportion to the state of development of the innovation and, above all, to the position of the provider developing it.

In *Phase 1*, the AI system is developed by a provider that still has the typical characteristics of an innovative entity in its early stages. The system may be subject to frequent changes and may still be in a phase of development and market validation: at this stage, the product cannot be considered definitively stable. Even where the system has already entered the market, this does not, in itself, affect the provider's classification within *Phase 1*, as long as the provider maintains the characteristics of an early-stage innovator.

In light of this, applying the entire set of obligations would be economically unsustainable and probably conceptually inadequate. In particular, the *ex-ante* conformity assessment provided for in Article 43 of the AI Act presupposes a “finished” product that is suitable for *Conformité Européenne* (CE). Applying this procedure means imposing certification on an unstable object, which would require a new conformity assessment for each substantial modification as provided for in Article 3, point 23, and Article 43 itself.

In this phase, on the other hand, the obligations relating to technical documentation (Art.11), transparency and information provided to users (Art.13), as well as the requirements for record-keeping and logging (Art.12) and human oversight (Art.14) shall apply, in a flexible manner, to the extent necessary to ensure the traceability and verifiability of the AI system.

The objective at this stage of regulation is not to formally validate the system, but to make the development process transparent and accountable, and verifiable where necessary *ex-post*.

*Phase 2* begins when the provider no longer exhibits the typical characteristics of an innovative entity, and consequently the most onerous obligations apply as well. The transition to this phase is not determined by a technical risk assessment or computational thresholds, but simply by the cessation of the *innovative provider* status.

The provider is subject to the application of the risk management system (Art. 9), governance and data quality requirements (Art. 10), accuracy, robustness, and cybersecurity requirements (Art. 15), as well as the adoption of a quality management system (Art. 17). In this context, Article 43 assumes its function, requiring a formal compliance assessment.

In this context, certification does not represent a disproportionate burden, but rather a tool consistent with the provider's potential maturity and ability to absorb the costs and effects. The transition to *Phase 2* should be interpreted as a natural process linked to the growth of the company.

The use of this concept allows the model to be applied to *General Purpose Models AI* (GPAI) as well, on the condition that the provider falls within the established parameters. Therefore, as suggested for high-risk systems, it is recommended to exempt GPAI from certain obligations and to revise the  $10^{25}$  FLOPS computational threshold which, as previously noted, is at risk of becoming obsolete due to rapid technological advancements.

### 5.2.2 Regulation and Investment: Risk Realignment

Returning to one of the most significant findings to emerge from the analysis carried out in Chapter 4, the relationship between regulation and investment in the European AI ecosystem would change with the application of *Graduated Compliance Model*. European AI regulation operates in a context characterised by structural fragility: limited availability of risk capital and weak execution capacity in transforming research and development into marketable products and services.

From an investor's point of view, the central issue of regulation is the nature of the regulatory risk it introduces. Investments in technology sectors, typical of venture capital, generally take place progressively, through technical and commercial validation phases. Capital allocation is based on the possibility of spreading risk over time: the investor's commitment grows in proportion to the consolidation of the product/service, increasing exposure only where initial uncertainties have been partially overcome. Intuitively, it would not be wrong to think that imposing burdensome regulatory obligations, without considering the initial design phases, would affirm regulatory risk as an additional component of investment assessment.

In this scenario, the model introduces an element of realignment between regulation and investment logic. The distinction between the two phases makes it possible to reduce regulatory risk where market and technological uncertainty is greater. By postponing the application of the most burdensome obligations to the next phase, the model allows investors to evaluate the *innovative provider* before having to bear high regulatory costs. It is a risk system that is more consistent with the innovation financing cycle.

A further effect concerns the concept of institutional predictability. For investors, the ability to anticipate when compliance becomes burdensome reduces the degree of uncertainty.

This paradigm takes on further importance when assessing the European Union's position in the global innovation landscape. As the results show, the European market, in addition to being almost devoid of future disruptive technological players, is perceived as less attractive than other contexts, particularly in the eyes of non-European investors, due to its current regulatory rigidity and bureaucratic complexity. A regulation perceived as more “flexible,” despite addressing the absence of strong investment policies, becomes a tool for attractiveness, also reducing the incentive to move innovative projects to more favourable contexts.

The proposed model aims to strengthen the European ecosystem’s capacity to attract capital, support innovation and encourage the emergence of new players in the artificial intelligence sector.

### 5.2.3 The Third Way

In order for the model to be applied within the current institutional framework, it is necessary to identify the tools through which the temporal reallocation of obligations is possible. In this regard, regulatory sandboxes, the role of the AI Office and enforcement methods emerge as the main candidates for introducing the proposed flexibility.

In the text of the AI Act, as repeated many times in the work, sandboxes are tools designed to allow controlled experimentation with AI systems under the supervision of the authorities. In the *Graduated Compliance Model*, they take on a more structural function: they are interpreted as the natural institutional space for *Phase 1*. Sandboxes operate as a regulated environment in which the most burdensome obligations are temporarily suspended, while the necessary ones remain in place. This offers a twofold advantage: for *innovative providers*, who can navigate a more flexible regulatory regime, and for institutions, which acquire relevant information to predict the potential impacts of technologies.

As a second element for the implementation of the model, the AI Office acts as a coordinating body responsible for providing ongoing guidance to regulated entities. In this context, it promotes a *co-regulation* approach between the public and private sectors: a governance model in which institutions and private entities (companies, trade associations) work together to define and implement rules and standards. In particular, the AI Office should define – after consultation with the ecosystem – the criteria for the transition from the development phase to the mature phase, identifying which articles and obligations should be applied. This perspective represents a third way between traditional (top-down) regulation and private self-regulation. The aim is to integrate the technical expertise of the sector with the principles of public accountability, formalising a shared path to compliance that protects innovation without compromising the protection of fundamental rights.

The third pillar is represented by the enforcement methods provided for in the model. The current approach envisages enforcement in rather binary terms: compliance and non-compliance. This approach does not consider the maturity of the project, compromising innovation from the outset. In the proposed model, enforcement is progressive and proportional. In the first phase, the authorities should only apply obligations that allow for a sufficient level of traceability and verifiability, favouring a path of gradual alignment with regulatory standards. With the transition to *Phase 2*, the enforcement provided for is that currently in force in the legal system.

In summary, regulatory sandboxes, the AI Office’s approach and progressive enforcement constitute a sort of operational infrastructure for the model. It does not introduce a formal non-application of the obligations provided for, but suggests a temporal modulation of the enforcement methods and their enforceability, consistent with the principles of proportionality.

### 5.3 Ethics as a competitive advantage?

Today, in Europe, ethics occupies a somewhat paradoxical position in the debate on the regulation of artificial intelligence. On the one hand, it represents the most profound justification for regulatory intervention: ethics is invoked because AI is rightly considered a technology capable of affecting human behaviour. On the other hand, ethics is frequently used in the language of competition, as if it could actually constitute a lever of strategic differentiation for the European Union.

Considering Professor Luciano Floridi’s perspective, modern technology means that human beings no longer live in a world divided between online and offline, but in a hybrid, dimension defined as the “onlife” condition; a new reality that takes shape within the infosphere, a relational space in

which the analogue and digital dimensions merge seamlessly, rendering obsolete the binary distinction between physical and virtual reality (Floridi 2014). In such a context, it seems right that ethics should operate as a guiding principle, as a criterion through which society establishes not only what is permitted, but what is acceptable and, above all, what is desirable (Floridi 2022).

Yet, precisely when ethics is presented as a “competitive advantage”, its meaning changes. The lexicon of advantage belongs to the logic of economic rivalry, where decisions are evaluated in terms of performance. On the contrary, ethics does not arise as a tool for maximisation, but rather indicates within what limits and according to what criteria it is “right” to act. Using it as a strategic asset is anything but ethical. As if, suddenly, the tool for guiding collective action had become a topic of institutional marketing: a redefinition of the very role of ethics.

The perceptions of the actors in the ecosystem reflect this tension. Ethics is rarely challenged *a priori* when it comes to protecting citizens or safeguarding fundamental rights; however, it becomes a source of scepticism when it is interpreted as an identity narrative, or worse still, a competitive lever. It is not surprising that, in the eyes of some, Europe may be considered weak: “*Ethics? ...people cling to ethics out of weakness...*” (INT4).

Ethics is not designed to strengthen the position of a player, but to make the evolution of technology responsible. The idea of “competitive advantage” seems counterproductive. This does not mean that ethics cannot generate positive economic effects, but that such effects cannot constitute the normative reason for its adoption.

Since AI shapes the daily life of human beings, it is right to talk about ethics, but it is not an advantage, but a condition; it does not serve competition, but precedes it.

## CONCLUSION

The thesis demonstrates a conclusion that is, unfortunately, not surprising: although the AI Act is a commendable initiative aimed at promoting reliable technology, it risks compromising the very competitiveness that the European Union claims to want to preserve and encourage. The answer to the central question of this research – *Does the AI Act succeed in balancing protection and innovation?* – reveals a critical asymmetry. Excessive regulation, characterised by an *ex-ante* approach and marked interpretative uncertainty, translates into a structural brake on the spread of emerging technologies on European soil.

Beyond the specific regulatory framework, this dynamic reflects a broader institutional tension: when regulation precedes technological consolidation, it may inadvertently reshape innovation incentives and firms behaviour. The lack of advanced AI models, the lack of attractiveness and the growing competitive gap with the United States and China are indicators of a trend in which regulation has preceded innovation, compromising the market before it could even consolidate. In this scenario, regulation actively influences the geography of technological development.

The most significant distortion, however, lies in the asymmetric burden imposed on start-ups and SMEs. By imposing compliance structures designed primarily for large incumbent operators, the regulatory architecture risks discouraging the emergence of a European-led technology. This dynamic creates a paradox: a continent that has historically been a driver of innovation risks laying the foundations for its future dependence on foreign technology leaders. The absence of cutting-edge technology would render the European Union a mere referee in a game dominated by powers that have prioritised innovation over regulatory rigour.

The recent intervention by the European Commission through the Digital Omnibus AI of November 2025 represents a belated but necessary admission of the existence of an imbalance that requires urgent corrective measures.

In view of this, the thesis proposed the *Graduated Compliance Model*, a “third way”, as an innovative solution. By distinguishing between stages of development and stages of corporate maturity, this model suggests that compliance should not be a binary event, but a progressive path supported by co-regulation mechanisms and the active role of the AI Office.

Only by implementing a system capable of accompanying innovation will Europe be able to transform ethics from a “lifeline for the weak” into a condition for attempting to enter the arena of technological sovereignty, while respecting the individual. Without such a rebalancing, the AI Act risks being remembered not as a turning point for “Trustworthy AI”, but as the document that sealed the EU’s definitive competitive marginality in the global race for artificial intelligence.



## BIBLIOGRAPHY

- Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Boston: Harvard Business Press.
- Ahern, Deirdre. 2025. “Operationalising AI Regulatory Sandboxes under the EU AI Act: The Triple Challenge of Capacity, Coordination and Attractiveness to Providers.”
- Albert Heijn. 2019. “Albert Heijn Zet Kunstmatige Intelligentie in Tegen Voedselverspilling.” *Albert Heijn Nieuws*, May 20, 2019.
- Anderson, Marc M. 2022. “Some Ethical Reflections on the EU AI Act.” *CEUR Workshop Proceedings 3221*. Presented at the International Conference on AI, Law, and Society (IAIL).
- Baird, Aaron, and Laurence M. Maruping. 2021. “The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts.” *MIS Quarterly* 45 (1): 315–341.
- Barney, Jay B. 2001. “Is the Resource-Based ‘View’ a Useful Perspective for Strategic Management Research? Yes.” *Academy of Management Review* 26 (1): 41–56.
- Boccardelli, Paolo, and Francesca Brunetta. 2014. *Istituzioni, mercati e imprese: Il ruolo delle istituzioni nelle strategie di crescita*. Milano: McGraw-Hill Education.
- Bostrom, Nick. 2016. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- Bradford, Anu. 2020. *The Brussels Effect: How the European Union Rules the World*. Oxford: Oxford University Press.
- Camuffo, Arnaldo, and Roberto Cappellari. 1996. *Le teorie neoistituzionaliste*. Torino: UTET.
- Carey, Samuel. 2025. “Regulating Uncertainty: Governing General-Purpose AI Models and Systemic Risk.” *European Journal of Risk Regulation*: 1–17.

Chen, L. Y., and L. Liu. 2023. “China Ends Tech Crackdown with Fines on Tencent and Ant Group.” Bloomberg.

Chui, Michael, Eric Hazan, Roger Roberts, Alex Singla, Kate Smaje, Alex Sukharevsky, Lareina Yee, and Rodney Zempel. 2023. *The Economic Potential of Generative AI: The Next Productivity Frontier*. McKinsey Global Institute.

Chun, Jon, Christian Schroeder de Witt, and Katherine Elkins. 2024. *Comparative Global AI Regulation: Policy Perspectives from the EU, China, and the US*. Kenyon College and Oxford University.

Coursera. 2023. “AI vs Deep Learning vs Machine Learning: A Beginner’s Guide.”

Davtyan, Tatevik. 2024. “The U.S. Approach to AI Regulation: Federal Laws, Policies, and Strategies Explained.” SSRN.

DeepMind. 2019. “Machine Learning Can Boost the Value of Wind Energy.” DeepMind Blog, February 26, 2019.

Dellermann, Dominik, Philipp Ebel, Matthias Söllner, and Jan Marco Leimeister. 2019. “Hybrid Intelligence.” *Business & Information Systems Engineering* 61 (5): 637–643.

Deloitte. 2024. *EU Artificial Intelligence Act Deep Dive*. Deloitte Netherlands, Digital Regulations Team.

Dennett, Daniel C. 2019. “What Can We Do?” In *Possible Minds: Twenty-Five Ways of Looking at AI*, edited by John Brockman, 41–53. New York: Penguin.

DiMaggio, Paul J., and Walter W. Powell. 1983. “The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields.” *American Sociological Review* 48 (2): 147–160.

Dong, Xingshuai, and Massimiliano L. Cappuccio. 2024. “Applications of Computer Vision in Autonomous Vehicles: Methods, Challenges and Future Directions.” arXiv.

Draghi, Mario. 2024. *The Future of European Competitiveness: A Competitiveness Strategy for Europe*. Brussels: European Commission.

Du, Z. Y. 2018. "A Literature Review on Institutional Environment and Technological Innovation." *American Journal of Industrial and Business Management* 8: 1941–1950.

European Commission, High-Level Expert Group on Artificial Intelligence. 2019. *Ethics Guidelines for Trustworthy AI*. Brussels: European Commission. (Nota: presente due volte nel testo originale).

European Commission, Joint Research Centre. 2020. *AI Watch: Historical Evolution of Artificial Intelligence – Analysis of the Three Main Paradigm Shifts in AI*. Luxembourg: Publications Office of the European Union.

European Union. 2023. "Recital 12." *Artificial Intelligence Act*.

European Union. 2024. "Article 3: Definitions." *Artificial Intelligence Act*.

European Union. 2024. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 (Artificial Intelligence Act). *Official Journal of the European Union*, July 12, 2024.

Felsmann, Balázs. 2016. "Do Institutions Matter in Business Strategy? The Changing Focus of Strategic Management to Institutions: A Literature Review." *Budapest Management Review*.

Fenwick, Mark, Erik P. M. Vermeulen, and Marcelo Corrales Compagnucci. 2021. "Business and Regulatory Responses to Artificial Intelligence: Dynamic Regulation, Innovation Ecosystems and the Strategic Management of Disruptive Technology." In *Flexible Regulation of Digital Innovations*, 1–21. Singapore: Springer.

Feuerriegel, Stefan, Jochen Hartmann, Christian Janiesch, and Patrick Zschech. 2023. "Generative AI." *Electronic Markets* 33 (4).

Floridi, Luciano. 2014. *The Fourth Revolution: How the Infosphere Is Reshaping Human Reality*. Oxford: Oxford University Press.

Floridi, Luciano. 2019. "What the Near Future of Artificial Intelligence Could Be." *Philosophy & Technology* 32 (1): 1–15.

Floridi, Luciano. 2021. "The European Legislation on AI: A Brief Analysis of Its Philosophical Approach." *Philosophy & Technology* 34: 215–222.

Floridi, Luciano. 2022. *Etica dell'intelligenza artificiale: Sviluppi, opportunità e sfide*. Milano: Raffaello Cortina Editore.

Giangiaco, Vincenzo, and Milena Lorenzi. 2024. "Direttiva UE 2024/2853: Il Nuovo Regime della Responsabilità per Danno da Prodotti Difettosi." CMS, December 4, 2024.

Gioia, Dennis A., Kevin G. Corley, and Aimee L. Hamilton. 2013. "Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology." *Organizational Research Methods* 16 (1): 15–31.

Glaser, Barney G., and Anselm L. Strauss. 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine Publishing Company.

Hacker, Philipp. 2024. "AI Regulation in Europe: From the AI Act to Future Regulatory Challenges." In *Oxford Handbook of Algorithmic Governance and the Law*. Oxford: Oxford University Press.

Hartmann, Simon, Thomas Lindner, Jakob Müllner, and Jonas Puck. 2022. "Beyond the Nation-State: Anchoring Supranational Institutions in International Business Research." *Journal of International Business Studies* 53: 1285–1300.

Horowitz, Julia. 2023. "JPMorgan Restricts Employee Use of ChatGPT." CNN, February 22, 2023.

Ingram, Paul, and Brian S. Silverman, eds. 2002. *The New Institutionalism in Strategic Management*. Amsterdam: Elsevier.

Kulothungan, Vikram, and Deepti Gupta. 2025. "Towards Adaptive AI Governance: Comparative Insights from the U.S., EU, and Asia." IEEE 11th Conference on Big Data Security on Cloud.

- Lawton, Thomas C., and Steven M. McGuire. 2001. "Supranational Governance and Corporate Strategy: The Emerging Role of the World Trade Organization." *International Business Review* 10 (3): 227–248.
- Lynn, Leonard H., N. Mohan Reddy, and John D. Aram. 1996. "Linking Technology and Institutions: The Innovation Community Framework." *Research Policy* 25 (1): 91–106.
- Maestri, Alberto. 2024. "Algovetica: Le sfide dell'Intelligenza Artificiale per un Nuovo Umanesimo Digitale secondo Padre Paolo Benanti." 24ORE Business School, March 26, 2024.
- Majumdar, Sumit K., and Alfred A. Marcus. 2001. "Rules versus Discretion: The Productivity Consequences of Flexible Regulation." *Academy of Management Journal* 44 (1): 170–179.
- McKinsey & Company. 2024. "What Is Generative AI?" McKinsey Explainers, April 2024.
- Meyer, John W., and Brian Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83 (2): 340–363.
- Moens, Barbara, and Melissa Heikkilä. 2025. "How the EU Botched Its Attempt to Regulate AI." *Financial Times*, November 20, 2025.
- Moravec, Hans. 1988. *Mind Children: The Future of Robot and Human Intelligence*. Cambridge, MA: Harvard University Press.
- Nah, Fiona Fui-Hoon, Keng Siau, Jingyuan Fan, and Brian L. Zhang. 2023. "Generative AI and ChatGPT: Applications, Challenges, and AI–Human Collaboration." *Journal of the Association for Information Systems* 24 (2): 283–300.
- Nilsson, Nils J. 2009. *The Quest for Artificial Intelligence*. Cambridge: Cambridge University Press.
- North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

Nye, John. 2008. *Institutions and Institutional Environment*. Cambridge: Cambridge University Press.

Okoro, Uloma. 2025. “Artificial Intelligence Governance and Regulation: The Impact of the EU AI Act on Innovation, Accountability, and Global Compliance in a Digital Age.” University of East London – Royal Docks Business School. SSRN.

Paffenholz, David. 2025. “Juicebox Raises \$36 M Led by Sequoia Capital.” Juicebox Blog, October 26, 2025.

Peng, Mike W., and Eric G. Pleggenkuhle-Miles. 2009. “Current Debates in Global Strategy.” *International Journal of Management Reviews* 11 (1): 51–68.

Peng, Mike W., Sunny Li Sun, Brian C. Pinkham, and Hao Chen. 2009. “The Institution-Based View as a Third Leg for a Strategy Tripod.” *Academy of Management Perspectives* 23 (3): 63–81.

Priem, Richard L., and John E. Butler. 2001. “Is the Resource-Based ‘View’ a Useful Perspective for Strategic Management Research?” *Academy of Management Review* 26 (1): 22–40.

Qiao-Franco, Guang, and Rong Zhu. 2024. “China’s Artificial Intelligence Ethics: Policy Development in an Emergent Community of Practice.” *Journal of Contemporary China* 33 (146).

Roberts, Huw, Josh Cowls, Jessica Morley, Mariarosaria Taddeo, Vincent Wang, and Luciano Floridi. 2021. “The Chinese Approach to Artificial Intelligence: An Analysis of Policy, Ethics, and Regulation.” *AI & Society* 36 (1): 59–77.

Scherer, Andreas G., and Guido Palazzo. 2011. “The New Political Role of Business in a Globalized World: A Review of a New Perspective on CSR and Its Implications for the Firm, Governance, and Democracy.” *Journal of Management Studies* 48 (4): 899–931.

Sheikh, Haroon, Corien Prins, and Erik Schrijvers. 2023. *Research for Policy: Mission AI – The New System Technology*. The Hague: The Netherlands Scientific Council for Government Policy.

Shrinivas, Santhosh Kumar, and Pavithra Shetty. 2024. "Impact of Artificial Intelligence on Business, Business Operations and Strategy." *International Journal of Creative Research Thoughts (IJCRT)* 12 (8).

Simion, Ana. 2024. "What Does the EU Competitiveness Report Say about AI and Competition Law?" AI Accelerator Institute, December 6, 2024.

Sky News. 2020. "Tesla Criticised for 'Lack of System Safeguards' after Autopilot Crash." Sky News.

Smith, Marcus, and Seumas Miller. 2025. "Technology, Institutions and Regulation: Towards a Normative Theory." *AI and Society* 40 (2): 1007–1017.

Sukharevsky, Alexander, Eric Hazan, Sven Smit, Marc-Antoine de la Chevasnerie, Marc de Jong, Solveigh Hieronimus, Jan Mischke, and Guillaume Dagorret. 2024. *Time to Place Our Bets: Europe's AI Opportunity*. QuantumBlack, AI by McKinsey.

Szczepański, Marcin. 2019. *Economic Impacts of Artificial Intelligence*. European Parliamentary Research Service (EPRS) Briefing. Brussels: European Parliament.

Turing, Alan M. 1950. "Computing Machinery and Intelligence." *Mind* 59 (236): 433–460.

Veale, Michael, and Frederik Zuiderveen Borgesius. 2021. "Demystifying the Draft EU Artificial Intelligence Act." *Computer Law Review International* 22: 97–112.

Viola, Roberto, and Luca De Biase. 2024. *La legge dell'intelligenza artificiale: L'AI Act europeo per trovare un senso nel futuro della tecnologia*. Milano: Il Sole 24 Ore.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.

Zhu, Junhua. 2024. *China's Approach to AI Standardisation: State-Guided but Enterprise-Led*. Helsinki: Finnish Institute of International Affairs.

