Università Ca'Foscari Venezia

Master's Degree in Comparative International Relations

Artificial Intelligence as a General Purpose Technology

An exploratory analysis of PCT patents

Co-supervisor

Prof. Francesco Lissoni

Supervisor

Prof. Luca Di Corato

Graduand

Alessio Nardin

Student number

858966

Academic Year

2019/2020

Acknowledgements

Throughout the writing of this thesis I have received a great deal of support and assistance.

I would first like to extend my deepest gratitude to my supervisor, Professor Francesco Lissoni, whose expertise and guidance was invaluable in formulating the research questions and methodology. Despite working in another university, you have always been available to clear my doubts. Your insightful feedback pushed me to sharpen my thinking and brought my work to a higher level.

I would like to thank all those who have worked to create the Globalization Program of the Venice International University, in particular the Vice Dean and Globalization Program Scientific Director Luca Pes. It was during my participation to the program that I met Professor Lissoni and decided to focus my research on innovation policy. VIU allowed me to experience Venice from another perspective, discussing with students from all over the world.

I am also grateful to Professor Luca di Corato, my supervisor from Ca' Foscari University, who offered to guide me during the thesis writing and provided me with critical advice for its completion.

I would also like to acknowledge the colleagues from my internship at the Brussels Office of the Veneto Region. Thanks to the internship I had the possibility to experience firsthand the dynamics of European Institutions. It was during a job-related event that I understood the need to focus my research on AI and technology. For the next decades, technology policy is here to stay.

In addition, I would like to thank my mother Marisa, who supported me and pushed me to finally conclude this work, and my brother Denis, for his wise counsel and sympathetic ear. You have read this thesis multiple times, and helped me to reorganize my thoughts. Thank you for always being there for me.

Finally, I could not have completed this dissertation without the support of my friends: Aurora, Annachiara, Benedetta, Davide, Riccardo, Marco, Claudia, Daniela, and Anna, who have been invaluable traveling companions during the writing of this thesis and happy distractions to rest my mind outside of my research.

Short Abstract

English

This thesis aims to explore some of the issues concerning the impact of Artificial Intelligence on innovation and specifically whether AI should be considered both as a General Purpose Technology and a research tool that can itself be used to increase the rate of innovation, also defined in the literature as a General Method of Invention. AI can both cause severe disruption in the economy and society and foster economic growth and development. After a historical and conceptual introduction of the mechanics of AI and a presentation of the institutions that regulate innovations (intellectual property rights), the work focus on the impact of AI on innovation, both through a qualitative analysis of the different ways and in which AI can be deployed and through the deployment of traditional and new indicators to detect GPTs in patent data.

Italiano

La presente tesi si propone di esplorare alcune delle questioni riguardanti l'impatto dell'Intelligenza Artificiale sull'innovazione. In particolare, indaga se l'IA possa essere considerata sia come una General Purpose Technology che come un innovativo strumento di ricerca, una categoria tecnologica definita nella letteratura come "General Method of Inventing" . L'IA causerà perturbazioni economiche e sociali, mentre favorirà la crescita economica e lo sviluppo tecnologico. Dopo un'introduzione storica e concettuale sulle meccaniche dell'IA e una presentazione delle istituzioni che regolano l'innovazione (come i diritti di proprietà intellettuale), la tesi si concentra sull'impatto dell'IA sull'innovazione, sia attraverso un'analisi qualitativa delle diverse modalità in cui l'IA può essere impiegata che attraverso l'impiego di nuovi e tradizionali indicatori utilizzati per individuare le GPT nei brevetti.

Long Abstract

L'Intelligenza Artificiale (IA) si sta configurando come una delle principali priorità delle potenze mondiali, coinvolgendo settori che vanno dalla sicurezza, all'economia, alla gestione dell'opinione pubblica. Gli Stati Uniti e la Cina si sono lanciati in una corsa tecnologica, mentre l'Unione Europea sta cercando di recuperare terreno, anche attraverso l'utilizzo di strumenti di regolamentazione, quali il Digital Market Act e il Digital Services Act, proposti dalla Commissione Europea e al momento discussi in Parlamento. Tutti questi sono segnali che la tecnologia sta acquisendo un ruolo sempre più importante a livello internazionale, con effetti trasversali su tutti i livelli. Nonostante la sua importanza nel determinare le dinamiche tra i vari attori internazionali, sia dal punto di vista normativo che sostanziale, il fattore tecnologico è molto spesso trattato come un fattore esogeno e indipendente o come prodotto delle azioni e desideri dei vari attori, a seconda della prospettiva adottata. In questa tesi si adotta una prospettiva intermedia, che considera la tecnologia (e in particolare, l'Intelligenza Artificiale) sia come la causa di importanti trasformazioni economiche, sociali e politiche che come una delle conseguenze delle istituzioni che regolano il processo di innovazione e del comportamento degli attori internazionali.

Sotto il macro-termine di IA vengono incluse tutta una serie di tecniche matematico-statistiche che mirano a simulare processi, ragionamenti e comportamenti umani. In particolare, a partire dai primi anni 2000, la grande disponibilità di dati e l'abbassamento dei costi computazionali ha permesso di ottenere migliorare la performance di Machine Learning e Deep Learning, tecnologie che mirano a prevedere il corretto output y dato un input x senza un intervento umano. Mappare input a output (cause a effetti, domande a risposte) è una parte essenziale di ogni processo decisionale e un aumento della precisione della mappatura può all'automazione di alcune decisioni, specialmente quando il livello di precisione è ritenuto sufficientemente alto e il costo di errore sufficientemente basso. Nelle situazioni più complesse, dove il costo di errore è più alto, la capacità di previsione è complementare alla capacità di esprimere giudizi. L'IA si sta configurando come una tecnologia capace di sostituire la forza lavoro nelle mansioni monotone e ripetitive, ma allo stesso tempo capace di aumentare le "capabilities" di lavoratori, organizzazioni e pubbliche amministrazioni in situazioni complesse, aumentando il valore delle loro decisioni e il livello di responsabilità a loro associate.

La direzione e il tasso di sviluppo di queste tecnologie è fortemente limitato e indirizzato dalle istituzioni preesistenti che regolamentano il processo innovativo. La legislazione relativa ai diritti di proprietà intellettuale ha un ruolo fondamentale nell'incentivare o disincentivare inventori e autori nell'ideazione e sviluppo di nuovi prodotti e processi. Nel secondo capitolo le motivazioni filosofiche ed economiche per la creazione di tali istituti vengono trattate in dettaglio, per poi preferire il paradigma utilitaristico, che afferma che l'alterazione delle dinamiche di mercato è giustificata solamente nel caso in cui migliori il "benessere economico complessivo" della società. Questo approccio è preferito in quanto l'eccesso di protezione è spesso causa di esternalità negative, che rallentano l'innovazione e incoraggiano comportamenti di rent-seeking, a discapito del benessere generale. Vengono poi presentati i tre diritti di proprietà intellettuale legati all'Intelligenza Artificiale: brevetti, diritto

d'autore e diritto sui-generis relativo ai database, per poi analizzare se l'IA abbia effettivamente bisogno di ulteriori sistemi di protezione da un punto di vista rigorosamente economico, raggiungendo la conclusione che al momento non vi è alcuna necessità di creare nuovi diritti relativi all'IA, anzi vi è un fortissimo rischio che ciò porti a ulteriori distorsioni del mercato, favorendo gli attori che hanno già acquisito posizioni dominanti.

Nel terzo capitolo vengono introdotte alcune nozioni relative all'effetto che l'IA avrà nel processo innovativo. Nel corso degli ultimi anni, parte della letteratura, specialmente quella legata all'economia dell'innovazione, ha identificato l'Intelligenza Artificiale come una possibile "General Purpose Technology" (GPT), una tecnologia a scopo generale, ovvero una tecnologia in rapida evoluzione, con un largo scopo applicativo e complementare sia a nuove tecnologie che a tecnologie preesistenti. Nel caso in cui l'IA si configurasse come tale, vi sarebbero forti conseguenze dal punto di vista economico, in quanto potrebbe portare a un drastico aumento della produttività, unito a una sostituzione di parte della forza lavoro in seguito all'implementazione dell'IA nei processi produttivi. Inoltre, le GPT sono caratterizzate da cicli virtuosi di innovazione, in cui avanzamenti in un'area possono beneficiare ambiti applicativi completamente differenti con numerose esternalità positive, un fenoneno chiamato "innovation complementarities". Parte della letteratura poi ritiene che l'IA si possa configurare anche come "Invention of a Method of Inventing" (IMI), ovvero come l'invenzione di un nuovo metodo per inventare, che in passato ha portato all'apertura di interi nuovi paradigmi scientifici. In tal modo, l'IA entrerebbe a far parte di una nuova categoria tecnologica chiamata "General Method of Inventing" (GMI) o metodo a scopo generale per inventare, il che implicherebbe un ulteriore incremento della velocità dell'innovazione tecnologica. Tutte queste particolari caratteristiche, insieme alla maturazione del mercato digitale, hanno particolari conseguenze in termini di struttura del mercato e concorrenza, che, come Crémer et al. (2019) suggerisce, stanno portando a una progressiva concentrazione del mercato nelle mani di pochi colossi tecnologici.

Nonostante in passato siano stati effettuati vari studi relativi ad altre GPT (elettricità, motore a vapore, tecnologie ICT) e sulla diffusione dell'IA, è al momento assente uno studio quantitativo specificatamente diretto a stabilire se l'IA effettivamente possa essere considerata come GPT. Questa tesi mira a porre le basi per rispondere a questa domanda utilizzando dati provenienti dai brevetti emessi dall'Organizzazione Internazionale per la Proprietà Intellettuale (WIPO) e utilizzando il database PATSTAT 2018 Autumn compilato dall'European Patent Office. Vengono utilizzati sia indicatori tradizionali della letteratura sulle GPT, come l'indice di generalità, la varietà di classi tecnologiche riscontrate e la proporzione di brevetti depositati annualmente, che dei nuovi indicatori basati sulla scienza delle reti e la teoria dei grafi. In particolare, la rete formata dai brevetti contenenti tecnologie di Intelligenza Artificiale viene modellata sotto forma di un network bipartito formato dai brevetti e dalle classi tecnologiche a loro assegnate dagli esaminatori, per poi analizzare l'evoluzione della centralità delle connessioni tra le varie classi tecnologiche. I risultati dell'analisi confermano che l'Intelligenza Artificiale dovrebbe essere categorizzata come GPT. Sebbene siano necessarie ulteriori ricerche, questi risultati confermano che l'attenzione riservata all'IA da parte di regolatori e governi è più che giustificata, in quanto l'avvento di una rivoluzione tecnologica di tale portata rappresenta un'ulteriore scossa all'economia già provata dalla pandemia di COVID-19. Sebbene le GPT possano drasticamente aumentare la crescita economica, il loro carattere trasformativo porta alla divisione della società in vincitori e vinti, ovvero in coloro che riusciranno a sfruttare i benefici della trasformazione e coloro che invece rimangono fermi nelle loro posizioni. Tutto dipenderà dall'abilità dei governi di investire in politiche di formazione e aggiornamento che riescano a far acquisire alla forza lavoro nuove competenze tecnologiche, in modo da evitare l'ulteriore accentuarsi delle disuguaglianze, fonte di scontento e di instabilità politica con ripercussioni anche sul piano internazionale.

Contents

Introduction

Artificial Intelligence (AI) technologies are at the center of academic and public debate. Policymakers around the globe are striving to find strategies that maximize the benefits that such a technological revolution may bring, while at the same time minimizing the negative effects of change. Technological innovation is both seen as a solution to the slowdown in productivity of recent decades and as a danger of increased social tension. At the international level, the US and China are about to launch themselves in a race towards developing autonomous systems to boost their economic and military power, while in 2019 the European Commission headed by von der Leyen relaunched the concept of *"technological sovereignty"*, with the long-term goal of securing strategic autonomy and a greater control of the European digital market. Nevertheless, notwithstanding the importance assigned by policymakers around the globe, technology is the great absent in the analysis of international relations (Herrera, 2003, 2012; Weiss, 2015). Herrera (2003) affirms that technology is generally introduced in the analysis in the form of either technological determinism or social constructivism. Technological determinism considers technological change as an exogenous phenomenon: it affects society from outside and enters only to disrupt and alter the equilibrium. Social constructivism instead treats it as a consequence of human agency and interests, and thus it is shaped by cultural, economic, and political forces.

I argue that both of these approaches are correct. Technology is a multidimensional phenomenon that is both exogenous and endogenous. As Weiss (2005) suggested, it is tied with international affairs by multiple patterns of interaction. Even if it is indeed true that the development and diffusion of technology are shaped by institutional settings and economic and political forces, it is also true that in many cases, new technologies are the necessary condition for the birth of new institutions and changes in the power distribution of economic and political forces. Technology may influence the dynamics of the international system in the form of an *"escaped genie"* that rapidly spreads along different dimensions of society, prompting a rapid change in the relations between actors on multiple levels, thus creating a fundamental shift in the international system. Often the international community is not able to keep the pace with rapid technological advances, which generate a *"steady stream of new policy issues"* (Weiss, 2015) that permeate society, politics, and the economy. At the same time, new technologies may change the international system's dynamics, creating new winners and opening up new areas of conflict while blurring previously clear concepts of international relations theories, such as sovereignty and security. Information and Communication Technologies (ICTs) have increased the speed of economic exchange, increased the ability of firms to operate internationally, and enabled the formation of a platform economy (Kenney and Zysman, 2016), allowing the rise of digital giants that are not only necessary for the functioning of the market but also have become strong actors in the international setting, influencing policy issues of states of medium dimension, fundamentally changing the distribution of bargaining power between state and non-state actors. Simultaneously, the rate and trajectory of technological innovation are shaped by institutions such as Intellectual Property Rights (IPRs), universities' research agendas, and economic and political competition. Innovation policy here plays a crucial role in determining the rate and direction of technological activity, since its outcomes dynamically alter the distribution of power between the actors of the international arena.

Today, scholars are pointing to Artificial Intelligence as the protagonist of the next great technological revolution, capable of disrupting the economic and power equilibrium both between nations and between states and non-state actors. Economists of innovation define these key enabling technologies as General Purpose Technologies (GPTs), which act as multipliers of change in all aspects of society, economy, and politics, potentially driving waves of technological innovation. GPTs are characterized by being subjected to rapid growth, economic pervasiveness, and strong complementarities with existing and new technologies. Rather than analyzing the international dynamics that will characterize the development of Artificial Intelligence, with the risk of being prone to speculation, in this thesis I adopt a perspective grounded on the economics of innovation to understand why AI technologies are considered the next great revolution in technology, with consequences that can exceed those of ICTs both in scope and length.

Chapter 1 is focused on introducing AI technologies. Several problems occur when trying to define AI, above all, the problem of defining what intelligence is in the first place. Part of the mainstream debate of the AI field is presented, such as the distinction between Artificial Narrow Intelligence (ANI), a system designed to perform a limited and very specific task, and Artificial General Intelligence (AGI), a system that can perform a variety of different tasks with originality, capable of autonomously identifying problem and solution. While the development of ANI is a well-established practice in the AI community, scholars do not envision that AGI can be reached in the short-medium term, and some even affirm that it is not possible at all. The economic community considers contemporary AI as a merely improvement in *prediction technologies* (Agrawal et al., 2018b), an essential component in the decision-making process. Treating AI as prediction is essential to understand why this technology can be applied in a variety of fields, and thus why AI should be considered as a General Purpose Technology. This is followed by the presentation of a brief history of AI, from their initial formulation in the 1940s to the more recent developments, with the business success of machine learning and deep learning, distinguishing between symbolic systems, such as expert systems, and sub-symbolic systems, such as machine learning. Finally, the chapter focuses on some functional applications of AI, aimed at presenting some of its broad range of applications and limitations.

As Drezner (2019) affirmed, much international relations scholarship treats technology as an exogenous shock, an independent variable that changes the dynamics of international politics, shifting the balance of power between one pole to another. However, this perspective does not consider that this relationship works both ways, with specific institutional settings affecting the direction and rate of R&D efforts, influencing the trajectory and the pace of innovation. Chapter 2 focuses on policy tools that favor or discourage innovation: Intellectual

Property Rights (IPRs). Several reasons were proposed to justify the existence of IPRs, either philosophical, such as the natural rights argument and the desert argument, or economical, such as the utilitarian argument. With the awareness that IPRs impose strong externalities on the economy, a consensus was reached within the academic community that the introduction of IPRs makes sense only when they increase the economic welfare. This perspective has the advantage that does not take for granted the introduction of a monopoly right, as temporary as it may be. Unfortunately, the construction of the legal framework surrounding IPRs did not always take into account this perspective, and in many cases was influenced by outsider interests and by the cumulative evolution of legislation. Nevertheless, the economic analysis treats IPRs as a tool to incentivize innovation by awarding a temporary monopoly right to the innovator, striking a balance between society's dynamic economic welfare (that is more well-off if an innovation has taken place) and the static economic welfare (that decreases after the introduction of a monopoly right). In the second part of the chapter, the three IPRs involved with AI will be presented. While the main focus will be on patents (to pose the basis for the empirical analysis of chapter 4), copyright and the database sui generis right will also be treated. The remaining part of the chapter focuses on whether it makes sense to introduce additional IPRs specific of AI technologies. This reflection is needed because many legal scholars argued in favor of creating additional IPRs to box AI technologies and their by-products in current legislation. I argue that there is no economic evidence for such action, except in very limited and circumscribed cases, since it would decrease the economic welfare. Moreover, this would lead to the creation of a paradoxical situation where the technologies have to adapt to the challenges of dealing with an over-complex legislation and not the opposite. Altering the current legal framework of innovation would impose changes to an environment that has no need for incentives for innovation, potentially hindering current incentives and slowing down economic growth.

The third chapter focuses on how AI technologies will impact the innovation process and the economics of innovation. First, a theoretical framework surrounding the classification of technologies in different classes will be presented. *General Purpose Technologies* (GPT) will be analyzed in detail, since they have a broad impact on the economy, providing econometric models that describe the dynamics of important externalities caused by innovation complementarities. GPTs are generally considered the origin of virtuous cycles of innovation, in which innovations in one application sector leads to an increase in the rate of innovation in other contingent sectors, where coordination acts as an enhancing factor. Another important categorization is the one first formulated by Griliches (1957), of *Invention of a Method of Inventing* (IMI), which is the invention of a new technique for achieving scientific discoveries or technical innovations, which increase the productivity of R&D departments. The existence of technologies that can be classified as both GPT and IMIs has led the literature to draft a new category: *General Methods of Inventing*, that can be used to define a technology that has both functional applications and research capabilities across a large number of fields. These technologies are extremely rare, the literature agrees that digital computing and, possibly, AI may be classified as such. However, even if there is an agreement on the fact that GMIs have a deep economic impact, their inner dynamics are still unknown. The second part of the chapter will be focused on providing qualitative evidence regarding whether Artificial Intelligence should be considered a GPT (or even a GMI), exploring whether AI possess the three characteristics of GPTs (rapid growth, pervasiveness, and strong complementarities), and its role in the contemporary research process. The use of AI technologies for research and discovery is treated in detail, with a focus on the advantages of Machine Learning for theoretical modeling. Finally, the chapter concludes by examining some of the impacts that AI technologies have on market structure. The digital market is characterized by extreme returns to scale and strong network effects, which give rise to the economies of scope that provide an competitive advantage to incumbents, especially in terms of innovative capabilities. While this has beneficial effects on innovation, it also provides enormous prescriptive power on the way it is conducted and its direction, accentuating the process of market concentration, as suggested by Crémer et al. (2019). Moreover, the use of AI technologies in the market has several negative externalities, that favor producers over consumer, while creating new issue areas for regulatory authorities, such as algorithmic collusion.

Chapter 4 is centered around finding quantitative evidence in patent data regarding the technological categorization of Artificial Intelligence as a GPT and the process of market concentration. First, the literature on empirical analysis of GPTs and AI will be reviewed in detail, in particular the one regarding the strategies used for identifying AI-related patents and the indexes used to measure the General Purpose properties of a technology. To this end, I sampled the 2018 edition of PATSTAT, the Worldwide Patent Statistical Database maintained by the European Patent Office. I examined all patent applications filed through the Patent Cooperation Treaty procedure and I identified the AI-related patents by means of a mixed strategy based on both technological classification codes and keywords search. The first part of the analysis focuses on exploring different GPT indicators, both coming from the literature and innovative approaches based on network science. The use of network-based indicators was based on the theoretical modeling of technology evolution provided by Korzinov and Savin (2016), that considers a GPT a technology that is complementary to other technologies based on the number of cliques that it manages to be part of. The results of the analysis confirm that AI should be considered a GPT, since it is rapidly spreading in many areas of the economy and connecting with a range of different technologies to produce new products. The second part of the analysis is instead focused on verifying the claim of chapter 3 regarding the progressive concentration of AI-related patents in the portfolios of a small number of applicants. AI-related patents were broken down by filing year and the applicants were divided in asymmetrical classes to distinguish between strong and weak patentees, confirming that the proportion of patents filed by strong patentees is increasing, while weak patentees are becoming less and less relevant in patenting activity.

Proof that AI is indeed a GPT increases the evidence that this technology will have a profound impact on society, altering the already fragile dynamics between actors at all levels, from local to international. The most visible consequences will likely regard its effects on employment, and the concentration of property assets involving AI technologies would reduce the capabilities of unemployed workers without providing them the possibility to compete with incumbents. Given their general purposeness, imposing temporary monopolies on the use of such technologies risk of impeding the innovative abilities of new entrants. Rather than proposing new IPRs to explicitly target AI technologies, legal scholars should focus on deescalating the universe of IP regimes that hinder the dynamic AI market, favoring SMEs over oligopolies to reduce the negative externalities provoked by digital giants, which too often dodge responsibility, lack in transparency, and provide no effective opportunity of representation to their clients, while having a normative power comparable (and, in some cases, superior) to the one of states. We are at the dawn of the AI technological revolution and the rate of change is only going to increase.

Chapter 1

Artificial Intelligence

Artificial Intelligence technologies (AI) are increasingly becoming part of our daily life. They have become commonly used in academia and the business world alike, in a large range of fields, allowing the application of notions from both natural and social sciences. While AI is based on mathematics and computer science, its recent successes derive from a combination of three fundamental factors: the increasing availability of large amounts of data, growing computing power, and new and more efficient algorithms.

1.1 What is AI?

Although there are many different definitions of what exactly is the aim of AI studies[1], essentially the development of AI systems focuses on building machines capable of performing tasks that typically require human intelligence. On a more practical level, when we speak of modern AI systems we generally refer to prediction machines. AI systems have no cognitive state: (Searle, 1990) they simply are machines that, given some inputs, return some more or less determined outputs. We generally speak of AI because these machines are capable of returning accurate answers even when provided limited and partial information, thus predicting. AI systems exploit previously collected data to create a statistical model of the world capable of predicting an output corresponding to our expectations with a reasonable amount of accuracy. In this framework, intelligence (in its general meaning) is not involved at all.

In this context, some experts distinguish between Artificial Narrow Intelligence (ANI) and Artificial General Intelligence (AGI). ANI, or weak AI, involves AI systems specialized in performing a particular activity, such as playing chess, recognizing images, or transcribing a speech. As we are all aware, the diffusion of such systems has rapidly made these instrument parts of the essential toolkits of our daily life, automating repetitive tasks and increasing efficiency.

On the other hand, with AGI, or strong AI, scholars identify to a machine capable to autonomously perform any intellectual task that a human can perform, in other words, an artificially intelligent system capable of

[1] Russell and Norvig (2010) identify four main definitions of artificial intelligence: (1) systems that think like humans; (2) systems that act like humans; (3) systems that think rationally; (4) systems that act rationally. According to Haenlein and Kaplan (2019), this lack of agreement is due to two factors: first, it would require agreeing on how to define human intelligence; second, there is little evidence that proves how much machine intelligence resembles human intelligence.

achieving a broad and general objective, using a model of reality to make predictions and plan actions (Flowers, 2019). Many scholars, above all Searle (1990) and Dreyfus (1972), think that this objective is not achievable, nor desirable.[2]

1.1.1 An economic perspective of AI

Since modern AI technologies represent a drastic improvement of prediction technology (Agrawal et al., 2018b), let us consider the role of prediction in the decision-making process. According to Agrawal et al. (2018b), the decision making process can be split into the stages of prediction and judgment, where prediction is the act of mapping the possible outcomes of a variety of actions and judgment is the act of choosing the most well-suited path to reach an objective (Agrawal et al., 2018a).

In economic terms, given a predetermined state S, prediction consists of mapping all the set of actions X that can undertake, and their respective outcomes. Judgment, on the other hand, consists of the process of evaluating the outcome obtained from choosing an action χ from a set of possible actions X. The payoff for each action can be defined as a utility function μ (χ, θ), where θ is the realization of an uncertain state drawn from a distribution F(θ) (Agrawal et al., 2018a). The decision-making process is thus reduced to a *"standard problem of choice under uncertainty"* (Agrawal et al., 2018a) where prediction maps the likelihood of possible outputs and judgment ranks the desirability of each output. Since it is assumed that utility can only be determined by humans, who have to undertake a costly process that allows the mapping from (χ, θ) to a specific payoff value μ, the utility function is unknown to the AI. Evaluation requires time and effort, and it is strictly linked to the operator's final goal, which is not easily translatable to a machine.[3]

AI technologies thus have substantially reduced the costs associated with prediction. However, while in some cases this has led to the full automation of simple tasks, in others is unlikely that this will lead to substitution of human users, rather to the enhancement of human capabilities. In economic terms, this equals asking if judgment and prediction are complements or substitutes. Agrawal et al. (2018a) created an model to determine whether prediction can reach a level of accuracy able to entirely take over the judgment part, concluding that they are complements, provided that judgment is not too difficult. When complexity is introduced, *"the impact of improved prediction on the value of judgment depends on whether improved prediction leads to automated decision making"* (Agrawal et al., 2018a). As complexity increases, there is a tendency for humans to opt for default decisions and heuristics that, on average, perform well. This aspect reduces the economic value of judgment and increases the one of prediction at the margin. Agrawal et al. (2018a) analyzed complex problems such as automation, contracting, and firm boundaries. They suggest that the possibility of introducing automated decision-making through prediction largely depends on the specific features of the tasks, with the value of judgment increasing when the cost of mistakes is particularly high, as in the case of applications of AI to tasks that have a great impact of people's life, such as the automatic detection of tax fraud.

[2] Leaving aside the risks of such systems, there are important ethical implications with no easy answer.

[3] As Agrawal et al. (2018a) affirmed, *"We assume that this process of determining payoffs requires human understanding of the situation: it is not a prediction problem"*.

1.2 A Brief History of AI

As a formal science, AI was born less than 80 years ago, drawing from a variety of different fields, such as philosophy, mathematics, linguistics, psychology, economics, and many others (Russell and Norvig, 2010). Providing a complete review of the topic goes beyond the scope of the dissertation, therefore in the following pages, I will only present a brief summary of the milestones of modern AI.

1.2.1 Early developments: 1940s – 1950s

In 1943, McCulloch and Pitts published the first work that is now generally considered AI (McCulloch and Pitts, 1943). By combining the basic physiology of neurons, the formal analysis of propositional logic, and Turing's theory of computation, they demonstrated that networks of connected neurons could compute any computable function. This work is notably the forerunner of both the logicist and the connectionist tradition of AI. Subsequently, other scholars started working on these concepts, and in 1951 Minsky and Edmonds built the first neural network computer (Russell and Norvig, 2010). In 1950, Alan Turing published the seminal paper *"Computing Machinery and Intelligence"*, where he describes the process of creating intelligent machines and proposes a method of testing machine intelligence through a "game of imitation", where a human should be able to distinguish a human from an AI in a teletype dialogue. [4]. In 1956, John McCarthy organized a workshop at Dartmouth College that brought together all the significant figures of the growing AI field. While the workshop itself did not lead to any notable discovery, it made all these researchers know each other, fundamentally contributing to the evolution of the AI field (Russell and Norvig, 2010).

1.2.2 High expectations and the first AI winter: 1960s – 1970s

In the following years, the AI field experienced the first successes, such as the creation of the General Problem Solver (GPS),[5] the development of the high-level language LISP, and the invention of time-sharing systems for optimizing the use of computing resources (Nilsson, 2009). These early programs often contained little or no knowledge of their subject matter and often succeeded in using simple syntactic manipulations (as in the firsts machine translation efforts). It was a general belief that solving more complex problems was only a matter of faster hardware and larger memories, an assumption proven wrong afterwards (Russell and Norvig, 2010). This progressively led to what has been described as the first AI winter. The high expectations that the public had for AI gradually faded away, and in 1974, criticism by Sir James Lighthill on AI technologies added to the pressure from public research institutions to fund more productive projects, leading to the cut of AI exploratory research by the US and British governments (Gonsalves, 2019).

[4] A method that was later renamed as the "Turing Test"

[5] A system that applied human problem-solving protocols to solve simple problems

1.2.3 Expert systems 1980s – 1990s

The development of the first microprocessors at the end of the 1970s renewed interest in AI research, leading to the development of the DENDRAL program, an expert system that capable of inferring the molecular structure based on the information provided by a mass spectrometer (Nilsson, 2009). Its expertise derived from large numbers of special-purpose rules collected through the extensive interviewing of domain experts. In the 1980s, expert systems gained commercial success, and by 1985, the AI market had reached a value of over a million dollars (Russell and Norvig, 2010). Moreover, when Japan launched the Fifth Generation Computer Project, US and UK researchers feared that Japan could establish supremacy in the field and managed to generate a similar investment in the US (Ray, 1998). At the end of the 1980s, the first limitations of expert systems emerged. Their programming was extremely complex, and as the number of rules increased, a "black box" effect made the functioning of the machine hard to understand. Development and maintenance were challenging, and therefore faster, easier, and less expensive ways to achieve the same results progressively took over (Russell and Norvig, 2010). The last success of expert system was Deep Blue, IBM's AI that in 1997 beat the world chess champion Garry Kasparov. Nevertheless, as all expert system, it was based on a brute force algorithm that evaluated each move based on a set of rules extracted from chess books and previous grandmaster games (Campbell et al., 2002), and marked the abandonment of rule-based AI systems. Even a relatively simple task such as playing chess required setting a complex set of parameters, limiting their applicability to the vast complexity of the world (Nilsson, 2009).

1.2.4 Data + Computer power = Machine Learning 2000s – 2010s

In the late 1990s and early 2000s, increased computational power and data availability gave new impetus to AI research. A focus on the resolution of closed and specific problems and new ties between AI and other fields led to increased use of AI for logistics, data mining, and medical diagnosis: AI entered into the mainstream technological use (Haenlein and Kaplan, 2019).

The development of the internet and the dramatic increase in data availability drastically improved machine learning algorithms' performance. Unlike expert systems, where the programmer had to explicitly tell the machine how to act in individual situations, machine learning systems build a mathematical model based on input data. While the first learning systems were developed in the early 1960s, their performance improved only after large sets of training data became available (Nilsson, 2009). Machine learning is not about writing a set of instructions to perform a specific task but rather letting the machine determine the best prediction model for a particular problem, based on massive amounts of data and a limited number human-set parameters. Quality training sets became more important, marking a paradigm shift from the previous approach that focused on algorithms optimization (Russell and Norvig, 2010). The larger and more complex the training set is, the more computing power is required. Computer graphics card processors (capable of performing multiple parallel computing, called GPUs) allowed to build affordable ML models and, therefore, increased AI researchers' data processing capabilities (Lee et al., 2018). In the past years, machine learning has been used for performing more complex tasks, within the framework of narrow and clearly defined problems, such as board games, text classification, and image recognition. In 2011, Watson, the question-answering AI system powered by IBM, repeatedly won against two "Jeopardy!" champions (Ferrucci, 2012), while in 2012 Google X finally managed to have an AI that can recognize various objects (such as cats and dogs) in videos without any explicit programming and in 2016, AlphaGO beat the world champion of the Chinese game Go (Borowiec, 2016). More recently, deep learning (a subfield of machine learning based on multiple layers of artificial neural networks) has increased in popularity among AI engineers (Goodfellow et al., 2016). Deep learning is now commonly used for machine translation, image recognition, drug design, and many other fields, but it still presents some limitations, such as a "black box" effect that prevents human engineers from understanding how and why a final model does or does not work.

1.3 Two different approaches to AI: symbolism and connectionism

Since the beginning of the formal study of AI it is possible to discern two different schools of thought which influenced how the AI field developed: Symbolism and Connectionism.

The symbolist tradition of AI is based on the idea that intelligence can be achieved through the manipulation of symbols[6] (Newell and Simon, 2007), where symbols are a high-level abstraction of objects, problems and logic. They provide the AI with an already codified representation of knowledge, using a top-down approach. A typical example of such systems are expert systems, which emulate a human expert's decision-making ability by codifying it in a series of if-then rules by drawing from a previously acquired knowledge base[7]. These systems present the advantage that once the basic shell is developed, it is possible to instill knowledge through a format that is easily comprehensible to domain experts, reducing the cost of intervention of IT experts. Additionally, it is possible to obtain a prototype in a relatively short amount of time. However, the process of knowledge acquisition to achieve a valuable product may result in being extremely long (Greene, 1987), and as the size of the knowledge base increases, the harder it becomes to verify the consistency of the decisions, making prioritization and ambiguities common challenges of sizable expert systems. Finally, when it becomes necessary to update the knowledge base, the risk of disrupting the system persists, since every change in if-clause statements might potentially endanger the whole system (Partridge, 1987). Some scholars even affirmed that expert systems that are not able to learn should not even be considered AI (Schank, 1983; Waltz, 1983). While many agree with this proposition, they represented a milestone in the development of the field, and are generally considered as AI for historical reasons (Partridge, 1987).

The connectionist (or sub-symbolist) tradition adopts another approach, affirming that to achieve intelligence AI systems need to mimic the functioning of the human brain rather than just manipulating abstractions. In the words of Searle (1990): *"a computer has syntax but not semantics."* Sub-symbolic systems have a bottom-up approach, starting from the lowest level and building intelligence by *connecting the dots.* Machine learning (ML)

[6]This is also called the "physical symbol systems" hypothesis (Newell, 1980)

[7]From a structural perspective, they are divided into two subsystems: the knowledge base (that contains facts and rules) and the inference engine (that combines rules and known facts to deduce unknown facts).

is an example of such technologies, which allows AI systems to learn from examples without hard-coding the instructions, but rather adjusting a set of parameters in order to achieve a desirable outcome. ML optimizes a statistical model defined up to some parameters based on example data or past experiences, also called training data (Alpaydin, 2004). The model can be either predictive, descriptive, or both. Assuming that the data is accurate and the model was constructed correctly, ML systems are able to detect patterns and schemes with acceptable approximation, sometimes even outperforming human experts (Grace et al., 2018). Performance greatly relies on the training set: the more examples a model is given during training, the more information it has when it comes to making accurate predictions and correctly classifying unseen cases (Russell and Norvig, 2010; Jordan and Mitchell, 2015).

The main techniques used to build ML models can be classified in supervised, unsupervised, and reinforcement learning (Alpaydin, 2004). In supervised learning, the training data is labeled, meaning that a sample of input vectors (inputs) is coupled with a sample of corresponding target vectors (outputs).[8] Unsupervised learning instead focuses on identifying patterns in the training set where no labeling or feedback is provided (Dey, 2016). It is mainly used in cluster analysis, detecting *"potentially useful clusters of input examples"* (Russell and Norvig, 2010). The underlying assumption is that the structural properties (algebraic, combinatorial, or probabilistic) of data might provide insights essential to the prediction (Jordan and Mitchell, 2015). Finally, reinforcement learning maximizes a payoff function based on the actions or decisions that the AI system takes. labeled training examples are provided to indicate whether a decision in an unknown dynamic environment is correct or not for a given input (Jordan and Mitchell, 2015).[9] While this classification is helpful to understand the dynamics underlying machine learning, in most real-life applications, AI models are often trained using mixed strategies (generally referred to as semi-supervised learning).

Recently, Deep Learning, a ML technique that uses multiple layers of artificial neural networks, has proven effective in solving many complex problems. An artificial neural network is a collection of connected artificial neurons that mimic the behavior of biological neurons. An artificial neuron is a simple processing unit designed to be trained from data (through unsupervised induction) (Henderson, 2010). In an artificial neural network, each neuron receives an incoming signal and transmits it to the neurons it is connected to. Typically, neurons are aggregated into layers that may perform different transformations on their inputs (Russell and Norvig, 2010). These kinds of machine learning structures have been applied to many fields such as computer vision, speech recognition, natural language processing, audio recognition, and board game programs, where they reached results that could be compared to human experts and that, in some cases, outperformed them (Pouyanfar et al., 2018). Unfortunately, training this kind of structure is very complex, and deep learning algorithms often operate as "black boxes", since their complexity makes it difficult to ML engineers to understand their functioning (Castelvecchi, 2016).

[8]Common algorithms used in supervised learning are linear and logistic regression, naïve Bayes, support vector machines, and K-nearest-neighbor (Kotsiantis, 2007).

[9]An example of the output of a payoff function might be game scores: they can be given at every step, as in a ping-pong game, or only in the end, as in chess (Russell and Norvig, 2010).

1.4 Some AI capabilities

Often, media and popular culture depicts AI systems in the form of gigantic supercomputers, which seem far from our human experience. However, reality contradicts these colorful narratives. Different applications (such as search engines, microtargeting advertisement, news aggregators, recommendation systems, speech recognition, and so on so forth), products and services that use AI technology are entering our daily lives, providing alternative solutions to traditional tasks. With no pretense of completeness, I retained useful to provide a non-exhaustive list of different AI capabilities, along the lines of human abilities: seeing, listening, understanding, thinking, and moving.

1.4.1 Seeing: Computer Vision

Computer vision is a field of AI that aims to understand the content of visual inputs. It may involve extracting a description, either in the form of a text, an object, or a three-dimensional model. In other words, *"at an abstract level, the goal of computer vision problems is to use the observed image data to infer something about the world"* (Prince, 2012). Computer vision technologies acquire, process, analyze, and understand visual inputs to extract high-dimensional data from the real world to produce numerical and symbolic information (Rosenfeld, 1988).

The main tasks operated by computer vision systems are recognition, motion analysis, scene reconstruction, and image restoration. Image recognition determines whether or not image data contains a specific object, feature, or activity (Forsyth and Ponce, 2012). Applications of AI spans from optical character recognition (OCR) to pose estimation and facial recognition (Nilsson, 2009). Motion analysis instead processes an image sequence (a video) to estimate the direction and velocity of an object (tracking) or the camera itself (ego-motion). When tracking and ego-motion technologies are combined, they are generally referred to as optical flow (Russell and Norvig, 2010). Scene reconstruction deals with the construction of a model of a 3-dimensional space. Its application ranges from crime reconstructions to the digitization of cultural heritage and geospatial mapping (Trucco and Verri, 1998). Finally, image reconstruction is used to restore old or damaged visual content, removing noise, and augmenting quality (Banham and Katsaggelos, 1997). In this way, it is possible to obtain more precise image data that can later be subjected to human analysis or other computer vision techniques.

1.4.2 Listening: Speech recognition

Similar to image recognition, speech recognition systems transcribe speech to text. Audio signals are cleaned and processed to isolate the frequencies that represent a human voice. Then, with the help of linguistic models, a final text is produced (Ashri, 2020). Speech analysis can provide additional information regarding the speaker, such as identity, emotional state, health status, accent, and gender (Nassif et al., 2019). Speech-to-text technology represents a crucial asset for transforming orally recorded data into text data that can later be analyzed using Natural Language Processing (NLP) techniques, further diminishing information retrieval's transaction cost. Unfortunately, the performance of speech-to-text technologies is subjected to high variance depending on the settings in which the recording takes place. Various studies show how even when dealing with systems that have a high-quality output in controlled environments, if disturbing elements[10] are introduced, their performance goes below the human benchmark.[11]

1.4.3 Understanding: Natural Language Processing

Natural Language Processing (NLP) is a field of AI concerned with the interactions between computers and human (natural) languages, particularly with how to program computers to process and analyze large amounts of natural language data (Chowdhury, 2020). In the first days of NLP, the typical approach was based on expert systems, where grammatical features and heuristics were hard-coded (Nilsson, 2009). With the introduction of machine learning techniques the performance of linguistic models drastically improved. NLP is now used in fields such as *"machine translation, natural language text processing, and summarization, user interfaces, multilingual and cross-language information retrieval (CLIR), speech recognition, artificial intelligence, and expert systems, and so on"* (Chowdhury, 2020).

Whoever attempts to build software capable of extracting and manipulating natural language generally encounters three significant issues: the extraction of the thought process, the storage of the representation and meaning of the linguistic output, and the contextualization of information (Chowdhury, 2020). As a consequences, approaches may vary: some are based on lexical and morphological analysis (Part-Of-Speech tagging), others on semantic and discourse analysis (Semantic Role Labeling, chunking), and knowledge-based approaches (Named Entity Recognition). On a more practical level, NLP systems generally comprehends a combination of all of these, starting from the word level, extending it to the sentence level, and finally framing it in the context of the specific domain (Chowdhury, 2020).

Some typical NLP applications that are currently being investigated are concept extraction, machine translation, question answering, and natural language generation (Clark et al., 2013). Among those, concept extraction is the most problematic: while in some cases NLP systems showed promising results in very restricted domains, we do not know yet how to correctly compute the meaning of a sentence based on words and context (Chai et al., 2001). Machine translation is one of the earliest applications of NLP, but it still presents several complex problems since human language is ambiguous and is characterized by a large number of exceptions. Nevertheless, recently machine translation has reached a stage that allows people to enjoy its benefits (Chowdhury, 2020). Even if it is not always perfect and the translations are not as good as humans', the initial results are very encouraging, and machine translations may serve as a way for human translators to speed up the process and improve their performance. Question-answering systems and natural language generators are showing promising results [12].

Despite these advances, three challenges prevent NLP from reaching commercial success: scalability, porta-

[10] Such as noise (Gupta et al., 2016), foreign accents (Eskenazi, 1999), and children's voices (Potamianos et al., 1997).

[11] Poor performance may also depend on training with insufficient data of adverse situations. It is possible that if noisy data is provided during the training process, the model performance will improve (Catania et al., 2019).

[12] In particular, the autoregressive linguistic model GPT-3, developed by OpenAI and introduced in May 2020, with an incredibly large number of 175 billion parameters, can generate fiction, poetry, newspaper articles, programming code, and probably much more. At the time of writing, this kind of technology might prove to be a game-changer in NLP development.

bility and variability. NLP systems establish patterns that are valid only for a specific domain and a particular task. When the topic, context, or user, change, it is necessary to create entirely new patterns. Second, advanced NLP techniques such as concept extraction are too computationally extensive for large scale NLP applications (Sparck Jones, 1999). Third, human behavior and communication patterns are erratic and constantly evolving, while the NLP system requires extensive and stable corpora to produce effective results (Chowdhury, 2020).

1.4.4 Strategic Thinking: AI Planning

A fundamental functional application of AI system is AI planning, which aims to build systems capable to design the set of actions to perform to achieve a desired goal. Planning involves the introduction to AI systems of concepts such as time, causality, intentions, uncertainty, and multiple agents' actions. The classical formulation of the planning problem requires three inputs: a description of the state of the world, a description of the agent goal, and a description of the possible actions that can be performed (also called domain theory). The planner then outputs a sequence of actions designed to achieve that goal (Weld, 1999).

Recently there has been a peak in the interest in what has been called domain-independent planning (when no input domain is specified). Now machine learning algorithms are used to improve planners' speed and quality performance, but there are still many challenges. The high uncertainty of the domain makes *"completing a plan a significantly more difficult task than computing one"* (Leonetti et al., 2016). Relevant details and dynamics might be misinterpreted, creating imperfect models. While current systems manage to scale up antecedent problems, planning in uncertain situations still needs further research to reach acceptable results (Jiménez et al., 2012).

1.4.5 Moving: Robotics

While the previous AI applications can entirely be performed in digital environments, robotics involves the interaction of AI systems with the physical world, where uncertainty and external actors increase complexity (Brady, 1985). It is important to stress that robotics does not necessarily involve AI: we can distinguish between intelligent and non-intelligent robots. Non-intelligent robots are mechanical devices that perform operations based on instructions that are either hard-coded into their systems or transmitted to them by humans through a telecommunication infrastructure, while an intelligent robot is *"a mechanical creature which can function autonomously"* (Murphy, 2000). While in the early days of AI, robotics was considered an integral part of the field, it progressively diverged over the years. Robotics focused more and more on the manufacturing industry and the assembly line's automation, which required no intelligence to function. Some contact points remained, but it was relegated to applications where humans could not efficiently communicate with robots in any way, such as in space explorations (Murphy, 2000). Recently, advances in both fields have brought renewed interest in closing the gap between them. Robots are increasingly introduced in *"less engineered and more open environments. To do so, they need to rely on cognitive capabilities typical of AI, such as knowledge representation, learning, adaptation, and human-robot interaction "* (Rajan and Saffiotti, 2017).

Increase autonomy of robotic systems may prove to be particularly useful in applications where humans are at significant risk (such as in space, military, or health threats) or in trivial, physically harsh, and unpleasant tasks (such as in the service industry or agriculture).

On the other hand, intelligent robots' introduction raises ethical concerns of a different scale than other AI applications, especially in robot-human interaction. The deployment of Lethal Autonomous Weapon Systems (LAWS) has increased concerns in the international community. Still, even apparent innocuous intelligent robots such as autonomous vehicles present ethical dilemmas in contexts where there is neither a human nor time available to provide an answer (e.g. in the context of a car crash). In these cases, the determination of liability is not trivial and there is no shared position in the academic community (Lin, 2011).

1.5 Conclusions

In this chapter, I presented AI technologies from both an historical and a technical perspective. In particular, I clarified a common misconceptions regarding AI technology: AI systems are not intelligent machines and even if the development of AGI prove to be possible, it is not likely it will happen in the short or medium term. While AI systems may seem to perform intelligent actions or to reason intelligently, they do possess an internal cognitive state regarding the actions they are performing (Searle, 1990), making them ontologically indistinguishable from other software. The recent advances in AI can be circumscribed as improvements in prediction technologies, one of the critical elements of every decision-making process, but certainly not the only one. Framing AI as a prediction technology is key to understand why and how AI technologies are becoming pervasive in the economy. The progressive digitization of society (which improves the process of data gathering across various dimension of the human experience) increases the opportunities to apply AI technologies in the most diverse areas, thus expanding the range of possible innovations.

Chapter 2

Intellectual Property Rights

Intellectual property law is one of the main policy instruments used to guide the impact and direction of innovation efforts. Formally, its goal is to encourage innovation and creation of intellectual goods by facilitating the innovators' appropriation of the derived benefits. Following the rise in the rate of innovation that marked the 21st century, further enhanced by the digital revolution, IPRs have been increasingly contested by different stakeholders. IPRs incentivize or deter specific innovative behaviors, playing a fundamental normative function in determining the rate and trajectory of innovation efforts.

First, I will present the foundations of IPRs, starting from the philosophical and economic justification for creating a property right in intangibles. Second, I will present some of the IPRs involved in Artificial Intelligence, focusing on the economics of patents and presenting the legal framework surrounding copyright and the database sui-generis right. Finally, I will examine whether the claims regarding the potential introduction of new IPRs to AI technologies are justified from an economic perspective.

2.1 Foundations of Intellectual Property

Spence (2007) defined intellectual property rights as a *"right that can be treated as property to control particular uses of a specified type of intangible asset"*, suggesting IPRs protection covers only specific kinds and uses of intangible assets. In other words, the property right does not directly involve the intangible asset but rather the right to exclude others to make specific uses of it.

2.1.1 Justifications for the existence of intellectual property rights

A complete review of the philosophical and economic justifications of existence of IPRs would go well beyond the scope of this dissertation, thus I will focus on the three most common explanations: the natural rights argument, the desert argument, and the utilitarian argument, which has traditionally been the most influential in constructing a legal framework for IP [1].

[1] For a detailed study on this topic, see Spence (2007) and Menell (1999).

The natural rights argument

The natural rights argument, first expressed by John Locke in the Second Treatise of Government (1689), is based on the assumption that people are naturally entitled to the fruits of their labor, including the fruits of their intelligence. He affirmed that: *"The labour of his body and the work of his hands, we may say, are strictly his. So when he takes something from the state that nature has provided and left it in, he mixes his labour with it, thus joining to it something that is his own; and in that way he makes it his property"* (Locke, 1689). However, this approach to IPRs cannot be considered valid on at least two different basis. First, it justifies control over the elements of an intangible asset for which a creator is responsible. However, intangible assets are very problematic because it is impossible to determine which assets the creator is genuinely accountable for. Nozick (1974) presented a counterargument by comparing newly produced knowledge to tomato sauce spilled in the ocean. If somebody owns the tomato sauce (the intellectual labor) and drops it into the sea (the entire global knowledge), is it right for him to claim the ocean as a whole as his? Locke himself provided the second objection to the natural right argument. He assumed that this appropriation system could be applied only when appropriation does not leave anybody else worse off, which is not always the case. Even when a creator can claim a specific asset's origin, he is not necessarily entitled to control its use. Imposing a normative claim would preclude the user's autonomy and, therefore, make him worse off, invalidating the appropriation process.

The desert argument

Another commonly used justification of IPRs is that an intangible asset's creator deserves to benefit from his labor. This is generally referred to as the *desert argument.* On the other hand, this argument does not consider that property rights are not the only way to reward a creator. As Hettinger (1989) points out, laborers may perform intellectual work only for the end of performing it, such as genuine interest in that field of knowledge, society's progress, ethical reasons, and so on so forth. In those cases, other possible rewards are recognition, gratitude, or public financial support to continue pursuing their work. IPRs are only one of several possible means to provide incentives to intellectual laborers, and exclusive property rights may impose an unnecessary cost to society.

The utilitarian argument

The utilitarian approach is the one that has historically prevailed, and most contemporary legal and economic arguments regarding IPRs are based on this paradigm. Unlike other arguments, the utilitarian approach in the past was supported by policy objectives and allegations such as *"industrial progress is desirable, inventions is a necessary condition of industrial progress, not enough inventions will be made or used unless effective incentives are provided"* (Machlup and Penrose, 1950). It is based on the economic theory of public goods (Samuelson, 1954): rather than focus on the creator, its primary concern is the community as a whole, and it aims to improve the allocation of resources to maximize the benefits of society through the analysis of economic welfare.

Greenhalgh and Rogers (2010) affirmed that *"if a firm cannot charge all the beneficiaries of its innovation, then there is a problem of matching incentives to the value of the activity, which may lead to an undersupply of innovation."* In economics, this is commonly referred to as *market failure*. In the case of the allocation of benefits related to the production of knowledge, market failure may occur when knowledge is either described as a pure public good or as a private good with positive externalities.

2.1.2 The economics of intellectual property

Knowledge as a public good

According to Arrow (1962) and Stiglitz (1999), free of any artificial construct, knowledge is both non-rivalrous and non-excludable, thus making it a public good. Public goods are those *"which all enjoy in common in the sense that each individual's consumption of such a good leads to no subtractions from any other individual's consumption of that good"* (Samuelson, 1954).

A non-rival good in consumption means that the consumption of the good x by an individual A does not preclude the consumption of the same good of another individual B, thus making the marginal cost of consumption equals to zero. Common examples of public goods are sunlight, radio transmission, public roads, and national defense. Public goods are often defined in contrast to private goods, which are those goods y whose consumption from an individual A precludes the usage of another individual B, such as food, energy, or drinkable water. For example, an apple is a private good because if person A eats it, it prevents the usage of person B. Conversely, sunlight is non-rivalrous because both A and B can enjoy it without precluding the other from consuming it. When treated as a public good, knowledge has often been assimilated with information (Archibugi and Filippetti, 2015). When somebody consumes information, it does not reduce the quantity available to other individuals, making it non-rivalrous. While knowledge may be costly to produce, the marginal cost of sharing information with an additional individual is zero (Lévêque and Ménière, 2004). For example, there is a zero cost of providing the notion of a mathematical theorem to an additional individual. However, as pointed out by Greenhalgh and Rogers (2010), while the use of knowledge from one actor does not prevent other actors from using it, its diffusion may exhaust the profits that can be obtained from it[2]. While information in itself is non-rival, the value associated with it can be rivalrous.

The other fundamental characteristic of public goods is that they are non-excludable, which means it is challenging (if not impossible) to exclude a potential consumer from using them. Recalling the previous examples, while it is relatively straightforward for A to prevent B from eating his apple, it is almost impossible for A to prevent B from enjoying the sunlight.

Several methods were used to prevent access to knowledge in the past: Archibugi and Filippetti (2015) list three of them: secrecy, access codes, and IPRs. The first method, secrecy, is widely diffused both in the military and business sectors. It relies on the underlying assumption that the best way to ensure the appropriability of the economic returns of knowledge is to prevent its diffusion. However, secrecy provides only partial protection.

[2] Imagine a method for anticipating the behavior of financial markets

	Rivalrous	Non-rivalrous
Excludable	Private goods	Club goods
Non-excludable	Common goods	Public goods

Table 2.1: Different types of goods

Practices such as headhunting, espionage, and reverse engineering can circumvent business and governmental agencies' security measures and undermine their efforts. Moreover, since secrecy guarantees appropriability by impeding information diffusion, it has the downside of hindering the possibility for third parties to engage in cumulative innovation and encourage the duplication of R&D expense, reducing the overall economic welfare. Access codes are instead technological tools that prevent the unwanted diffusion of knowledge. While they increase the difficulty of access, they do not provide total protection. A single breach in the defense systems of businesses and governmental agencies, coupled with the non-rival nature of knowledge, may idle their protective measures. Finally, IPRs are a family of legal instruments that, with specific limitations, allow their owner to exclude other individuals from enjoying the benefits of the knowledge they safeguard.

When a good is both rival and excludable, it is generally considered a private good. In opposition, when it is both non-rival and non-excludable, it is usually considered a public good. However, there are also other dimensions, such as when products are non-rivalrous but excludable or when they are rivalrous but non-excludable. The former are generally referred to as club (or network) goods while the latter are referred to as common goods (Archibugi and Filippetti, 2015). A typical example of a club good is on-line streaming services. While the consumption by one individual of streaming services such as Netflix or Disney+ does not prevent other individuals from enjoying them, they are potentially excludable by imposing a paywall that restricts access to those paying a monthly subscription. On the other hand, examples of common goods are ocean fisheries and forestry. While the consumption of fish in the ocean or wood in forests is rivalrous because excessive fishing or deforestation can deplete resources, it is challenging to restrict access.

Knowledge as a private good with positive externalities

Opposed to Arrow (1962) and Stiglitz (1999), modern economics of science and innovation considers in his analysis the highly differentiated forms of knowledge. This perspective rejects the comparison between information and knowledge (Pavitt, 1987) because, depending on the specific traits of the knowledge transferred, the consumer may experience high transaction costs, undermining the assumption that all knowledge can be considered a public good. Callon (1994) distinguishes between freely available knowledge and knowledge that can be used without incurring costs underlined that, except for few specific cases (such as consuming a drug or using a computer program with an already known interface), the acquisition of knowledge requires additional efforts from the consumer (such as learning time or other resources). Later, Stiglitz (1999) dismissed this aspect, affirming that Arrow's definition of public goods does not consider the consumer perspective but only the producer's. However, if we include in the definition of public goods also the transaction costs bore by the consumers, knowledge may be considered either a public or a private good based on the specific kind of knowledge we examine. As formulated by Archibugi and Filippetti (2015), *"what makes knowledge different* *from public goods is not the related production process, rather its process of diffusion, which has been scarcely addressed in standard economic theory."*

The heterogeneous dimension of knowledge may influence the positioning of knowledge as a private or public good (Nelson, 1959). While basic research can be assimilated to a public good because it is possible to apply it in many different areas, knowledge regarding innovation in a product or a process directly applicable in the market can be assimilated to a private good with positive externalities (Greenhalgh and Rogers, 2010). Positive externalities arise when a producer's behaviors enhance other firms' profits, leading to a misallocation of benefits.

The traditional economic argument in favor of IPRs is that, without a mechanism of appropriation for intellectual work, sometimes creators will not have any incentive to innovate and produce knowledge, potentially leading to its sub-optimal provision. Coase (1960) suggested that this could be solved by introducing clear and defined property rights. Positive externalities could be compensated by requiring a fee from the beneficiaries, and negative externalities could be offset by charging a tax to the responsible. In the absence of appropriate incentives for innovation, then IPRs represents a way to ensure the appropriability of the benefits raised by the innovation process.

2.1.3 The traditional economic model for IPRs introduction

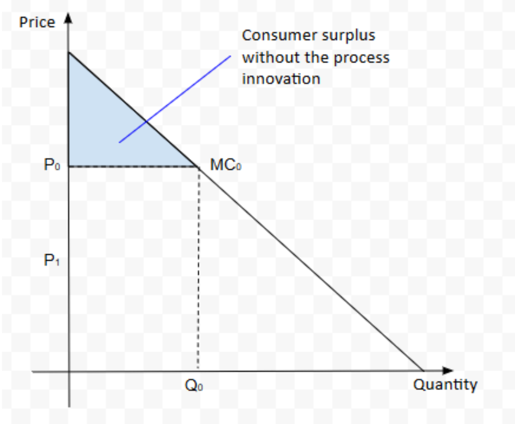
Imagine a perfectly competitive market composed of a large number of firms that produce a standardized product. The price P_0 is equal to the marginal cost of production ($P_0 = MC_0$). In perfect competition, economic welfare is entirely comprised of consumer surplus, while producer surplus is equal to zero. When a process innovation is introduced, it reduces the marginal cost of production from MC_0 to MC_1. If we consider knowledge as a non-excludable good, and the innovation can immediately be applied to the production process by all firms in the market, the increase in economic welfare involves only consumer surplus. In contrast, producer surplus remains equal to zero.

This may lead to the conclusion that, in the absence of an appropriability mechanism, producers have no economic incentive for innovation. The introduction of temporary monopolistic rights over the produced knowledge (such as IPRs), when assigned to innovators, aims to provide this economic incentive.

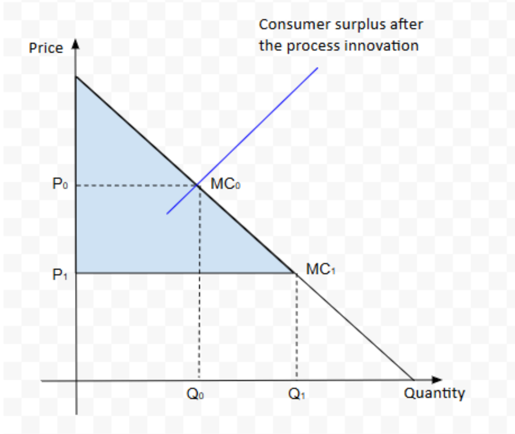
If we add IPRs to the model, its owner is the only individual that can benefit from the cost reduction created by innovation, either through direct application in the production process or by licensing it to other firms. Therefore, during the protected period, the price will remain P_0, while the marginal cost of production for the innovator will decrease from MC_0 to MC_1. The consumer surplus will remain stable, and the innovating firm will enjoy a profit equal to $(P_0 - P_1)Q_0$. However, the market will also experience a deadweight loss equal to

$$DW = \frac{(P_0 - P_1)(Q_1 - Q_0)}{2} \tag{2.1}$$

where Q_1 is the quantity of the good that would have been produced without IPRs. While the economic welfare



(a) Before the innovation



(b) After the innovation

Figure 2.1: Economic welfare before and after the innovation process

is higher than in the scenario without the process innovation, it is still lower than without IPRs, meaning that it is not Pareto efficient. However, after protection expires, the producer surplus is zeroed, and consumer surplus is maximized, leading to Pareto efficiency.

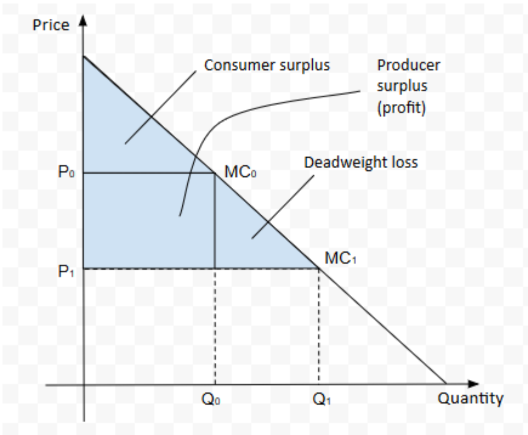


Figure 2.2: Economic welfare after the introduction of IPRs

Assuming that, in the absence of IPRs, no innovation would have taken place, while no protection mechanism provides static efficiency, IPRs encourage the system's dynamic efficiency. The economics of IPRs aims to balance the dynamic efficiency of continuous innovation and the static efficiency given by the absence of IPRs.

2.1.4 IPRs in an oligopolistic market

The trade-off between static and dynamic efficiency is sometimes solved using competition economics. However, given that dynamic efficiency is impossible to measure with a satisfactory degree of precision (Drexl, 2010), many economists of the neo-Schumpeterian school favor caution in applying competition law to IPRs' design, referring to Schumpeter's argument of *creative destruction*. According to Schumpeter, capitalism is an evolutionary process that is driven by *"new consumers' goods, the new methods of production or transportation, the new markets, the new forms of industrial organization that capitalist enterprise creates"* (Schumpeter, 1942).

Schumpeter affirms that it would be a mistake to evaluate oligopolistic structures without considering the evolutionary dimensions of the capitalist process. In particular, he identified innovation as a constant source of competition, explaining how even established oligopolies can be disrupted and countered by new forms of organization and technological revolutions. Additionally, Schumpeter examined the relations between IPRs and competition, affirming that competition policy should consider the positive effects of temporary monopolies to drive innovation and the creative destruction process activated by new technologies. He does not deny that, in some cases, competition policies are necessary to dismantle cartels that are effectively damaging the system but that drafting precise and effective policies is very difficult and should be done only on a case-by-case basis.

Economists advocating for neo-Schumpeterian approaches base their claims on the argument that monopoly power is necessary to enhance innovation and dynamic efficiency because monopolies' profit is needed for investments in future innovation (Drexl, 2010). So, even when IP confers monopoly power, the latter should be protected against state intervention because doing otherwise would reduce the innovation system's effectiveness. Competition policy should not collide with the allocation of IP rights because the long term effects of these interventions are challenging to quantify, and on a theoretical basis, it is never recommended to interfere with the distribution of the benefits of IPRs. *"When considering whether forcing the disclosure of companies' trade secrets or compelling them to license valuable IP, policymakers must therefore balance the gains from stimulating short-term competition with the losses from the reduced investment in innovation. They should keep in mind, however, that while the allocative efficiencies that arise in the short term as a result of intervention are relatively easy to measure, the long-term costs of such actions are uncertain and difficult to quantify."* (Geradin et al., 2006) A counterargument to the neo-Schumpeterian approach was already pointed out by Arrow (1962), which framed the applicability of Schumpeter's line of reasoning in terms of incentives to innovate. In an oligopolistic market, a firm is already making a profit, and innovation only offers a higher revenue. However, when compared to a second firm contemplating market entry, the incumbent's marginal gain[3] is lower than the one of the potential entrant. This argument leads to the conclusion that an oligopolistic system, in the absence of new entrants, is less conductive to innovation. However, as pointed out by Lévêque and Ménière (2004), *"this reasoning ignores competition between companies to conquer markets. The incentives are different if the monopoly knows that a competitor can enter its market by inventing and patenting a less expensive process or a similar new product."* When firms compete for conquering a market, a race towards innovation is very likely to happen. Competition in the market is then replaced by competition for the market (Geradin et al., 2006). If the incumbent fails, it will lose the profit provided by the monopoly and the R&D investments, while if the entrant misses the window, it will only suffer the loss of the development costs. In this case, the monopolist has a higher incentive to innovate, even if it will never introduce into the market the innovative product.

Summing up, competition law is a useful instrument to leave access to a market open by forcing oligopolies and new entrants to compete with each other and, coupled with compulsory licenses, promotes the diffusion of knowledge in the market. Simultaneously, even if IPRs create static inefficiencies due to their monopolistic nature, in many cases they maintain the incentives for the production of innovation high, promoting dynamic

[3] the monopoly profit after the invention minus the monopoly profit before it

efficiency, and discouraging the sub-optimal provision of innovation. However, such cases are not as frequent as one may think and that must be carefully evaluated on a case-by-case basis. As it was already mentioned, other incentives, economic or not, may already be in place. IPRs can alter these fragile dynamics, potentially hindering innovation, as it is claimed by Brüggemann et al. (2015) in the case of sequential innovation.

2.2 Three IPRs involving AI technologies: patents, copyright, and database

An in-depth analysis of the legal and economic features of all of them would go beyond the scope of this dissertation, thus I will focus on the main IPRs concerning AI technologies: patents, copyright, and the database sui-generis right.

2.2.1 Patents

Legal overview

Patents are a legal instrument that allows the patent owner to establish a temporary monopoly over particular uses of a patented invention. According to Guellec and de La Potterie (2007), a patent confers on its owner the exclusive right to prevent third parties to make, use, offer for sale, sell or import a patented product or the product resulting from a patented process. Patents only offer protection on a national level, requiring the patentees to individually seek protection in each of the countries where one desires to protect the intangible asset. To minimize legal fragmentation and reduce transaction costs, patent legislation has progressively harmonized at a global scale by several international treaties (Charnovitz, 1998).

In 1883 the Paris Convention for the Protection of Industrial Property was signed. It introduced two main provisions: national treatment and the right of priority. National treatment is a standard principle in international law that consists of providing to foreigners the same rights that it grants to its citizens, in our case, the ones concerning industrial property protection. The right of priority consists of conceding to the first applicant of a patent the exclusive right to extend it to other countries for twelve months. Before the Paris Convention, to ensure protection over his invention, the applicant had to file a patent in all states at the same time, creating several logistical issues (WIPO, 2004).

The second significant development in the international harmonization of patent systems was the implementation of the Patent Cooperation Treaty (PCT), signed in 1970 and entered into force in 1978, which created an international procedure for patent application. While the examination and grant of a patent remained a responsibility of the national offices, the PCT allowed inventors to file a single application stating a list of countries where he/she desired to obtain protection, and it provided additional 18 months of priority.

In 1994, the signature of the Trade-Related Intellectual Property Agreements (TRIPs) introduced the most favored nation principle, a standard duration of patents, a standard definition of patentable subject matter, compulsory licensing, and extended patent protection to test data in agricultural and pharmaceutical products

(Park, 2008). The TRIPs agreement represent the last development in international patent legislation and determine the minimum requirements for patenting. Article 27(1) of Annex 1C recites that: *"patents shall be available for any inventions, whether products or processes, in all fields of technology, provided that they are new, involve an inventive step and are capable of industrial application."* Additionally, all jurisdictions exclude from patenting artistic creations, scientific discoveries, mental acts, abstract ideas, animal and vegetal varieties, methods for medical treatment, and medical diagnosis. To this day, the patentability of business methods, software, and biotechnological products depends on the specific provision of national legislation. While in the US it is possible to patent business methods and software, in most European countries it is not (Wells, 2001).

Industrial application The *industrial applicability* criterion has been introduced to guarantee the practical utility of granting monopoly rights to the inventor. In this case, *industry* should be interpreted broadly, with the meaning that the invention must be applicable for practical purposes (WIPO, 2004). The economic rationale is to prevent the patenting of basic upstream applications that could slow down the development of commercially viable downstream applications, potentially reducing the net social benefit of society (Guellec and de La Potterie, 2007).

Novelty Novelty is an unquestionable requirement for patenting: an invention is new when the prior art has not anticipated it. *Prior art* is defined as *"all the knowledge that existed prior to the relevant filing or priority date of a patent application, whether it existed by way of written or oral disclosure"* (WIPO, 2004). The economic rationale at the basis of novelty is at the core of the utilitarian justification for IPRs. Since the introduction of an excludability mechanism involves a decrease in the static efficiency of the system to promote dynamic efficiency an innovation, there would be no reason to grant temporary monopoly rights for an invention that has already taken place (Guellec and de La Potterie, 2007).

Inventive step A third criterion shall be respected: inventive step[4]. According to Art. 56 of the European Patent Convention, *"an invention shall be considered as involving an inventive step if, having regard to the state of the art, it is not obvious to a person skilled in the art."* Given that an invention has industrial applicability and it is novel, why then require an inventive step? Again, the norm is based on economic considerations (Guellec and de La Potterie, 2007):

- **promote competition:** small and incremental innovation is part of the routine of many industries. Allowing a firm to patent a marginal improvement on a product or a process would prevent the competing businesses to innovate and would increase the probability of the development of permanent monopolies, hurting market competition;

- **reduce uncertainty:** requiring a low inventive step increases the risk of granting patents that are not novel, increasing the probability of raising society's costs without any benefit in the long-term. A high inventive step overcomes this issue and incentivizes radical innovation over marginal improvement;

[4]In the United States it is generally referred to as non-obviousness

- **ensure sufficient protection to inventors:** an inventive step which is too low reduces inventors' incentives because marginal improvements could immediately overtake any invention, dissolving the economic reward.

These three economic rationales have been included in models by O'Donoghue (1998) and Hunt (2004), who analyzed the impact of the size of the inventive step on innovation. According to their findings, a low inventive step reduces the lifetime of patents since patented marginal improvements would overtake them without risking of infringing them and incentivize investments in small, low-cost, marginal inventions while discouraging large and high-cost ones. On the other hand, a high inventive step increases the length of monopoly and the economic profits for a patented invention, while discouraging marginal improvements. Guellec and de La Potterie (2007) argue that determining the optimal length of the inventive step is not trivial and largely depends on the technological field of the innovation. They affirm that small inventions are valuable for customers and that since, in many cases, significant innovations are the product of marginal improvements, preventing patentability would incentivize secrecy and reduce economic welfare.

Economic rationale and optimal patent design

Adopting an economic approach to patents allows us to evaluate IPRs from an efficiency-based perspective, where their introduction is justified only if they encourage innovation and the diffusion of technology while minimizing the deadweight loss associated with the temporary introduction of monopoly rights (Langinier and Moschini, 2002). The net social benefit of an innovation is positive when the welfare it brings to society exceeds the cost required for its production. On the other hand, the net individual benefit of an innovation is positive when the innovator's returns are higher than the R&D cost he incurred during production. The underlying economic logic behind the patent system is to find a way to ensure that, when the net social benefit is ≥ 0 and the net individual benefit is ≤ 0, the innovation is produced anyway. Only when the trade-off between the net social benefit and the individual benefit is positive, a patent should be granted (Lévêque and Ménière, 2004). Providing a temporary monopoly right over the benefits of the invention was the historically chosen solution to this issue.

Many patent advocates affirm that, by ensuring high profits for the inventor, the disclosure of knowledge that would otherwise have remained secret and the creation of a market of tradeable assets would be encouraged, thus leading to the optimal allocation of productive resources (Guellec and de La Potterie, 2007). Unfortunately, this is rarely the case. The introduction of a patent involves two possible scenarios: the invention would have taken place even without the incentive provided by a temporary monopolistic right, or it would not. In the first case, from an economic perspective, the patent system should be structured in such a way that it is not possible to patent the innovation. However, determining that would require to evaluate factors such as the R&D costs of the invention, the risk of the research, the existence of alternative appropriability mechanisms for its returns, and the possibility of obtaining it through publicly-sponsored research. Generalizing, only if the invention is costly, risky, there is no other way of protecting it, and it is outside the interest of government-funded research, a patent should be introduced. Unfortunately, this identification process is costly, complex and largely depends on information that may be known only to the patent applicant, that has no interest in sharing it with patenting authorities, since it may result with having its patent application rejected.

Patents are an exclusionary tool that can potentially hinder innovation, thus betraying their original purpose: when a general-purpose invention is patented, it provides the holder with a strong market power that could prevent diffusion, follow-up research, and increasing the deadweight loss (O'Donoghue, 1998). This was the case of the patent granted to George Selden in 1895 for the invention of the automobile (Howells and Katznelson, 2016). Selden's application included the combination of an internal combustion engine in a 4-wheeled car. Selden's claim was very general and did not specify critical details regarding the engine. In 1899 he teamed-up with William Whitney and started collecting royalties from other automobile manufacturers up until Henry Ford and other four car makers challenged the patent in 1904. In 1911, Selden's patent was declared invalid by the appeal court. Merges (2009) suggests that Selden's rent-seeking behavior may have prevented access to the automobile market to small firms, discouraging follow-up research, and potentially impeding innovation. Another famous patent that arguably impeded innovation was the one granted to the Wright brothers, who made a modest improvement to flight technology and secured a patent in 1906. Later they engaged in a fierce legal battle to monopolize the US airplane industry, imposing prohibitive fees for economic gain (Boldrin and Levine, 2013) and effectively hindering follow-up innovation. When the US entered World War I, the US military did not have a competitive airplane industry, and the lawsuit was only solved through political channels by effectively absorbing the Wright patent in a patent pool composed by the other airplane producers (Shulman, 2003). On the other hand, the establishment of clear property rights facilitates the commercialization of the invention, ensuring the appropriability of returns (Coase, 1960). Many research institutes do not have the infrastructures needed to reach downstream users, and IPRs ease the transfer of these assets to industry actors through a license (Guellec and de La Potterie, 2007).

Moreover, in the absence of a patent system, disclosure rely on the willingness of the original inventor and the nature of the invention. While for product innovations there may be other systems of appropriability of the economic returns such as first-mover advantages or scale economies, in process innovations the most common alternative to patenting is secrecy. Therefore, in some cases, patents may favor the disclosure of process innovations and reduce the risk of the multiplication of R&D efforts (Archibugi and Filippetti, 2015).

"Applying this series of tests to individual patent applications or to entire categories (e.g. technical fields) would make the patent system of high quality from an economic point of view" (Guellec and de La Potterie, 2007). Unfortunately, applying these criteria to patent applications is not straightforward. Sometimes disclosure and exclusivity are in contrast with each other since greater exposure would facilitate competing firms to *"invent around"* the patent, thus reducing the returns of the innovative firm. In other cases, the information required to evaluate the impact of an innovation is known only after years of market introduction or it is not measurable at all. Additionally, the design of an optimal patent system requires greater coordination with the authorities in charge of innovation policies and competition, well beyond the reach of patent offices. Evaluating the grant of a patent on a case-by-case basis would increase the application cost, and it would reduce legal certainty, raising the

risks associated with unwanted infringement or of having the patent revoked (Guellec and de La Potterie, 2007). Since firms' investments in R&D are made depending on the probability of covering costs and appropriating the returns for profit, increased uncertainty in the evaluation process has negative effects on both competition and investment, amplifying business risk and raising the costs of innovation. Some institutions even explicitly exclude the use of economic considerations for patent evaluation, such as the European Patent Office. *"The EPO has not been vested with the task of taking into account the economic effects of the grant of patents in specific areas of technology and of restricting the field of patentable subject matter accordingly"* (European Patent Office, 2019).

Patent breadth Patent breadth (or scope) measures the use that the innovator can make of the patent compared to his competitors (Lévêque and Ménière, 2004). It is not determined by law but results from the combination of the number and genericity of novelty claims and the inventive step. Patent offices identify what an inventor has found or not based on the claims present in the patent application but the definitive scope is determined in court during a case of alleged infringement[5].

According to economic theory, the scope of a patent can be divided into *lagging breadth* and *leading breadth*. Lagging breadth affects horizontal competition, determining how far competitors should position their products in order not to infringe. Klemperer (1990) frames it as the maximal degree of substitutability between the patented invention and a competing one while Tandon (1982) defines it as the cost of inventing a substitutable product. In both scenarios, the competing products do not infringe on the patent, and the lagging breadth determines how much competition its owner will face. Taking Klemperer's perspective, a broader scope decreases the substitutability degree of competing products, giving the patent owner higher market power, while taking Tandon's one, it results in a higher entry barrier to the market and, therefore, weaker competition. Conversely, leading breadth affects vertical competition, which is the length of the inventive step needed in downstream research in order not to infringe the patent. It generally arises in the context of cumulative innovation (Guellec and de La Potterie, 2007), where an invention is needed for another one to happen [6]. In a two-stages development scenario, where two different research entities contribute to innovation and the revenue is generated in the downstream sector, leading breadth determines how the profit be shared between them. *"Granting too narrow a patent at one stage will reduce the negotiating power of the inventors at that stage as opposed to the inventors of the other stage, lowering their reward and, therefore, their incentive to invent in the first place, possibly blocking the innovative chain. Conversely, too broad a patent might lead inventors to require too high a share of the total reward, at the expense of other inventors, who might then choose to keep out, also breaking the chain"* (Guellec and de La Potterie, 2007).

Summing up, an increased lagging breadth reduces the possibility of competitors entering the market and

[5]In jurisprudence, there are two mainstream approaches: the doctrine of equivalents and the doctrine of literal interpretation. The doctrine of equivalents, extends protection to any product which *"does the same work in substantially the same way to accomplish substantially the same result"* (Wawrzyniak, 1990), while the doctrine of literal interpretation restricts the protection only to what is explicitly claimed in the patent document (Meurer and Nard, 2004; Lee, 2010). Both these approaches might be problematic, especially in rapidly emerging fields of technology, where in addition to real findings, the inventor may have intuitions regarding further applications of his invention.

[6]Such as innovations that improve the quality of a preexisting product or process or the discovery of new applications of an already existing technology

offering similar products, while a higher leading breadth makes it more difficult for competing innovators to patent marginal innovations based on the first patent (Scotchmer, 2004). Balancing out leading and lagging breadth is not trivial and economists suggest three main approaches.

The first approach advocates for the creation of deep patents that cover all innovations that follow an initial discovery. Kitch (1977) proposes that deep patents allow the inventor to organize research efficiently, discouraging patent races and excessive investments. Other firms can then identify new applications and propose them to the patent owner and request a license or offering research partnerships as they would represent a new flow of revenue. However, these arguments are highly controversial. High uncertainty and asymmetrical knowledge may increase the complexity of the negotiations (and therefore transaction costs), since no party may be willing to share valuable information regarding the specifics of their invention for fear of imitation. Broad upstream patents may affect the early disclosure of fundamental innovations, incentivizing secrecy and preventing other researchers from investigating other applications of the invention (Matutes et al., 1996). These rent-seeking behaviors may lead to patent races: firms may compete to be the first one to breakthrough and appropriate the totality of the revenues of the invention. Additionally, in many fields of research, the issue is not over-investment, rather under-investment. Even if the problem were over-investment, promoting broad upstream patents would not solve it but rather move it from the downstream to the upstream stage, reducing the profitability of downstream research and therefore risking an under-provision in that sector.

A second approach has been taken in the field of software. While computer programs *as such* are excluded from patenting[7], it is still possible to patent software applications which use a mathematical formula in a specific way for a particular purpose. This resulted in the incentivization of extremely narrow patents, which constitute prior art (Neuhäusler and Frietsch, 2019), incentivizing sequential innovation.

The third approach involves not applying intellectual property at all when innovations are cumulative, imposing the same non-patentability conditions as a requirement for licensing. Innovators directly compete with each other in the market, having a lower incentive to invest in R&D but having the possibility to draw from all existing innovations, reducing the costs derived from licenses, and eliminating the risks associated with possible infringement (Bessen and Maskin, 2009). This is another common approach used in the software industry, promoted by the open-source movement. In copyright legislation it is referred to as copyleft, while in patent legislation is referred to as patentleft (Dusollier, 2007).

Patent duration TRIPs agreements fixed the maximum duration of a patent at 20 years from filing, making it almost uniform all over the world. Limiting the temporal length of protection is the easiest way for legislators to control the monopoly rights given to innovators. A longer duration would grant additional profits to innovators, and therefore it would increase their economic incentives. On the other hand, from an economic perspective, protection should be eliminated as soon as the innovator as received a reward equals to the profit it would have obtained if he invested the same amount of resources elsewhere, such as in bank deposit. In this framework, the discount rate is the first limit on the efficiency of long patents: the more distant in time the expected profit

[7] They are considered mathematical processes and fall in the public domain

from the innovation, the less able it is to compete with the cumulative interest that a bank account would generate (Scotchmer, 2004). Another reason for limiting the duration of patent protection are transaction costs (Langinier and Moschini, 2002). Among time, subsequent inventions dilute the contributions to the technology provided by each inventor, and establishing the participation of an inventor to the new knowledge becomes more and more expensive as technology evolves (Lévêque and Ménière, 2004).

Having determined that limited duration is essential for patents, how long should protection last? In other words, how is it possible to balance out the dynamic efficiency of a patent? Nordhaus (1969a) developed a model to determine optimal duration that minimizes the discounted value of the deadweight loss generated by the patent over its entire life under the constraint that the discounted profit of the innovator exceeds his R&D expenses. He affirms that duration and the scope of patents should depend on the particular characteristics of the reference market since different products and processes may have different optimal extents of protection. Thus the optimal solution would be to grant different protection lengths to innovations in different sectors, based on the cost of the invention and the shape of the deadweight loss as a function of price and time. Unfortunately, the current patent system establishes a standard maximum length to all patents, even if the R&D expenses greatly vary between different sectors. While there is a renewal system based on fees (De Rassenfosse and van Pottelsberghe de la Potterie, 2013), in which, periodically, firms can choose to extend the duration of protection, the deadweight loss imposed by a patent may still be higher than the benefit to society. The profits derived from a patent may be higher than the renewal fees even after the complete recovery of the R&D costs, thus blocking the diffusion of socially valuable inventions and encourage rent-seeking behavior, undermining the efficiency of the system.

To determine the optimal patent design, different combinations of length and scope can be used. The same economic reward could be obtained by either increasing the scope or the length of protection. Different approaches have been taken to strike a balance, notably by Tandon (1982), Gilbert and Shapiro (1990), and Gallini (1992). According to Tandon (1982) and Gilbert and Shapiro (1990), the optimal combination of duration and scope largely depends on the elasticity of demand, which influences the impact of price on the deadweight loss. Patent protection is framed as the power it confers on the market. Following this understanding, a broad patent reinforces the monopoly of the innovator by excluding from the market products that are substitutes for it but that are substantially different. So, if the product elasticity is high, deadweight loss increases faster with breadth than with length, leading to the conclusion that narrow and long patents are preferable. Conversely, when the product is highly inelastic, deadweight loss increases faster with duration than with breadth, making short and broad patents more desirable. A different approach was taken by Gallini (1992), which instead took into consideration the effect of patent breadth on competition. She defined breadth as the R&D costs of imitation of an innovation without infringing its patent. Patents are framed as the conditions under which the innovator shares the market. *"The longer the patent, the more incentive imitators have to invest in the creation of alternative technologies. By contrast, a broad patent makes it costly for imitators to enter the market. In other words, a long patent attracts imitators by giving them the time to recover the cost of their imitation, whereas a broad patent dissuades imitators by increasing the cost of imitation."* (Gallini, 1992) However, the

R&D expenditure undertaken by imitators is wasted because the patent technology has already been developed. Because of these considerations, she concluded that short and broad patents are preferable because it is better to have a strong monopoly for a short period than an oligopoly for a longer period coupled with worthless imitation costs.

2.2.2 Copyright

Copyright is a legal instrument used to protect *"any literary, dramatic, musical or artistic original work, provided that is recorded, either in writing or otherwise"* against imitation (Spence, 2007). Contrary to patents, it only protects against the exact reproduction of the work and not independent creation. Copyright gives authors an exclusive right over the reproduction, performance, adaptation, and translation of their work. The criteria used to determine whether an intangible asset can be subject to copyright protection are very inclusive. Work is defined as any quantity of material assembled with some conscious ordering. There is no judgment of quality or contents in the definition of literary: originality does not require any inventive step, only an intellectual effort of some kind (Spence, 2007). Additionally, copyright legislation also covers a range of related rights, which are rights of creative works not connected with the works' authors. They include the rights of performers, phonogram producers, broadcasting organizations, and, within the European Union, the rights of film producers, database creators, semiconductors, and industrial design. Provided these general requirements, it is clear how the subject matter of copyright protection is extensive. It encompasses many industries, such as visual arts, publishing, performing arts, and, since the digital revolution, digital products such as software code and databases.

Civil law and common law regimes adopts two different approaches. In common-law copyright is understood solely as a set of economic rights that aim to guarantee the economic appropriability of the production of knowledge. They include the exclusive right for the reproduction, distribution, renting, lending, public performance, and making adaptations of the work, such as translating or dramatizing a literary work and transcribing or producing an altered version of a musical work (Spence, 2007). On the contrary, in civil law regimes, other rights (commonly referred to as moral rights) such as paternity, false attribution, and integrity, are associated with the economic rights. The distinction between common law and civil law approaches to copyright is reflected even in terminology. While in the US and former Commonwealth countries, the legal term to define these IPRs is copyright, in civil law countries, more emphasis is put on the author figure, and therefore they are referred to as author's rights (generally using the French term *droits d'auteur*).

Differently from patents, that have a registration system that requires the filing of an application to obtain protection, copyright does not have a registration system. Protection is granted since the act of creation of the work, and it is automatically awarded to the author (Spence, 2007). However, sometimes the identification of the author is not so trivial as one may expect. According to VerSteeg (1995), an author is an individual who has actively created something copyrightable, as the writer, the photographer, or the composer.

From a historical perspective, the author is a legal figure born in XVII century England, in the context of the dispute Donaldson v. Becket. Donaldson was a Scottish bookseller whose business model consisted of reprinting classics of English literature. When he published James Thomson's book *"The Seasons,"* Becket and a group of London booksellers sued him for infringement, since they allegedly claimed they detained the copyright (Rose, 1988). At the time, copyright in the United Kingdom was protected by the Statute of Anne (1709), which granted limited protection of fourteen years to books with a possible second term if the author was still living. The Statute was based on the Stationers' Company regulations, a publishers' association that regulated the printing of books in London. It introduced a temporal limit of protection and the possibility for authors to become copyright holders, effectively giving birth to the author as a legal figure. In England, copyright was a right created for publishers, responsible for the diffusion of literary works. While at the dawn of copyright legislation it was relatively easy to determine the author of a work, with the progressive expansion of copyright subject matter and complexity, it has become increasingly challenging to assign attribution. In the past, business models such as work-for-hire and technological innovations such as film-making questioned the traditional association of property rights with authorship, leading to the creation of models of attribution tailored to the specific issues. Today, AI-created works are putting these models again to the test.

Length

Copyright protection starts from the moment the work is created, and it proceeds until some time period after the author's death. The ratio of this provision is to allow the author's successors to benefit economically from the author's work. In some legislation, moral rights continue in perpetuity. In the Berne Convention, the minimum duration of copyright protection is fixed at 50 years after the death of the author. However, more and more countries are extending the length of protection to 70 years, to harmonize with US and EU IP regimes. In some cases, specific categories of works (such as anonymous, cinematographic, and posthumous works) are awarded a different length of protection to minimize the deadweight loss to society. When copyright protection expires, works enter into the public domain and are freely accessible by everyone.

International copyright legislation is highly fragmented, and it is currently formed by an interlocking system of treaties formed by the Berne Convention for the Protection of Literary and Artistic Works of 1886, the Universal Copyright Conventions of 1952 and 1971 and the WIPO Copyright treaty of 1996. Additionally, neighboring rights are regulated by ad-hoc treaties such as the Rome, Geneva, and Brussels Conventions and the WIPO Performances and Phonograms Treaty. In 1994, the adoption of the TRIPs agreement contributed to the harmonization of international copyright law, but the global legal framework is still disjointed. Some countries have adopted all the latest multilateral treaties, while others abode only the minimum standard required by previous agreements (Goldstein, 2001).

Compared to patents, which in almost all legislation have a standard limit of protection of 20 years from filing, copyright has an unusually long term of protection. The economic rationale is that the success of a work is highly uncertain, and it can come many years after the first edition. So, a long span of protection represents an additional guarantee for the author to obtain the profit from his work, even if it is delayed in time (Landes and Posner, 2003; Diderot, 1767). In this way, the creator's lower profit in each period is compensated by a higher number of periods, to enable him to recover his cost and provide him with sufficient incentive (Landes and Posner, 2003). On the other hand, the unusual length of copyright protection may have also arisen from the intense lobbying of interest groups in the show business industry. Initially set at 14 years, copyright in the US has been gradually extended to the current 70 years after the author's death. While this unusual length may be partly justified in the context of literary works, with the expansion of the scope of copyright to works that are technology themselves, or essential inputs for the production process, such as software or databases, 70 years after the author's death greatly increase the deadweight loss, suggesting that a revision may be needed for specific products.

Scope

Copyright protects the exact reproduction of the author's works and not the act of using the ideas expressed, considered the building blocks of knowledge and fundamental for transmission. Word strings and music are protected, as other ordered strings of symbols such as software, but not algorithms. In the US legislation, the dichotomy between ideas and expression is solved by the doctrine of *"comprehensive nonliteral similarity,"* (Mohler, 1999) where the line between infringement and autonomous creation lies on the evaluation of mechanical changes are made, such as changing the names of the characters in a short story or renaming the variables in software.

Copyright protection is also subjected to several limitations. First of all, some categories of work might not be protected at all. International legislation requires that a work is fixed in a tangible form, and therefore non-registered works fall out of copyright protection. Additionally, some legislation explicitly exclude from protection works such as laws, court sentences, and administrative decisions (Goldstein, 2001). Second, some forms of exploitation of a work that theoretically would require a license may be carried out without explicit permission. Article 9, par 2 of the Berne Convention affirms that reproduction rights may be granted, *"in certain special cases, provided that such reproduction does not conflict with a normal exploitation of the work and does not unreasonably prejudice the legitimate interests of the author"*. Third, it is possible to grant compulsory licenses regarding certain individual cases, such as the mechanical reproduction of musical work and broadcasting. Similarly to patents, the author receives an economic compensation, but the amount is generally decided by a court or by the consumer (WIPO, 2004).

The reason for these exceptions lies in economics. When a consumer gives a low but positive value to a work, and they are not willing to meet the cost of a transaction with the copyright holder, the diffusion of the work may be reduced to a sub-optimal level, diminishing the overall economic welfare (Landes and Posner, 1989). To solve this problem, the US developed the *fair use* doctrine which, by analyzing the nature and the purpose of the usage of the work, determines whether a copyrighted work can be used without the permission of its author (Depoorter and Parisi, 2002). The doctrine has no equivalent in Europe, but it is considered admissible in some national law regimes. Fair use exceptions generally include citations, news reporting, and use of the works for teaching purposes. Additionally, depending on the legislation, copyright protection may be limited by a range of defenses such as fair use for non-commercial purposes (private study and research), public interest, parody, review, and criticism.

2.2.3 Databases

The recent technological advancements and the development of a data-driven economy prompted many business and industries to push their government for the introduction of a new right for protecting the investments made in databases. Database production is a long and expensive process, but ordered and easily searchable data-sets are essential prerequisites for many information services, and the basis for ML-powered AI.

Potentially, databases could be covered by copyright, under the definition of literary work and in compliance with the Berne Convention. Additionally, article 5 of the WIPO Copyright Treaty (1996), states that *"compilations of data or other material [...] which by reason of the selection or arrangement of their contents constitute intellectual creations, are protected as such. [...] This protection does not extend to the data or the material itself and without prejudice to any copyright subsisting in the data or material contained in the compilation"*. However, copyright protection requires a creative effort and originality, and it is conceptually difficult to stretch these elements to databases. This opinion was first expressed with the sentence of the US Supreme Court in the case Feist v. Rural Telephone Service Company in 1991. Rural was a public utility providing telephone services to various communities in Kansas and published databases of telephone address. Feist extracted information from the database in question, and Rural accused them of being in infringement. The US supreme court sentenced that the database was not eligible for copyright protection because it lacked the requirement of originality, and its creation did not involve any creative effort, rather the opposite: the intellectual capital of databases depends on completeness, ease of access, and the standardized categorization of information. Databases are hardly created from scratch. On the contrary, they draw information from pre-existing databases and other sources of data. On the other hand, applying copyright legislation to databases has raised more than one eyebrow: given the particular nature of the subject matter, applying the entire length of copyright protection to databases could potentially lead to a disproportionately higher welfare loss.

Still, a part of the database industry lamented the absence of a system of protection for uncopyrightable databases as a potential cause of market failure, where free-rider competitors who contributed nothing to the data collection could appropriate the returns of the producer's investments. Therefore, starting from the 1960s, some European states created a series of legislative instruments to protect databases that fell outside the scope of copyright protection. Since legal discrepancies between the EU Member States could potentially harm the free-flowing of goods within the Digital Single Market (DSM), in in 1992 the European Commission drafted a proposal for a Directive on the Legal Protection of Databases that introduced a sui-generis right for databases, and was later amended and entered into force in 1996. The Directive was not self-executable, and it needed the implementation law by the Parliaments of the Member States. While within the EU, the harmonization was concluded in 1998, at the international level the reciprocal character of the Directive has created pressures for the acceptance of the EU model of database protection.

The European Directive on databases

Article 1(2) defines a database as *"a collection of works, data or other independent materials arranged in a systematic or methodical way and capable of being accessed by electronic or other means."* The directive excludes *"computer programs used in the manufacture or operation of databases which can be accessed by electronic means."*

In Art.3(1), the Directive distinguishes two different scenarios for a database. When the database already enjoys copyright protection, since the creative effort required for its creation makes it eligible for copyright protection and when it does not. When the database enjoys copyright protection, the Directive provides the author the standard copyright legislation, depending on how Member states structured the national law of application. Still, this provision cannot unreasonably prejudice the rightholder's legitimate interests or conflicts with the normal exploitation of the database. When a database does not enjoy copyright protection, it falls within the scope of the sui-generis right, that is the novel element of the European Directive, conditioned to the ability of the *compiler* of demonstrating that there has been a *"qualitatively and/or quantitatively substantial investment in either the obtaining, verification or presentation of the content to prevent extraction and/or re-utilization of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database."*

Extraction and re-utilization are precisely defined in the Directive. The former is defined as *"the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form,"* while the latter is *"any form of making available to the public all or a substantial part of the contents of a database by the distribution of copies, by renting, or by other forms of transmission, including on-line."* This rule grants protection to the producer of a database, irrespective of the database's eligibility for copyright or other protections. However, actions concerning the non-substantial part of a database are non-infringing, and the sui-generis right is limited by first-sale exhaustion. Public lending is expressly excluded from the scope of protection. The national law of application determines additional exceptions.

The protection accorded through the sui-generis right is granted for fifteen years, starting from the completion of the database. It is possible to renew the term of protection for another fifteen years if the database has been subjected to a substantial change that required a significant new investment. However, this creates issues concerning the incentive provided to database producers, since renewal creates an incentive to update the database even when there is no real need with the objective of extending protection (Koboldt, 1996). Moreover, in the final draft of the Directive, the compulsory licensing for sole source databases originally proposed by the European Commission was deleted. Such a license would have made it possible in those instances where the stored information could not at all or not without great practical or financial difficulties be created or gathered independently (Grosheide, 2002). From a purely economic perspective, however, the absence of a compulsory licensing scheme equals to granting an indefinitely long monopoly. Data can be kept reserved but used within the firm, providing a comparative advantage. Databases are essential for Artificial Intelligence systems, since machine learning, the game-changer in the field of AI, needs organized and ready-to-use data-sets are to build effective predictive models. As a consequence, inefficiencies in economic welfare are created, increasing the deadweight loss caused by IPRs in the data economy.

2.3 AI and Intellectual Property Rights

2.3.1 Limits of IPRs legislation

Unfortunately, the review of the main economic arguments on the best structure for IP systems remains on a theoretical basis. Since its first developments in the XIV century, the legal framework that surrounds patents, and, in general, IPRs, has demonstrated a lack of flexibility that is in striking contrast with what it aims to protect, that is, technological innovation and creation. While economists such as Nordhaus (1969a) call for an increase in differentiation in the nature and degree of IP protection, based on industry and market needs, the fundamental economics of existing laws remain fixed on old assumptions and paradigms. This is because legal institutions evolve incrementally and are focused on preserving the integrity and consistency of rules, even when there is a dire need of radical transformations to remain faithful to the rationale and motivations that gave them birth in the first place (David, 1993). When patents first appeared in the XIV century, they were designed to attract foreign inventions to boost development. They incentivized innovation in the sense that they promoted the dissemination of precious information in a national economy, encouraging its growth in the long run, but they were conceived as a preindustrial system of protection that served preindustrial or industrial needs. The development of a knowledge-based economy sharpened the possibility of information to act both as a capital and a consumer good, and has made evident that the patent system needs a revision, since the incentives to innovate have changed (as the success of models such as patentleft and copyleft demonstrated).

The fixed character of IP law has increased the level of uncertainty regarding the adaptation of IP law to new technologies (David, 1993). While patents were designed to promote innovation, new technologies and economic structures have changed their role in the economy. Invention is often a cumulative process and the enforcement of patent rights can interfere with further discovery. Patent races and inventing-around strategies may deflect resources and incentivize rent-seeking behavior, while at the same time discouraging complementary inventions because previous patentees may be able to extract revenues in downstream innovation (David, 1993). The current legal framework is far from being efficient and, in the meantime, new technologies keep on posing new challenges and requiring emergency adjustments to the new circumstances. In particular, AI technologies are posing critical challenges to the IP legal framework, while at the same time IPRs influence the structure and rate of innovation, introducing distortions in the market through the introduction of incentives or disincentives to innovators.

Recently legal scholars have started a lively debate on how to face the challenges presented by AI to the current legal system. However, as it was previously mentioned, the legal perspective is intrinsically limited. Recalling the utilitarian justification, IPRs should be introduced only when they contribute to the promotion of innovation. Thus, I argue that the introduction or modification of IPRs regimes needs to prioritize the economic effects of introducing new norms rather than continuity with previous legislation. As mentioned in section 2.1.1, policymakers should maximize the net social welfare when shaping the rule of property rights. This requires to strike a balance between incentives that stimulate innovation and the tendency of IPRs to create monopoly rights and dysfunctional effects.

In particular, the general statements concerning utilitarian perspective on IPRs need to be reassessed with the concrete case examined, in our case the AI market. Unfortunately, since creative and innovation markets related to AI function under unequal conditions, this is not an easy task, and, even if IPRs are required to stimulate investments in R&D, current regimes need to be re-calibrated to take into account both market and non-market considerations, favoring an increase in the differentiation of the scope and length of IPRs. For example, a key factor is the elasticity of product demand, as it was suggested in section 2.2.1 in relation to optimal patent design (Nordhaus, 1969a). IPRs are not justified a priori: if in the market there are already sufficient incentives to promote R&D efforts, the introduction of monopolies on technology would only reduce society's welfare, either through monopolies or by adding the administrative costs of the IP system. Hilty et al. (2020) explored various economic paradigms to justify the application of new IPRs to AI technologies: the general incentive theory, the investment protection theory, and the prospect theory.

2.3.2 General incentive theory

General incentive theory is the original argument of IPRs, that we examined in section 2.1.3. It suggests intellectual outputs would not be produced without the incentive provided by property rights. However, it would be short-sighted to affirm that economic rewards are exclusively tied to mechanisms of appropriation such as IPRs. In particular, as it was mentioned in section 2.1.4, maintaining a stable market position and/or accessing successful market opportunities may be determinant for a firm to invest in R&D. When we apply this theory to the AI market, there is no evidence of lack of motivation to invest in innovation, rather the opposite: AI-related innovation is thriving. While an under-provision in innovation in the AI field is still theoretically possible, this may likely be caused by other bottlenecks in the system, such as the scarcity of AI experts (McGowan and Corrado, 2019) and data fragmentation (Martens, 2018), issues that cannot be solved by using IPRs.

When we transpose this approach to the by-products of AI, legal scholars often claim that without gaining rights to the outputs of AI, developers have no sufficient incentives to create such AI systems in the first place (Abbott, 2017). However, this does not take into account that there are different ways through which developers may limit the use of these systems, such as product-as-a-service (PaaS), in which the use of the AI system residing in cloud platforms can be controlled by limiting the number of API calls, the duration of the task and the amount of the processing power used. Another commonly used business model is typical of the platform economy, where developers may provide the opportunity to access to higher-quality, personalized AI systems in exchange for a periodic fee[8]. AI system developers can potentially never fully disclose the AI model to the end-user, thus making their work profitable without the need to claim ownership on AI by-products.

[8]Through the use of a technique called transfer learning

We can also observe that the AI market has also adopted a self-regulation model, where actors voluntary share their intangible assets (Gonzalez et al., 2020) under various kind of licenses. This approach is typical of software industry, where many important demanders actively contributed and distributed open source projects as a source of complementary innovation (Wang et al., 2020), while individual programmers often collaborated for prestige or good-will, leading to a process where innovation *feeds* itself. To a lesser extent, we are assisting to a similar process in the context of the data market, with pleas of the liberation of data gathered by public administrations in standard, freely accessible formats (under the open data paradigm) that may lead to increases in innovation rate and economic competitiveness for private companies (Alderete, 2020).

2.3.3 Investment protection theory

Investment protection theory focuses on the need to protect the investments made by the innovators by minimizing free-rider behavior (Hilty et al., 2020). It was modeled by Nordhaus (1969b), who applied this framework to technologies whose innovation processes are long and require high investments in markets with slow innovation cycles and high probability of imitation and misappropriation. This paradigm is extremely targeted to market scenarios where incentives to innovate are low, suggesting the introduction of IPRs as a way to increase them.

Although AI technologies require a certain degree of investments (such as high-quality data and computing power), it varies greatly depending on the specific application sector and scale of the industry. However, in most cases, AI innovation works through incremental progress which generally does not require high investment (Hilty et al., 2020). Even if imitation is possible, other systems of protection may already be in place, such as the database directive. Moreover, Teece (1986) affirmed that, in certain cases, investments may already be protected even when IPRs are not granted at all. Specifically, he suggested that, in some cases, such as when new products are difficult to copy, there is no need for state-provisioned protection. He made the case of cospecialized assets, whose value is tied by a bilateral relation between two goods: possessing one without the other is of no use. He affirmed that in the context of innovations whose value depend on cospecialized assets that can easily be kept secret, IPRs are of no use, since a mechanism of appropriability is already in place, thus making the introduction of IPRs detrimental to economic welfare. A typical example of cospecialized assets in AI technologies is characterized by algorithms and weights. While algorithms are often published in scientific papers, weights[9] are generally kept secret (Hilty et al., 2020). An AI system without correlating weights would be valueless for potential free riders. Moreover, imitation through reverse engineering is possible only to a certain extent because of the black box effect of ML and, even when it is possible, it is often more costly than building an original AI system. This further reduces the incentive for imitation in favor of autonomous creation. As a consequence, it seems that the investments undertaken for developing AI systems are already sufficiently protected and that the introduction of IPRs on the basis of investment protection theory is not motivated by any economic justification.

Someone may suggest that this is only applicable to AI systems and not to the by-products of AI. However, the extent to which it may be reasonable to introduce additional IPRs to AI-aided products largely depends on

[9] Trainable parameters that are optimized during the learning process.

the rate of innovation of the specific industry, which determines the rate of substitution of new products. The utility of AI outputs is strictly correlated to the time it takes for them to become obsolete. In other words, when AI-based innovation and creation surpass a certain threshold, the utility of each AI output reduces, as effect of incremental innovation, a situation where investment could not be recouped even with the introduction of IPRs[10]. A higher rate of innovation reduces the temporal difference between dynamic efficiency and static efficiency, making IPRs detrimental to economic welfare. The rate of innovation depends on the specific economic sector of the AI output, that cannot be determined before and can only be evaluated on a case-by-case basis (Hilty et al., 2020). So, while IP protection for AI outputs may potentially be justified under the investment protection theory, affirming that a completely new IP regime is needed for AI by-products is merely speculative and it may be counterproductive, since it would increase the legal costs associated with innovation. Considering that current legislation already provides the possibility to protect AI-aided outputs by mentioning a human as inventor or author[11], there is no need to create new IPRs.

2.3.4 Prospect Theory

In section 2.2.1, I already presented the basis of prospect theory when referring to Kitch's proposition for determining optimal patent breath (Kitch, 1977). He suggested granting to the inventor short and broad patents to improve R&D coordination and reduce over-investment in innovation. In the context of AI technologies, over-investment may seem relevant, but it would also interfere with the ability of other innovation to conduct parallel R&D efforts, thus unnecessarily blocking the way of other innovators. This may subsequently lead to *"races to create or invent, which may lead to wasteful duplication of research effort. Furthermore, instead of enabling the original inventor to coordinate efficiently the exploitation of the technology, a quasi-monopoly may lead to satisfying behaviour and thus to an inefficiently narrow focus on improvements related to the primary AI creator's or inventor's principal line of business"* (Hilty et al., 2020). The introduction of IPRs may reduce the incentive to proceed with cumulative innovations and favor rent-seeking behavior. As we will see in chapter 3, this is particularly relevant, since AI technologies are a General Purpose Technology, and enhancing the scope of protection on the basis of Kitch's prospect theory is very likely to slow down or exacerbate virtuous cycle of complementary innovation, reducing economic welfare.

2.4 Conclusions

This chapter presented Intellectual Property Rights, a legal institution that has the declared goal to incentivize innovative behavior. First, the philosophical foundations for the creation of property rights on intellectual creations were presented, distinguishing between the natural rights arguments, the desert argument, and the utilitarian argument, and accepting the latter as the most reasonable one, since it is agnostic to the introduction of IPRs. The utilitarian argument accepts IPRs only to the extent to which they present benefit to society and

[10] Which would only increase the deadweight loss and reduce economic welfare.
[11] Here it is important to underline that AI-produced inventions do not exist, as it will be discussed in chapter 3.

it is based on economic analysis. Second, I analyzed three IPRs strictly involved with AI technologies: patents, copyright and database sui-generis right. The section was mainly focused on the economics of patents, the first IPR introduced in the economy, on which other IPRs, such as copyright, were modeled. Third, I applied the economic analysis of IPRs to AI technologies. In contrast to the legal literature on the relations between IPRs and AI, I observed that new IPRs would only be detrimental to economic welfare and innovation. I determined that the dynamic character of the AI market and the presence of alternative systems of appropriation already provide sufficient incentives to embark in innovative behavior. The introduction of additional IPRs specific of AI technologies would only represent a burden to the innovative process, potentially hindering innovation.

Chapter 3

AI and Innovation

Technological change and innovation have long been recognized as crucial drivers of economic growth. Today this necessity became ever more urgent to face the challenges of the post-pandemic recovery and climate change. More certain is the extent to which a specific technology will have the desired effect, but there is a consensus on the fact that the more sectors will be invested by change, the deeper the changes in the economy will be. Since prediction technologies are a fundamental element of every activity, AI technologies are configuring as a source of profound transformation both in industry and research setting, possibly changing how innovation itself take place.

Throughout human history, technologies with similar characteristics, such as electricity, steam engine, information and communication technologies (ICT), have been classified as *General Purpose Technologies* (GPT). According to Lipsey et al. (2005), *"a GPT is a single generic technology, recognizable as such over its whole lifetime, that initially has much scope for improvement and eventually comes to be widely used, to have many uses, and to have many spillover effects"*. Another fundamental category of technology studies is the *Invention of a Method of Inventing*, or IMI (Griliches, 1957). Generally they are research technologies that cause paradigm shifts within the scientific community, increasing research productivity and easing bottlenecks in development and innovation. Understanding the different ways in which technologies' impact society is vital to guide policymakers to choose the most appropriate policies to foster economic growth and social balance.

The first part of this chapter is dedicated to building a theoretical framework for the technological categorization of General Purpose Technologies, Methods of Invention, and analyzing how they spread in the economy. In the second part of the chapter, I will analyze whether Artificial Intelligence enters in one or more of these categories from a qualitative perspective. Artificial Intelligence is impacting a wide variety of application sectors, affecting productivity, employment, and competition, suggesting that AI should be considered a GPT. Additionally, some scholars have also highlighted that AI can revolutionize the innovation process itself, increasing the number of tools available to researchers and knowledge producers, suggesting it may also be classified as an IMI. This has prompted the scientific community to define a new category that combines the characteristics of GPTs and IMI: *General Method of Inventing*, or GMI.

Many scholars also affirm that AI can represent a solution to the issue of stagnant productivity. However, the use of AI in the market has both direct and indirect consequences on the market structure, increasing the concentration of capital in a small number of very large firms. Goldfarb and Trefler (2019) affirmed that given the specific characters of the digital economy, AI-powered businesses are changing the way competition takes place, with firms competing for the conquest of new markets. The third part of the chapter focuses on analyzing these phenomenon and presenting some of the externalities caused by the use of AI as a commercial technology.

3.1 A theoretical framework for classifying technologies

3.1.1 General Purpose Technologies

The notion of *General Purpose Technologies* was first introduced in the economic literature by Bresnahan and Trajtenberg (1995) to identify technologies that have a broad impact on society by promoting both qualitative and quantitative changes in institutional structures, industrial dynamics, allocation of resources, and skill requirements, influencing the macroeconomic dynamics of the international economy. The assumption behind the idea of GPTs is that *"major technological changes are the main determinants of cyclical and non-linear patterns in the evolution of the economy"* (Cantner and Vannuccini, 2012). However, the idea that technology represents an exogenous factor that uniquely influences the direction and rate of change was formulated well before the concept of GPT itself: Schumpeter (1934) noticed that drastic innovations were often followed by long periods of economic growth while Mokyr (1990) uses the term *macroinventions* to refer to those innovations that were essential complements for the development of a much-increased number of *microinventions*.

In the economic literature, GPTs are usually identified by the presence of three features:

1. **wide scope for improvement and elaboration**: to cause important disruption in the economy, a technology needs to be able to go through a process of rapid technological advance after its first introduction. It is fundamental for the technology to provide the possibility to build knowledge around it cumulatively;

2. **pervasiveness**: for a technology to be *General Purpose*, it must be possible to be used in a wide variety of products and processes, thus making it able to spread across different economic sectors and increase its impact on different industrial compartments;

3. **strong complementarities with existing and new technologies**: it must be possible to apply the candidate GPT in combination with existing or future technologies, further increasing the opportunities to develop new technologies. In other words, the productivity of R&D in any downstream sector may increase as a consequence of innovation in the GPT.

Features (1) and (3) are often incorporated under the concept of *innovation complementarities*, a virtuous cycle of complementary innovation between the GPT sector and the application sectors, as presented by Bresnahan (2010). During a first phase, the GPT technology sparks innovation in the application sectors while at a later stage, the new technologies developed in the application sector generate further innovation in the original GPTs, increasing its technological innovation speed and triggering a positive loop of technological advancement. The overall effect that innovation complementarities have on economic growth largely depends on whether application sectors are capable of implementing the GPT. Suppose the actors in the application sectors manage to do so. In that case, Bresnahan (2010) suggests that the *"social increasing returns to scale (SIRS) arise across an entire cluster of technical change in the GPT and technical change in the AS [application sectors]"* and that GPTs *"may overcome diminished returns because innovation is inherently an increasing returns activity"*. However, this virtuous cycle of economic growth is not exempt from obstacles and externalities, which will be later discussed.

Some words on pervasivesness

Pervasiveness is a GPT quality that describes its *"generality of purpose"*. From a theoretical perspective, the Schumpeterian model of innovation (Schumpeter, 1934), the theory of recombinant growth (Weitzman, 1998), combinatorial technology models (Arthur and Polak, 2006), and recent studies on technological diffusion (Korzinov and Savin, 2016) that consider knowledge as a combination of "pieces" of information that can be connected in one or the other way are helpful to codify such a complex phenomena.

According to Weitzman (1998), new ideas arise from existing ideas in a combinatorial process. He developed the theory of *"recombinant innovation"*, where old ideas are reconfigured and recombined to produce new ideas. In this framework, the set formed by new ideas is modeled as a subset of the combinatorial set of already existing ideas. The researcher/innovator's task is to explore the set of *old* ideas to detect the ones that are more likely to lead to the discovery of *new* ideas.

Arthur and Polak (2006) operationalize this concept with the help of network science by simulating the process of technological evolution using simple logic circuits. Their work confirms the hypothesis that technological systems can be modeled as a dynamic network of nodes where new nodes are formed as the consequence of the combination between two or more existing nodes. The simulation they performed showed that the formation of new nodes is not only driven by the availability of previous technologies but also by the collection of human needs (provided by the researchers) and the needs of the technologies themselves (the intermediate steps needed in order to reach the research goal). Additionally, some combinations of technologies can become obsolete and practically disappear from the network, paralleling what Schumpeter defined as *"gales of destruction."*

Korzinov and Savin (2016) built on this literature to create a network-based model of technological evolution. They modeled the relation between technologies and final products using a bipartite network of two-layers: final products and technologies (also referred to as intermediates). In the beginning, technologies are not connected. However, to find practical application (and therefore create a final product), they need to become interconnected and form a fully connected component (also referred to as clique).

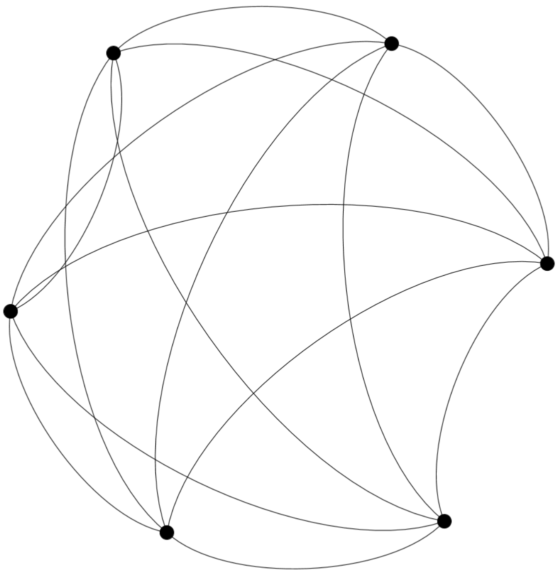


Figure 3.1: A technological clique composed by six different technologies (nodes) connected with each other

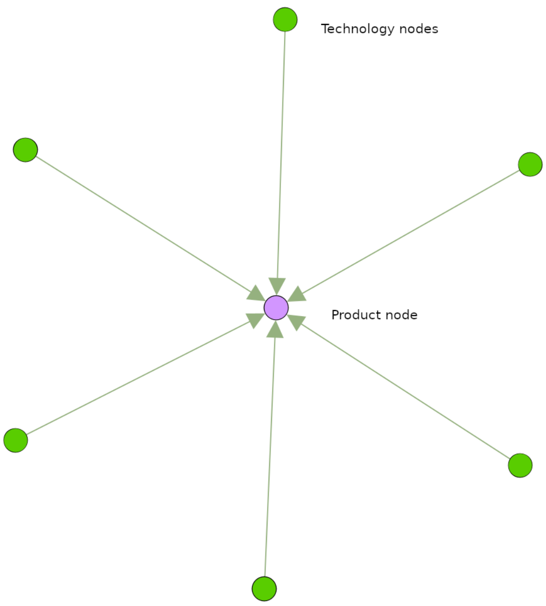


Figure 3.2: The bipartite representation of the technological clique. The node at the center of the graph represents the product created by the combination of different technologies

In their model, each link between technologies represents the process of knowledge discovery, where *"existing technologies are combined with one another in new ways to produce value for a consumer."* Each product that satisfies a specific need can be produced using different technologies (such in the case of substitute goods), and the same combination of technologies may serve to produce different final goods (the complementarity arising from combining different technologies can have more than one application).

By combining all possible technology combinations, one obtains a *"potential technology network"*. In their model, the *potential technology network* is not equivalent to a fully connected graph because the structural properties of the technologies limit the network development: not all technologies can be combined, and therefore not all links are possible. In this context, a technology's pervasiveness may be measured based on the number of links it has with different products. The extent to which a technology may be considered a GPT can then be measured on the degree of its pervasiveness. Alternatively, this pervasiveness can be measured based on its ability to facilitate new bridges between technologies that would not be connected otherwise, or in other words, its *brokering* capabilities.

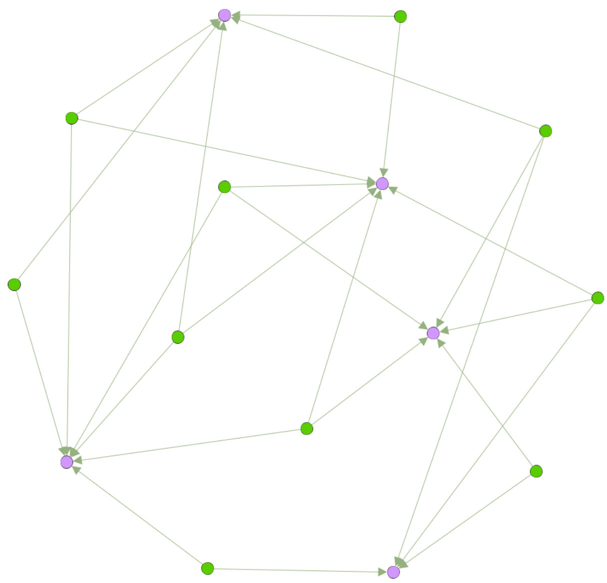


Figure 3.3: A bipartite technology network. Technologies are represented by yellow dots, while products are represented by blue dots

Innovation complementarities

A key feature of GPTs is their ability to spur innovation in the application sectors. The initial assumption is that most technologies are originally introduced as unrefined versions of their best self. As Bresnahan and Trajtenberg (1995) suggested, continuous efforts to innovate a primitive version of the GPT may lead to either a reduction of the ratio between input prices and performance (through learning) or to an increase in the generality of its functional applications. Consequently, the costs of implementation of the GPT for the downstream sector diminish, which may lead to further investments in R&D efforts to profit from the advantages provided by the new technology.

The dynamics of innovation complementarities have been formalized by Bresnahan (2010). Define the technological level of a GPT as T_G, and its rate of change $T_G^{'}$. Apply the same definitions to an application sector a out of A application sectors (AS), bringing its respective technological level and rate of change T_a, $T_a^{'}$. The rate of change in the social returns to innovation in a is a function of the rate of innovation in the GPT, the application sector, and a dummy variable X that stands for other exogenous factors, such as industry and market structure.

$$V_a'(T_G', T_a', X) \tag{3.1}$$

Note that the social returns for innovation in this function do not take into consideration potential reductions in production costs caused by the technological improvements. Note that the function V_a' has increasing differences in (T_G', T_a')[1], suggesting that there are social increasing returns if there is coordination in the investments in both the GPT and the application sectors.

Generalizing the effect over time and the entire GPT cluster, the social returns can be expressed as

$$\sum_{a \in A} \int [V_a^{'}(T_G^{'}, T_a^{'}, X) e^{-rt}] dt \tag{3.2}$$

The social returns *will be larger if* (T_G', T_a') *are increased together, and will be larger if all of them are increased in a coordinated fashion than if there is not a coordinated increase* (Bresnahan, 2010).

The success of the virtuous cycle of innovation presented above largely depends on factors such as the specific GPT and application sector's industry structure, and how private incentives for innovation are distributed between innovative actors. The GPT and applications sectors' intellectual property regime may increase or reduce the speed of adoption while at the same time having an important influence on the innovation rate. Bresnahan (2010) affirms that, other than these complex dynamics, some basic incentive structures are present in all of them and provides a summary model concerning the returns of innovators. Label λ_a, λ_G, λ_c the fraction of social value received by the innovators in the application sector, the innovators in the GPT sector, and the consumers, respectively. Consider $\lambda_a + \lambda_G + \lambda_c \leq 1$, to take into account the deadweight loss caused by systems of appropriation such as patents or trade secrets. Then the returns to investments for an agent operating in an application sector a is

$$\lambda_a \int [V_a^{'}(T_G^{'}, T_a^{'}, X) e^{-rt}] dt \tag{3.3}$$

"A higher rate of technical progress in the GPT sector [...] (T_G'), increases both private return to the innovator in a and the marginal return to increases in with T_a' [...]. Increasing differences mean that increases in T_G' will increase the incentive for innovators in AS a to make increases in T_a'. Similarly, increases in T_a' will increase the incentive for GPT innovators to make increases in T_G'" (Bresnahan, 2010). This virtuous cycle offers the opportunity to overcome diminishing returns into investments in technical change. However, it also creates a trade-off related to the division of the returns between the upstream and downstream sectors. While the upstream innovation makes it possible for the downstream innovation to occur, a higher λ_G reduces λ_a and the incentive to invest in innovation in the application sector. At the same time, a higher λ_a, reduces λ_G. Overall, considering that T_G' depends from the rate of technological innovation in all the sectors A (T_A'), then the aggregate incentive for investing in GPT innovation may be compensated. From the perspective of a GPT developer (upstream sector), the higher the number of application sectors, the higher the demand, increasing

[1] Name $T_G' = g$ and $T_a' = a$, thus $V_a'(g, a, X)$. Suppose $g \leq g'$, and $a \leq a'$ for all X. Then $\Delta_{g,g'} V_a'(a) = V_a'(g', a, X) - V_a'(g, a, X)$, making the function $\Delta_{g,g'} V_a'$ non-decreasing for all $a \leq a'$. Thus $\Delta_{g,g'} V_a'(a) \leq \Delta_{g,g'} V_a'(a')$, that can be shaped as $V_a'(g', a, X) - V_a'(g, a, X) \leq V_a'(g', a', X) - V_a'(g, a', X)$

the value of investments in the GPT. Formulating the private returns of investments in the GPT sector:

$$\sum_{a\in A}\left\{\lambda_{G,a}\int\left[V_a'\left(T_G'\left(T_A'\right),T_a',X\right)e^{-rt}\right]dt\right\} \tag{3.4}$$

where $\lambda_{G,a}$ represents the fraction of the social returns obtained by the GPT sector for a specific application sector, out of A application sectors. In this model, innovators in the application sector pick T_a' to maximize $\lambda_a V_a'$. If instead they aimed to maximize the rate of innovation at all levels, they would pick T_a' to maximize the effects on the whole production function $(\lambda_a + \lambda_G)\, V_a'$.

Other than this "vertical" externality, Bresnahan (2010) and Bresnahan and Trajtenberg (1995) affirm there is also a second, horizontal externality, that arises from the interaction between multiple application sectors. Any innovation or increase in investments in the GPT makes them better off by raising the GPT quality. The same is true for R&D investments in any other application sector because innovations in any of the application sector increase the rate of innovation of the GPT sector, which, in turn, increases the rate of innovation in the application sectors, including those that invested in the first place, as an effect of the positive cycle of innovation complementarities. Simultaneously, innovators in the application sectors act in such a way to maximize the returns for their investments in innovation, which largely depends on the degree of investments in other sectors. However, if every application sector follows the same strategy this behavior can result in a standoff, where application sectors invest the minimum necessary amount, slowing the pace of innovation. This issue can be partially solved by improving coordination, but that can prove complicated because the large number and variety of application sectors increase information asymmetries. So, in the presence of innovation complementaries between the GPT and application sector, the lack of incentives in one sector can create an indirect externality that results in a system-wide reduction in the innovative investment itself (Bresnahan and Trajtenberg, 1995). In the past, coordination was partially achieved because of the existence of predictable large demanders that had other incentives to sustain conspicuous fixed costs such as the strategic importance of the industry (in the case of public procurement) or to boost their technological dynamism (in the case of the private sector).

Time implementation

The concept of a time delay between the moment a technology is invented and when it enters into the economy is well established in the innovation literature. Considering that the innovation dynamics of a GPT are marked by the complex dynamics described above, it is not surprising that GPTs implementation lag is generally larger than average. As Bresnahan (2010) noted, *"the very idea of a GPT draws a distinction between raw technical progress (GPT invention) and the further innovation needed to create value-in-use (AS [application sector] invention.)"*. In the economics of innovation, implementation is generally modeled as an S-shaped diffusion curve (slow at the beginning, followed by a period of rapid growth, and then reaching an innovation plateau) that may depend on a variety of factors, such as supply or demand constraints, adjustment costs related to the need to substitute previous technologies, or information asymmetries.

Helpman and Trajtenberg (1994) suggest that GPTs spread in two phases: one in which the resources are diverted to the development of complementary inputs needed to take advantage of the new GPT, where productivity and stage output growth are negative, and a second phase, after the complementary inputs are already developed, where it is worthwhile to switch manufacturing with the new, more productive, GPT, that is characterized by an enormous increase in stage output, rising wages, and profit. In a later paper (Helpman and Trajtenberg, 1996), they expanded the model to consider the fact that multiple application sectors are operating contemporaneously in the real world. They determined that, while the initial stages of the spread of the new GPT are characterized by intermittent growth[2], after a "critical mass" of application sectors have made the necessary complementary investments, we assist to a surge in growth, together with a rise in wages and profit. Their models show that there may be various factors affecting this diffusion process and that shape the rate and the trajectory of the transition towards a new GPT, such as the rate of adoption, the technological distance of the new GPT from the previous one (such as the transition between steam power and electricity), the propensity to invest in R&D, the order of adoption of the various application sectors, and other exogenous factors. Moreover, it is essential to note that different adoption orders in the application sectors may trigger multiple second phases (those characterized by rising productivity).

3.1.2 Methods of Inventing

Research technologies are essential tools for innovation. They provide new ways to produce knowledge, opening new scientific fields and expanding the knowledge frontier of what it is possible to discover. Griliches (1957) was the first economist that brought to attention the critical role that the development of new research tools has on the innovation process. He did so during a study on hybrid corn, founding the technological category of IMI. However, even if he formalized this concept, the definition that he provides: *"It was not a single invention immediately adaptable everywhere"*, together with the econometric evidence that shows multiple waves of innovation, appears to blur the differences between GPTs and IMIs. Some scholars, such as Darby and Zucker (2003), misapplied Griliches' breakthrough concept to the field of nanotechnology, limiting it to be an invention that provides a business opportunity across a variety of ranges. On the other hand, (Cockburn et al., 2018) suggest a different idea of IMI, where they are closer to innovations in research tools or innovation in processes. This seems to be more in line with the initial intuition of Griliches since the value of the invention of hybrid corn was not in the hybrid corn itself rather than in the invention of a *"widely applicable method for breeding many different new varieties"* (Cockburn et al., 2018). Hentschel (2015) analyzed how some scientific instrumentation shifts from the specific context in which it is developed, with a particular use, towards a more general use, mentioning the invention of optical lenses. Putting two lenses in line allowed for the first time to observe things that were not possible to discern at the naked eye and that completely changed the scientific paradigms of entire research fields, such as medicine, astronomy, and chemistry, and allowed the development of new research fields. This reduced the gap between theory (what was thought to be) and the empirical observation of real phenomena (what it is possible to experience), altering the preindustrial scientific community's scenario. *"[..]*

[2]Since the complementary investments are used only by a few early adopters

much of the seventeenth-century scientific revolution was made possible by better instruments and tools" (Mokyr, 2018), which, in turn, were the prerequisites for the explosion of productivity of the first industrial revolution. Another famous example of IMI is the much-debated Oncomouse patent, whose applicants aimed to expand the patent scope to the method for genetically engineering. This sparked outrage in many researchers, which refused to pay royalties for a fundamental and general method of biotechnology.

Differently from the concept of GPTs, the applications of an IMI are limited to a single (or a small number of) application sector, such as agriculture (in the case of hybrid corn) or biotechnology (in the case of the Oncomouse). IMIs can reduce the cost of specific innovation activities while shaping the development of new *"innovation playbooks"* (Cockburn et al., 2018). They expand the number of tools available to the researcher, thus extending the addressable research question's horizon and enhancing the rate of innovation and discovery.

Recalling the theory of recombinant innovation by Weitzman (1998) and the network-based technology model developed by Korzinov and Savin (2016), new IMIs act as "facilitators" in finding useful connections between different pieces of knowledge. In network science terms, nodes in the network represent ideas, and edges represent their combination. The introduction of a new IMI increases the likelihood of creating an edge between previously interconnected nodes and of reinforcing existing ones. While traditional indicators of innovation such as patents and scientific papers seem to be increasing at increasing rates, recent empirical evidence indicates that research productivity has been decreasing and that new ideas are getting harder to find (Bloom et al., 2020). This seems to confirm what Weitzman (1998) proposed in his original paper, stating that the limit for the generation of new ideas (innovation) is marked not by the potential to generate them but rather by the ability of researchers to process them and evaluate their validity. The number of possible combinations of ideas increases at a much higher rate than the research community's ability to evaluate them, but the cognitive abilities of human beings can only process a limited amount of information. The knowledge landscape is not only becoming vaster and vaster but also more fragmented. The expansion of science leads to increased specialization, the flourishing of sub-disciplines, and researchers can focus on fewer and fewer topics (Jones, 2009). Thus, in the absence of improvements in research methodologies, the pace of innovation is bound to decrease. Improvements in IMIs may partially solve this issue.

While the impact of new research tools on innovation is a well-established concept in innovation studies, currently there is no econometric model that tries to explain its complex dynamics. However, formalization is fundamental to understand how the evolution of research tools influences technological advancements, especially now that the literature has started noticing several technologies that possess features typical of GPTs and IMIs. Drafting a formal definition for IMIs:

1. they expand the knowledge frontier of science by increasing the set of tools available to researchers;

2. they reduce the distance between recombinable ideas, increasing the rate of scientific advancement;

3. they have genericity of use: they can be applied for inventing different products in a limited number of application sectors.

At first glance, it may seem that the genericity of IMIs can be assimilated to the pervasiveness of GPTs. However, while the former involves the possibility to use a method or a technique as an intermediate step *within an application sector* for the development of new methods and technologies, the latter is about the ability of a technology to spread *across different application sectors* of the economy as a final product.

3.1.3 General Methods of Invention

In rare cases, some technologies present features of both GPTs and IMIs, leading scholar to come up with a new technological category: General Method of Invention. *"A GMI is a [I]MI that is applicable across many domains. Similar to GPTs, development of the GMI and complementary developments in application domains are mutually reinforcing"* (Bianchini et al., 2020). Given that such technologies are extremely scarce, there is not a shared terminology yet, and they have been called using different nomenclatures, such as General Method of Invention, General-Purpose Invention in the Method of Inventing or General Meta-technology (Bianchini et al., 2020; Cockburn et al., 2018; Agrawal et al., 2018). For this dissertation, I retain that the term suggested by Bianchini et al. (2020) is particularly descriptive, thus I will refer to this class as General Method of Invention (GMI).

As one can imagine, GMIs have an immense potential for exponentially increasing the innovation rate. At present, the innovation literature agrees only on one of such technologies: digital computing. Computing technologies had an enormous impact across economies and societies, of which we are still experiencing its full potential. Recently, AI is increasingly regarded as a GMI candidate. Unfortunately, the absence of previous examples of such technologies and the complex nature of how they affect the economy prevents scholars and researchers from identifying precise characterizations.

3.2 Is AI a GPT?

The theoretical framework presented in the previous section is crucial to understand how AI technologies will impact the economy. After the commercial success of the new generation of AI technologies, several papers have suggested that AI should be categorized as a GMI, since it is possible to identify features of both GPTs and IMIs. Recalling the definition provided in section 3.1.1, a technology is considered a GPT if it has a broad scope for improvement and elaboration, if it can be applied in a large number of sectors of the economy, and if it has strong complementarities with existing and new technologies. From a qualitative perspective, AI fulfills these the GPT requirements:

Rapid growth AI algorithms can improve themselves over time through learning, further enhancing their potential applications. Pratt (2015) identified as one of the core capabilities of AI its ability to share knowledge and experiences almost instantaneously, provided an internet connection. In this way, once a machine learns a new skill, that piece of knowledge can be shared with all the other devices in the network, thus increasing the data available and improving performance, further extending the possible fields of application.

Pervasiveness As it was mentioned in chapter 1, modern AI technologies are essentially prediction technologies, which are a crucial enabler for improving decision-making, which is present in any sector to different degrees. Decision-making may involve choosing where to store materials in a storage unit, whether or not to operate a patient suspected of having a tumor, or selecting the best financial assets for a long-term investment. AI technologies may improve speed and precision, desirable performance indicators in a variety of different areas. Implementations of AI technologies already surround us: machine translation has spread widely; search engines use AI to select the best results according to our query; it is possible to use speech-to-text to control devices remotely; we increasingly rely on navigation apps to find the best route in unknown environments. AI is impacting several sectors of the economy, and the number of its applications is increasing steadily.

Strong complementarities with existing sectors The introduction of AI in various application sectors has the potential to spur complementary innovation. New methods must be found to increase productivity, and new challenges need to be faced. Solutions found in the downstream sector can later be applied in the upstream sector, starting a virtuous cycle of innovation in algorithms and data management.

3.2.1 The role of AI in the research process

Artificial Intelligence applications are not limited to commercial settings. Since their early developments, AI technologies found most of their applications in academia. Most quantitative research regards finding patterns and causal relationships, and ML systems are particularly efficient when dealing with high-dimensional data. The development, optimization, and application of multi-layer neural networks to these sets of data allowed researchers to expand the set of research questions that can be feasibly addressed, thus changing the way research is conducted (Cockburn et al., 2018). Agrawal et al. (2018) distinguished two different uses of AI in research: search and discovery. The *search* stage concerns finding relevant information for the researcher, drawing from different fields. In other words, AI tools involved in the search process aim to optimize the initial stage of the researcher's work, facilitating the process of identification of the critical factors to take into account when investigating a phenomenon. In the *discovery* stage, AI tools are more focused on predicting which *"combinations of existing knowledge that will yield valuable new knowledge across a large number of domains"* (Agrawal et al., 2018). In the following paragraphs, both of these aspects are exposed in detail.

Limited resources, unlimited knowledge

According to Bornmann and Mutz (2015), scientific papers are subjected to average growth of 8 to 9 % per year. Considering the increasing investments in research activities, the upward trend in the number of PhD graduates, and the tendency towards knowledge fragmentation, it is doubtful that individual researchers can deal with the constantly increasing knowledge at their disposal. A temporary solution was found by expanding the size of research teams, but it is not sustainable in the long-term. In this context, AI assistants can prove to be a valuable asset. Various AI-based tools have been developed to help researchers navigate significant amounts of unstructured knowledge, typically in the form of search engines or knowledge graphs.

Perhaps **Google Scholar** is the most known AI-based tool for identifying papers and studies given a keyword-based input, but also the less sophisticated one. It is generally used to navigate papers and books, find cited/citing papers, and obtain bibliographic information. In recent times, **Semantic Scholar**, another search-engine developed by the Allen Institute for Artificial Intelligence in Seattle, has gained popularity. Whereas the ranking function of Google Scholar is mostly based on citation indices, the approach of Semantic Scholar relies on a semantic analysis of papers' titles and abstracts, thus making it possible to obtain more accurate results and optimizing it for scientific discovery, allowing for a more relevant and efficient view of the literature of choice. It was initially used for computer science, geoscience, and neuroscience, but it has been progressively extended to other academic areas. It provides additional functionalities, such as the construction of personalized *Research Feeds* based on a small group of papers chosen by the users, or the *TL-DR* function (Too Long, Didn't Read), that summarizes the abstract of scientific articles in maximum two lines. Used in the context of biomedical research, **Meta** is a machine-learning web platform that aims *"to map biomedical research in real-time"*. By analyzing and connecting millions of scientific outputs, it aims to provide researchers with a comprehensive view of how biomedicine sub-fields are evolving. Substantially different from the other mentioned tools, **SourceData** is a data discovery tool promoted by the European Molecular Biology Organization (EMBO) that explicitly targets biomedical scientists. It tries to address the issue of biomedical data locked away in published scientific illustrations that are not indexed by traditional search tools. Starting from the metadata that describe scientific figures, SourceData converts them in a standardized format that allows for the linkage of data originated by different experimental papers. In this way, it aims to reduce the inefficiencies caused by the lack of structured communication within the biomedical scientific community.

AI as a research technology

In the second stage of the research process, discovery, Artificial Intelligence technologies can effectively be used to speed up experimental validation and/or dealing with high dimensional data.

Recalling the definition of the knowledge space provided by the theory of recombinant innovation (Weitzman, 1998), where the successful combination of ideas originates new ideas, improvements in prediction technologies (such as AI) can result in fewer resources wasted during the research process. In any research effort, prediction [3] is used to map all the possible combinations of ideas to prioritize those that are more likely to yield new knowledge. Assuming that inputs are the already existing ideas and data over past combinations, the prediction function provides as output the probability for each possible combination to lead towards a successful new idea. These probabilities can be ranked to promote the prioritization of the combinations that are more likely to lead to a successful idea, helping scientists to choose the right path. Consequently, every improvement in prediction technology represents a potential source of increased productivity in the research output.

Compared to other prediction technologies, such as parametric modeling, AI has the advantage of being particularly suited for discovering non-linear relationships in high-dimensional data. While parametric modeling techniques are constrained by the need to extract features (explanatory variables) by hand before the statis-

[3] In its different forms, such as educated guesses, theoretical models, statistical parametric modeling, or others

tical analysis, algorithmic modeling considers the relations between input data as complex and unknown. In parametric modeling, the analyst specifies the dependent variable, the predictor variables, the functional form of the relationship between them, and the stochastic form of the disturbance term. Using theory, the researcher explicitly formulates the association of real-world data to the variables. While this approach has the advantage of allowing total explainability, when it is used to analyze more complex phenomena, the human brain cannot manage so many variables and information, undermining the model validity.

Instead, DL-based prediction aims to find the function $F(x)$ that from x predicts the response y. The validity of the model is measured by the accuracy through which the algorithm predicts the correct output. Advances in AI architecture allowed for substantial improvement in high-dimensional spaces with complex non-linear interactions, giving AI a comparative advantage over parametric modeling in analyzing large amounts of real-world data (Agrawal et al., 2019). Theory is present, but it enters the research pipeline in a different stage. Rather than being used to build a predictive model (as in parametric modeling or educated guesses), it is used during previous and later phases such as data gathering, experimental design, and result interpretation. Recently, these steps have started a process of automation under the supervision of the researcher. Representation learning automatically extracts what the algorithm identifies as the most relevant features. Robotic laboratories are used to design and perform experiments. Theory is used to understand why AI chooses them, to rule out algorithmic confusion and biases and build new conceptual frameworks.

When the challenges presented are combinatorial, the discovery process has the potential to be entirely automated. Examples of such research questions are *"how do I combine molecules to create a new material with these characteristics?"* or *"how do I build a drug that has these specific effects?"*. One of the most promising applications of such closed-loop platforms is material discovery, where the combinatorial space of the molecules can be analyzed using Deep Learning to predict the structure of new materials and their properties. This process is known as inverse design (Sanchez-Lengeling and Aspuru-Guzik, 2018). An initial library of molecules selected by the researcher is filtered based on focused targets such as ease of synthesis, solubility, toxicity, and stability. Closing the loop would require integrating these simulations in a robotic laboratory able to conduct experiments. According to Aspuru-Guzik and Persson (2018), current research related to closed-loop systems for material discovery is focusing on improving *"AI-based predictions of new materials and their properties, autonomous robotic systems for synthesis and experimental data collection, data analysis [techniques] such as feature extraction, AI-based classification and regression of results, and decision modules to drive optimal experimental design for subsequent experimental iterations"*. Currently, the bottlenecks of the process are represented by experimental design, where human intuition still plays a fundamental role, and the lack of data on *"failed"* experiments, crucial to teach the AI how to build an integrated research pipeline. Machine learning methods such as deep learning are indeed a promising tool for discovery, especially where the complexity of the phenomena makes it difficult for the researcher to adopt more traditional methods. However, even in these limited cases, the analysis they provide still needs to be subjected to experimental validation and thorough examination by human researchers.[4]

[4] In appendix A is discussed the current debate on Autonomous Discovery Systems

3.3 AI in the market

The extent to which Artificial Intelligence can be considered a GPT or a GMI has significant consequences for economic growth and innovation but these technologies are already likely disrupt the economy in unexpected ways. In recent times, some scholars started investigating the effects of the deployment of AI systems on market structure, competition, and the consumers' decision-making processes.

3.3.1 AI and market structure

Without the introduction of specific legislation, AI technologies are likely to enhance current disruption provoked by the expansion of the scope of the digital market. Coupled with an enormous decrease in labor per customer ratio, the digital economy is already marked by extreme network externalities and strong economies of scale and scope that may result in anti-competitive behavior. In a report on competition Crémer et al. (2019) explored the mechanisms of the digital economy in detail, identifying three key features of the digital market:

- **extreme returns to scale:** after having surpassed a threshold, digital products have a marginal cost of consumption that is close to zero. This brings a significant competitive advantage for an already established business. Moreover, data present increasing returns to scale. While in terms of performance, data exhibits decreasing marginal returns (accuracy improves at a decreasing rate according to the quality and quantity of data supplied during training), a firm may experience a comparative advantage in providing a service that is just marginally better than its competitors. A slightly better prediction leads to a better service, which in turn leads to a higher share in the market Goldfarb and Trefler (2019);

- **network externalities:** in many cases, the attractiveness of a product or service relies on the number of users that already use it, such as in the case of messaging services or social networks. After a business has conquered market shares, it may be challenging for newcomers to convince users to migrate to its products, even when their products are cheaper or more innovative. Additionally, the incumbent may be actively promoting its own complementary service;

- **data advantage:** first-movers have no incentive to promote easy transfers of users' data towards competitors, and the current IPRs framework on database rights leave space for ownership claims regarding compilation of users' data, thus giving grounds for the systematic obstruction of the transfer of users' data and information towards competitors. Additionally, users' data can be used to develop new innovative products, thus providing them with a competitive advantage in innovation (Goldfarb and Trefler, 2019).

These mechanisms are leading towards increased concentration of the profits generated in the digital economy in a few dominant platforms that have strong incentives to reinforce their dominant position (Crémer et al., 2019). Bloom et al. (2014) suggest that widespread use of AI technologies may speed up the concentration process.

Additional concerns related to the widespread use of AI technologies are related to data ownership. Since data are an essential input for all AI-based technologies, firms that own large datasets may create barriers to entry, further increasing economic externalities. Goldfarb and Trefler (2019) suggest that *"competition in the market is substituted by competition for the market"*. This is because, in AI-powered business, both the expenses related to personnel and data acquisition can be optimized to reach a variety of objectives such as the development of new products, optimizing the total costs. The lack of exploitation of already gathered data would result in firm-level inefficiencies because businesses would not have extracted their full value. Then the firm has two choices: either sell the data or invest in R&D to expand the pool of goods that it puts into the market. However, privacy legislation may impose severe restrictions on buying and selling data, thus disincentivizing the creation of a data market. This mechanism increases the firm's incentive to find new ways to produce economic value from the unused data in its possession through innovation and expansion in new markets. However, while this mechanism may have a positive effect on innovation (Nuccio and Guerzoni, 2019), the absence of a data market impedes other actors from enjoying the innovation benefits derived from data repurposing and thus favors incumbents over new entrants. Multipurpose firms such as Google, Microsoft, Apple, Facebook, and Amazon can only allocate a certain amount of resources to R&D compared to data repurposing possibilities. The exclusive ownership by a firm of a particular set of data can potentially erect data-driven entry barriers to markets that have not come into existence yet, thus impairing the speed of innovation. Additionally, this provides corporations possessing large datasets with an extreme prescriptive power over the direction of innovation (Cockburn et al., 2018). On the other hand, forcing the incumbent to share data could diminish the incumbent incentive to invest in data creation, obtaining a reduction in the pace of innovation. An alternative proposition made by Korinek and Ding Xuan (2018) points out that one way to reduce inequality and anti-competitive behavior may be to reduce the rents earned by innovators derived from IPRs, specifically patents, but this may result in a lack of technology disclosure, thus slowing down innovation and reducing transparency over the treatment of users' data.

3.3.2 Some market externalities of Artificial Intelligence

The benefits brought by AI technologies are affecting both firms and consumers. Notably, there are cost-reductions in search, replication, transportation, tracking, and verification (Goldfarb and Tucker, 2017). Algorithms can improve consumers' decision-making by organizing information based on price information and other criteria, such as product quality and consumers' preference. However, the spread of AI technologies can also provoke market failures, such as behavioral biases and cognitive limits. Decreased search costs do not necessarily favor consumers. As argued by Goldfarb and Tucker (2017), service providers and companies can elaborate strategies to manipulate the search process to trick consumers and increase profits. While AI technologies positively affects the supply side of the market by reducing production costs, increasing the quality of existing products, and improving resource allocation, it is also used to provide optimal commercial strategies to sellers that can use Recurrent Neural Networks to maximize their profits. While a complete review of these mechanisms is outside of the scope of this dissertation[5], I will briefly introduce how AI may impact the market dynamics through the use of dynamic pricing.

[5]See Abrardi et al. (2019) for a more complete review.

Dynamic pricing

Dynamic pricing technologies optimize and adjust prices based on factors such as stock availability, capacity constraints, competitors' prices, or demand fluctuations. On the bright side, this allows the market to constantly be in equilibrium, thus preventing the mismatch between demand and excess of supply. On the other hand, this puts sellers that do not use these technologies (for various reasons, such as investment cost, low innovation capabilities, or others) at a disadvantage, and makes consumers' decision process more difficult, since they are forced to buy in an environment characterized by constant price fluctuations. While for some aspects, this has a positive effect since it puts companies under the pressure to innovate (Cockburn et al., 2018), there is a concrete risk of algorithmic collusion and thus a reduction of the consumers' surplus.

Moreover, for the first time in economic history, dynamic pricing provides an opportunity to systematically apply first-degree price discrimination[6]. Through extreme personalization obtained by gathering data on users' online behavior, platforms can often deduce the consumer's reservation price and differentiate the price (Milgrom and Tadelis, 2018). Gautier et al. (2020) presented some examples of such practices[7], even if it seems that, for the moment, they are not the industry standard. They suggest that this may be attributed to different factors, such as technical barriers and consumers' perception[8].

The use of dynamic pricing also creates issues related to algorithmic collusion. Several experiments were drawn on sample market structures to understand how the interaction of multiple automatic pricing systems may take place. Last-generation algorithmic pricing is based on ML techniques, such as Recurrent Neural Networks, in which the software learns the optimal strategy by trial and error without requiring any specification related to the economic model. Klein (2019) showed that well-established algorithms learn autonomously to set supra-competitive prices in a sequential game. Calvano et al. (2020) in another experiment show that when multiple dynamic pricing algorithms are left to interact with simultaneous moves, they learn to play sophisticated collusive strategies, difficult to detect. Unlike human-based collusion, AI-based collusive strategies are *"robust to perturbations of cost or demand, number of players, asymmetries and forms of uncertainty"*. This is particularly problematic for antitrust authorities since algorithmic collusion has no trace of concerted action (automatic pricing systems do not communicate with one another), thus escaping the current regulatory framework. Additionally, liability determination is problematic, as in all situation where AI systems actively take decisions autonomously.

3.4 Conclusions

This chapter touched on many arguments related to the effects of the widespread diffusion of Artificial Intelligence technologies in the economy. First, it provided a theoretical background for the classification of

[6] First degree price discrimination occurs when prices at which goods and services are sold are equal to consumers' reservation price

[7] Such as in the case of Amazon's scandals regarding price discrimination over DVDs or mahjong tiles. See Cavallo (2018) and Townley et al. (2017) for reference.

[8] Since first-degree discrimination is widely considered an exploitative practice, when aware of these mechanisms, consumers tend to behave strategically, either by limiting the amount of information they reveal or by creating multiple accounts and masking their IP address.

technologies based on their impact on innovation. We have seen that General Purpose Technologies may bring an increase in productivity and, as Bresnahan and Trajtenberg (1995) suggested, they can become *"engines of growth"* creating a virtuous cycle of innovation. *"Inventions of a Method of Inventing"* identify research tools as one of the key elements for increases in the rate of innovation. Combining these two categorizations, many scholars have argued in favor of the development of the concept of *General Method of Inventing*, which include the technologies with characteristics of both IMIs and GPTs. Second, some qualitative evidence was provided regarding how Artificial Intelligence technologies should be included in any of these categories. The evidence suggests that AI can be both used in various commercial settings and as a research technology, indicating that it is a potential GMI candidate. On the other hand, given the rarity of these technologies, their economic effects are challenging to forecast, even if they likely increase the rate of innovation and technological change, increasing the productivity of both research and industrial output. Finally, a point was made on the effects of AI in market structure and some of the externalities of the widespread use of AI as a market tool. In the absence of appropriate regulation, AI technologies will likely increase the speed of concentration of market power in a few players that do not compete with each other in markets but that primarily compete for the conquest of new markets, constantly expanding the scope of their economic activities. These conglomerates are in an advantageous position when compared with new entrants because of the repurposing capabilities of data. Moreover, when AI technologies are introduced in market transactions, they indeed bring some benefits to both consumers and producers, but they can also create negative externalities caused by their widespread use, such as dynamic pricing, first-degree price discrimination and algorithmic collusion.

Chapter 4

Measuring Artificial Intelligence with patent data

In chapter 3 I presented how episodes of acceleration in economic growth can be driven by particular technologies, often referred to as General Purpose Technologies (GPTs). Artificial Intelligence is configuring as such a technology, while simultaneously being widely used in R&D efforts, thus making it a good candidate for the new technological category of General Method of Invention. However, while there is sufficient qualitative proof on the GPT character of AI technologies, at present there is no research that verifies these claims from an empirical perspective. It is common practice for scholars studying the economics of innovation to rely on the quantitative analysis of patent documents to investigate the diffusion of technology in society. This chapter aims to build on the rich literature on GPTs and Artificial Intelligence to explore whether AI could be considered as such, and to serve as the basis for further empirical research regarding the categorization of AI as a GMI.

4.1 Literature review

In the past few years, the gain in popularity of AI technologies has awaken the interest of patent offices and scholars at a global scale. As a consequence, several studies were made to analyze how AI is impacting the innovation panorama, to understand its possible effects on the economy, generally using patent document and research papers as primary data.

4.1.1 Studies on GPTs

After posing a theoretical foundation on the concept of General Purpose Technology, the economic literature started developing ideas on whether it was possible to develop quantitative indicators to operationalize it and detect possible GPTs at an early stage.

Moser and Nicholas (2004) applied criteria proposed by Henderson et al. (1995) to identify whether electricity could be considered a GPT. They found that between 1920 and 1928 there was an enormous increase in the rate of patenting, and thus applied a range of indicators (originality, longevity, and generality) to test this hypothesis. Originality is a backward-looking citation measure that aims to identify the exact arrival dates for influential innovations, longevity is the speed of obsolescence for inventions in different industries, and generality is a classical Herfindahl-Hirschman index that measures the technological diversity range of citing patents.[1] They also proposed to use the number of technological classes assigned to each patent as a measure of patent scope. Their findings contradicted the expectations that electricity should be considered a GPT.

Hall and Trajtenberg (2004) focused their efforts on the development of better suited indicators to detect GPTs using patent data. They used citation-based measures as a proxy for determining the GP-ness of a random sample of patents. According to their model, GPT innovations should have many citation form outside their particular technology areas or from industries outside the one in which the patented invention was made. Simultaneously, since GPTs should also be able to spawn innovation in their original application sector, following a pattern of cumulative innovation, GPT patent should also have many citations within their technology area, thus making citing technologies to be subject to increases in the rate of innovative activity. They also suggested that, since GPTs take a long time to spread in the economy, citation lags (the distance between the cited and citing patents) could be longer than average, but that, in the long run, GPT patents should turn up to be highly cited. They expanded the use of the Generality measure, correlating with the absolute number of citations, and they suggested to examine also citations of the second degree[2]. It is important to note that, in computing the Generality index, when a patent had no citations, G is not defined, while if the number of citations is equal to 1, they chose to assign $G = 0$. In their analysis, they decided to test the validity of the Generality index based on five different classification systems, three based on technological classification[3], and two related to industry[4]. This choice was motivated by the fact that *"it may be argued that a GPT is not likely to manifest itself as a series of citations by patents in different technology classes, but rather as citations by firms in different industries"* (Hall and Trajtenberg, 2004). To measure the phenomenon of innovation complementarities, they proposed to search for the patent classes that presented a rapid growth in patenting, and then to compare it with the growth rates of the citing patents.

Jovanovic and Rousseau (2005) adopted a different approach, basing their GPT identification process on historical data on the diffusion of electricity and information and communication technologies (ICTs) in the economy. As indicators of pervasiveness they identified the shares of total horsepower in manufacturing by power source, and the share of IT equipment and software in the aggregate capital stock for ICTs. They identified the technical advancement as a function of the cost for an decline in prices, an increase in quality, or both. This choice was motivated by the fact that capital as a whole should be getting cheaper faster during a GPT-era, especially if it involves the new technology. Finally, they used the rate of patenting and investments to identify the technology's ability to spawn innovation, finding a worldwide increase after the introduction of both technologies. The latter is motivated by the fact that a GPT should increase the rate of investment

[1] The standard formula for generality is computed as $G_i = 1 - \sum \frac{(s_{ij})^2}{NClass_J}$ and it is presented in detail in section 4.3.2.
[2] The citations of a citation
[3] The US patent classes, Hall-Jaffe-Trajtenberg technology subcategories, and the International Patent Classification
[4] Industry of manufacture and Industry of use, based on Silverman (2002)

because it requires the refurbishing of both capital and the skills of labor.

Feldman and Yoon (2012) drew from previous literature to verify whether a specific technology, the Cohen-Boyer's rDNA splicing techniques, could be considered a GPT. They built three indicators aiming to verify three characteristics of a GPT: technological complementarity, technological applicability, and technological discontinuity. Technological complementarity was computed by counting the number of industrial sectors (identified using Silverman (2002) correspondence table) of first-degree forward citations, technological applicability was instead computed using the proportion of the total forward citations present in patent classes outside the original patent class, while technological continuity is tied to verify the hypothesis that complementary innovation spawns in geographical areas in proximity to those of the original patents. Feldman and Yoon (2012) computed these indicators and then compared them to a control sample.

The most recent work on determining alternative indicators to identify GPTs in patent data is the one of Petralia (2020). This paper is particularly interesting because it is based on different conceptualization of GPTs. Rather that considering a technology either as a GPT or a non-GPT, he speaks of a degree of GP-ness that every technology potentially has. This suits well with the network based view of the technological space mentioned in chapter 3, where the extent to which a technology can be considered a GPT is measured on the basis of how much it is capable to potentially connect to other technologies. Although he used the rate of patenting rates in specific fields to identify the first characteristic of GPTs, he adopted a different strategy to identify its pervasiveness and the complementarity between new or existing technologies. In particular, he focused on the degree to which Electric and Electronic (E&E) and Computing and Communication (C&C) technologies are GPTs. Starting from a list of selected keywords for each category (E&E and C&C), he identified technological classes that included the use of these technologies by counting the co-occurence of technological classes based on either the originating patents (Use Complementarity, or UC), or citing patents (Innovation Complementarity, or IC).

4.1.2 Studies on Artificial Intelligence

The ways AI technologies are spreading in the economy were the subject of several empirical studies, that aimed to map the AI industry or understand whether they could be considered a GPT, or a GMI. Since AI technologies are particularly difficult to define, much attention has been devoted to correctly identify AI-related scientific publications and patent documents.

The first study specific of AI technologies was conducted by Cockburn et al. (2018), and aimed to map the different fields of AI to evaluate how it could be characterized. Their analysis *"draws on two different datasets, one that captures a set of AI publications from Thompson Reuters Web of Science, and another that identifies a set of AI patents issued by the U.S. Patent and Trademark Office"*, (Cockburn et al., 2018). To identify patents related to AI technologies, they assembled a dataset of patents associated with AI-specific codes of the U.S. Patent Classification System and a second dataset of patents by conducting a title search on patents, using a list of keywords. By assembling these two subsets and dropping duplicates, they created a sample of 13615 patents from 1990 to 2014. Subsequently, they subdivided the sample in three major categories of AI technologies: learning system patents, symbolic system patents, and robotics patents, with the objective of mapping the distribution of different kinds of AI technologies. They only focused on the pervasive character of GPT, and they encountered empirical evidence that learning-oriented AI systems seems to have some of the characteristics of being a GPT, since they are applied in a variety of sectors, especially in the *"most technologically dynamics parts of the economy"* (Cockburn et al., 2018).

Klinger et al. (2018), used a combination of open-source paper databases (arXiv, Microsoft Academic Graph, and Global Research Identifier) and CrunchBase, to measure the geography and impact of Deep Learning (DL), the most promising field of AI. They argue that AI could change world economic geography, affirming that governments now have launched in an *AI race*. To identify AI-related papers, they used Natural Language Processing techniques, and then matched the scientific areas with CrunchBase industry sectors. To verify whether DL could be considered a GPT they measured the growth in publication of Deep Learning-related papers, finding that the share of DL papers in the total has increased from 3% before 2012, to 15% afterwards. Then they measured the number of DL-related papers in different arXiv subjects, finding that there is a visible upward trend in the relative importance of DL in many computer science subjects. Finally, they evaluated the impact of DL in other fields using citations as a proxy of knowledge diffusion, finding that in most arXiv subjects, DL-related papers occupied a larger share than average in highly cited papers.

In a similar fashion, Bianchini et al. (2020) conducted an analysis of citation patterns of scientific publications (based on the combinations of arXiv and Web of Science data). They noticed that in recent years there has been an increase in the research activity involving Deep Learning technologies in all areas of science taken into examinations. They confirmed Klinger et al. (2018) findings and affirmed that Deep Learning publication activity also grew *"relative to the overall number of papers in scientific areas"*. Moreover, they observed that different research fields reacted in different ways to the introduction of Deep Learning, as a signal that there may have been field-specific bottlenecks that were solved in various points in time. Additionally, they propose that the fact that the proportion of DL articles containing "computer science" as one of their labels is progressively diminishing suggests that *"the diffusion of DL into application domains began with an interdisciplinary effort involving the computer sciences, breaking its way into 'pure' field-specific research within the various application domains"*. This observation adds to the precedent evidence in confirming that the mechanism of innovation complementarities is taking place in, at least, academia.

Two additional empirical studies on AI technologies are worth mentioning, not because they are targeted to determine whether AI should be considered a GPT, but rather because they were commissioned to assess the state of AI technologies in the economy. They are based on patent analysis and were commissioned by important players in the innovation scenario: WIPO, and USPTO.[5]

The WIPO study on AI was commissioned in the context of the 2019 WIPO technology review (Benzell et al., 2019) to assess the state of AI research and technology. WIPO analyzed patent data to identify the patent families related to AI technologies based on the Questel Orbit patent database. Of particular interest is the query

[5]World Intellectual Property Organization and United States Patents and Trademark Office.

strategy used to identify AI-related patents, since the one I adopted in the empirical study for this dissertation is largely based on it[6]. To overcome the limits of keywords-based queries, WIPO focused on used a combination of technology classification codes and keywords that led to the gathering of 339828 patents related to AI. Then, using a hierarchical clustering scheme, WIPO was able to divide the technology macroarea in different clusters based on an adaptation of the ACM classification scheme (Association for Computing Machinery, 2012), first by subdividing the different technological areas of patents in three categories (AI techniques, functional applications and application sector) and then by clustering the technologies based on different subcategories, allowing them to map the evolution of the AI industry. Among their findings, WIPO reported that AI-related inventions are increasing at an impressive rate, since over half of the identified AI inventions were published after 2013. They identified the area of Machine Learning as the fastest-growing technique, boosted by implementation of Deep Learning and neural networks, and that, starting from 2002, there has been an acceleration in the number of patents regarding AI technology.

Finally, the study commissioned by the USPTO and produced by Toole et al. (2020), aimed to map the AI technological scenario of US patents. This study is of particular interest because it used a supervised machine learning algorithm to identify patents involving AI technologies. The performance for individuating AI patents was later evaluated against the techniques of both Benzell et al. (2019) and Cockburn et al. (2018) and compared with a human-based examination by AI experts. It resulted that this AI-based classification system outperformed traditional retrieval strategies, reaching an accuracy comparable to that of human examiners. Among its key findings, the study reports that AI has become increasingly important for invention and it spread across technologies, inventor-applicants, organizations, and geography. The share of annual AI patent applications from 2002 to 2018 has increased by more than 100%. Additionally, the USPTO report shows that AI technologies are increasingly expanding towards new areas of invention, since the percentage of technology subclasses present in at least two granted AI patents increased from 10% in 1976 to 43% in 2018.

These studies seem to confirm that Artificial Intelligence technologies are rapidly developing and spreading in the economy, thus favoring the interpretation that AI should be considered a GPT. Moreover, the increase in research output and the widespread diffusion of DL in academia provides an empirical basis to the claims that suggest that AI has the potential to revolutionize the innovative process, by building a new *"innovation playbook"* that involves the use of large datasets and learning algorithms to both increase the efficiency of individual researchers and discover patterns in complex sets of multivariate data. As Brynjolfsson et al. (2017) suggested, Artificial Intelligence *"has the potential to be pervasive, to be improved upon over time, and to spawn complementary innovation, making it a candidate for an important GPT"*.

4.2 Research goals and data

Building on the literature presented above, the analysis that follows aims to:

[6] A summary of WIPO query strategy can be found in appendix B

1. find additional evidence in patent data of the fact that AI technologies should be considered a General Purpose Technology;

2. explore the industry structure of the AI field in terms of patenting activity.

The chosen data source are patent documents drawn upon a subset of PATSTAT 2018 Autumn, that provides access to worldwide patent statistical information. Since 2018 data is incomplete, the search was restricted to applications filed from 01/01/1995 to 31/12/2017 under the Patent Cooperation Treaty (PCT), which provides the possibility to seek patent protection for an invention simultaneously in a large number of countries by filing one *international* application instead of filing several separate national applications. This choice was motivated by the assumptions that:

- patents that are filed using the international patent application scheme are the ones most likely to being applied in a variety of different countries. The PCT application procedure is more expensive than national applications so, for an applicant to opt for the international route, the revenues generated by the protected invention needs to be sufficiently high to justify the expenses;

- it is only with the harmonization of the international patent legislation through TRIPs that the PCT route started to be consistently used by companies and PCT patents became representative of global patenting patterns. As a consequence, the sample size was restricted to the year following the signing of TRIPs agreements (1995).

The query resulted in the retrieval of 91796 unique applications related to AI technologies from 1995 to 2017 over a total of 3198252 patents filed using the PCT route. The methodology used for the selection of the patents involving AI technologies is based on the WIPO study (Benzell et al., 2019) and it is described in detail in Appendix B.[7]

Some comments on the quantitative analysis of patents Patents are an invaluable source of data for innovation studies, but since they were not designed for the econometric analysis of technology, rather to act as property assets, the information they provide must be treated with caution. First of all, patents do not cover all aspects of a technology. When an innovation does not fulfill patenting requirements, it cannot be detected by solely using patents. This is often the case for AI technologies, which greatly rely on algorithms and mathematical methods, which private-led innovators may decide to protect through means of trade secrets. If there is a public record of these innovations, it is likely to be present in academic papers, and therefore an analysis exclusively focused on patents would not include this information. Nevertheless, the *functional application* of AI technologies can be patented, informing on how AI technologies are spreading to a higher degree of abstraction.

[7]The complete code used during the analysis can be found on the MyThesis repository on Github `https://github.com/AlessioNar/MyThesis` (Nardin, 2021).

Another factor to take into account is that the process of obtaining a patent is designed to be particularly long and expensive, thus increasing the barrier of entry for protection. As a consequence, only the technologies and techniques that are more likely to be a source of consistent profits are going to be patented.

Moreover, the heterogeneous and ever-changing nature of AI technologies may complicate the process of patent identification. In chapter 1 are provided some qualitative definitions of what can be considered as Artificial Intelligence and a short history of how the concept evolved over time. For the purposes of this analysis I decided to use the widest definition of AI possible, thus including technologies that a part of the literature does not considered AI anymore, such as expert systems.

Information regarding specific technologies in patent data can be found by analyzing technological classification codes. Classification codes are used by patent examiners to classify patent applications and other documents according to their technical features, to facilitate search and examination. Technology classes are generally classified hierarchically, giving the possibility to cut the tree at different levels of abstraction. For the purpose of this analysis I used the 4-digit classification level. Several patent classification systems are used by the various patent offices. In the analysis I mainly used the Cooperative Patent Classification (CPC), jointly developed by the United States Patent and Trademark Office (USPTO) and the European Patent Office (EPO), while in some cases I also used the International Patent Classification (IPC), maintained by WIPO, both of which contains more than 100000 codes.

To identify the evolution of technology, it is common practice to use patent citations. Patent citations have an important legal function, because, as presented in chapter 2, they delimit the scope of the property rights awarded by the patent. From this it is generally inferred that, if patent B cites patent A, patent B builds on the knowledge of patent A, over which patent B cannot have a claim. Patent citations are in this case representative of the links between patented innovations. To conduce a quantitative analysis on patents, citation data is relevant because, as Hall and Trajtenberg (2004) affirms, *"citations [made] may constitute a 'paper trail' for spillovers"*, in other words, we assume a knowledge transfer from the cited to the citing patent. In this framework, citations offer insights regarding the evolution of a particular line of technology and whether a particular invention is used in a wider variety of applications. However, the analysis of citations, especially when dealing with recent data, is limited by truncation: the more recent the patent examined, the less forward citations it will have, as the pool of potential citing patents progressively diminish.

4.3 Artificial Intelligence as a General Purpose Technology

Recalling the definition presented in chapter 3, a technology should be considered as a GPT if it follows these three requirements:

1. technological dynamism;

2. pervasiveness;

3. strong complementarity with existing or new technologies.

These three characteristics are complex and intertwined with each other. In particular, measuring both pervasiveness and technological complementarity require the use of multiple indicators, that will be developed in detail in the following sections.

4.3.1 Technological dynamism

Perhaps the easiest characteristic to examine is technological dynamism, since it is, at its core, a growth-based measure. As Hall and Trajtenberg (2004) and Petralia (2020) suggested, I analyzed the percentage of PCT patent applications from 1995 to 2017.

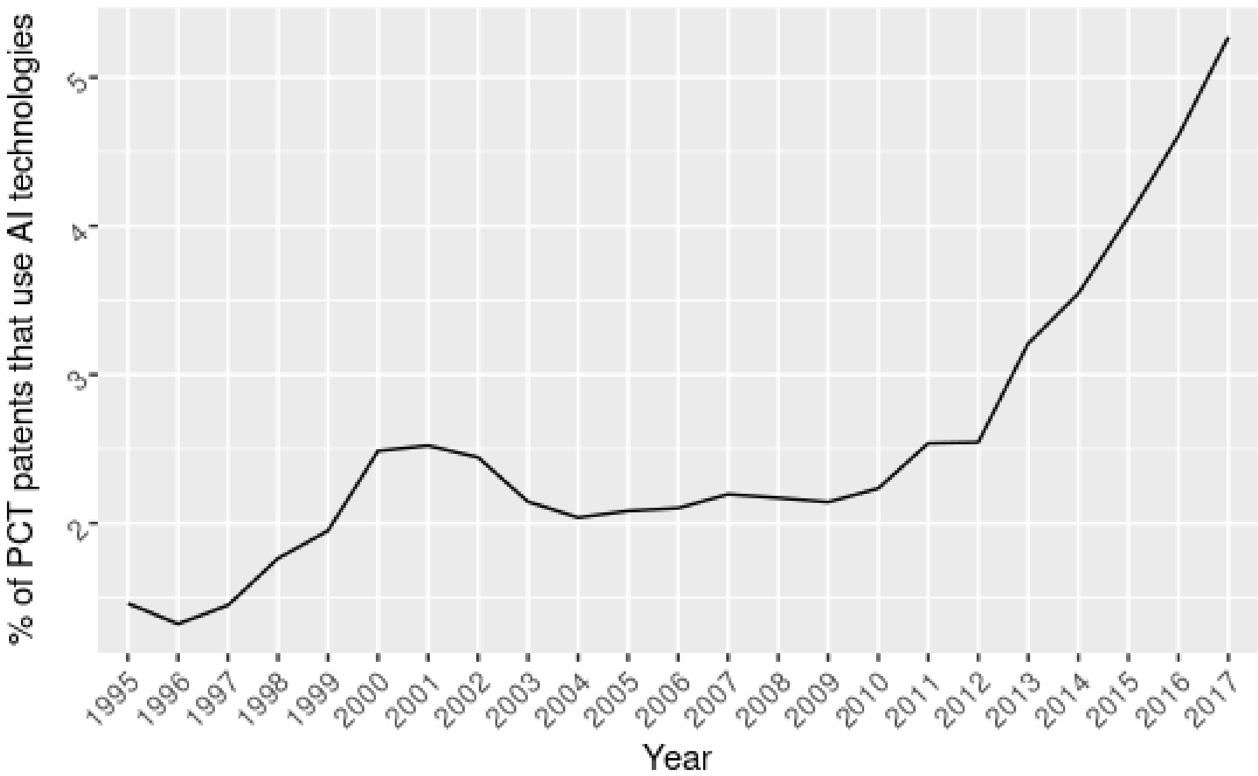


Figure 4.1: % of AI patents filed over the total number of PCT patents

Figure 4.1 shows the percentage of AI-related patents over the overall PCT patents, suggesting that, starting from 2013, we assist to a progressive increase in the percentage of AI-related patents filed following the PCT route every year. From more or less stable plateau comprised between 1.5 % and 2.5 % from 1995 and 2012, the percentage increases to 3.24% in 2013, up to reach more than 5% in 2017[8]. The timing of this surge in AI patenting is particularly significant. If we recall the historical evidence mentioned in chapter 1, 2012 was the year when the new generation of AI technologies (those related to Machine Learning and Deep Learning) took off commercially. Companies realized that AI as a prediction tool could be applied in a variety of different fields and increased their investment in R&D, subsequently filing an increased number of patents, possibly to prevent competitors to claim exclusive rights over specific applications. AI technologies are indeed experiencing a rapid increase in patenting activity, that become more marked from 2012 onwards, up to become the 5.3% of

[8]In appendix D it is possible to find the values used to draw these graphs, respectively in tables D.1 and D.2

the overall PCT patents in 2017.

4.3.2 Technological pervasiveness and innovation complementarities

Determining the other two features of GPTs (pervasiveness and technological complementarity) using patent data is more complex. First of all, distinguishing the two phenomena is not trivial. Since patented technologies are not composed by general upstream principles, but rather intermediate goods or final products, likely what can be measured is the degree of innovation complementarities caused by GPTs. Nonetheless, basic indicators of pervasiveness may be created by analyzing the percentage of technological fields in which a related technology can be found. Toole et al. (2020) used this strategy to find evidence of how much AI is diffused in the economy. They considered the percentage of technological classes assigned to at least two patents over the total number of available technological classes, broken down by years. In figure 4.2 the evolution of this indicator is shown using both the IPC and CPC technological classification systems at the 4-digit level.

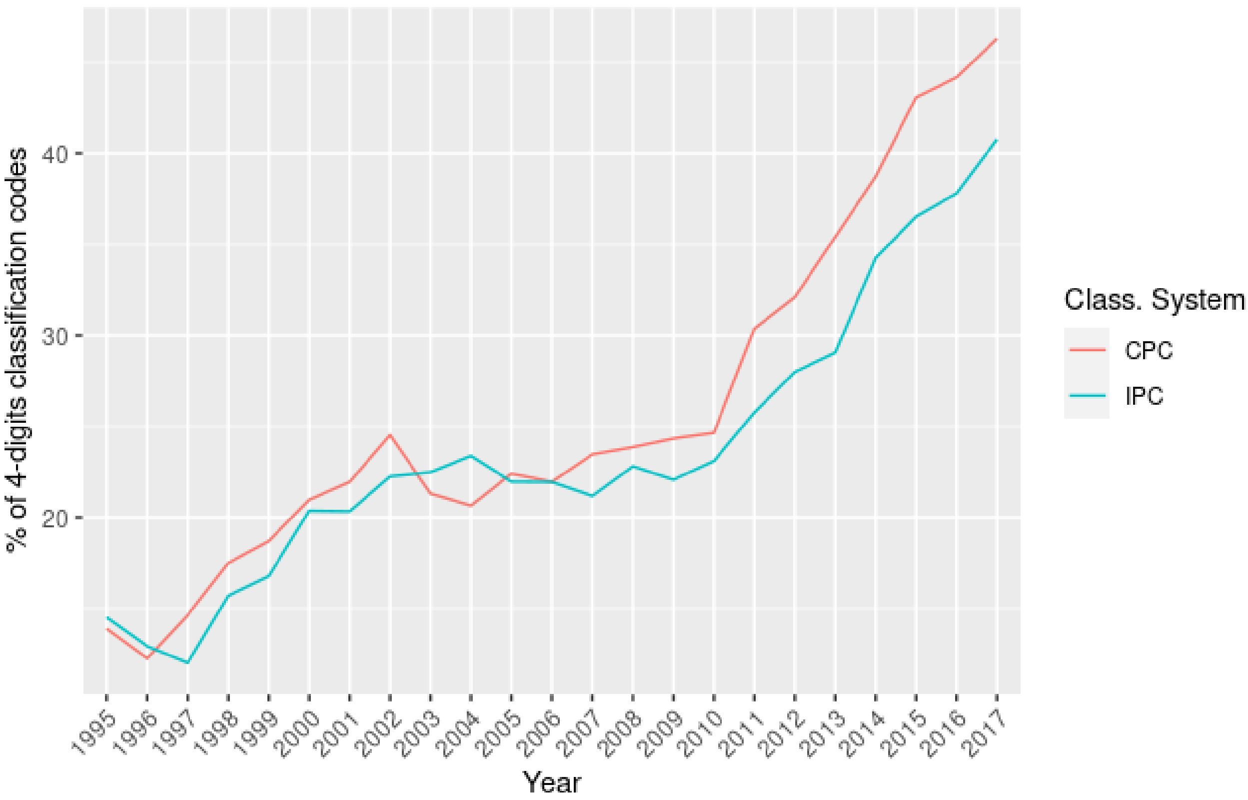


Figure 4.2: Percentage of 4-digits CPC and IPC classification codes present in AI-related patents

As it can be seen in tables D.3 and D.4, the number of technological classes of AI-related patents increased consistently in both classification systems and so their proportion, up to reach almost half (0.45) of all of the CPC symbols and well above one third of all of the IPC symbols in 2017. Again, it is interesting to note that, after a plateau in the 2000s, we assist a consistent and stable increase in the number of technological classification codes used from the early 2010s up until the most recent data available. This constant increase suggests that AI-related technologies are spreading across the economy, involving a variety of final products.

The second patent-based indicator commonly used to identify GPT is the generality index, that aims to determine the degree of innovation complementarities of the technology. Hall and Trajtenberg (2004) suggest that patents that involve GPTs should have many citations from outside their specific technology area while having many citations within their technology, caused by patterns of cumulative innovation. However, given that GPTs take more time to pervade the economy, citation lags for patents in this area may be longer than average, and thus the burst of innovative activity caused by complementary innovation in application sectors may take longer. In a seminal article published in 1993, Jaffe et al. introduced the generality index, an indicator based on the Herfindahl concentration index, that investigates the range of technology fields (and industries) that cite the original patent.

Given a patent i, the generality index is defined as:

$$G_i = 1 - \sum_{k=1}^{N_i} {s_{ij}}^2 \tag{4.1}$$

where s_{ij} is the percentage of citations received by patent i that belong to the technology class j, out of J technology classes. s_{ij} is computed as

$$s_{ij} = \frac{NClass_{ij}}{NClass_J} \tag{4.2}$$

where $NClass_{ij}$ is the number of citations the patent i received from a patent having technology class j, and $NClass_J$ represents the total number of technological classes associated with the citing patents. So:

$$G_i = 1 - \sum_{k=1}^{N_i} \left(\frac{NClass_{ij}}{NClass_J} \right)^2 \tag{4.3}$$

Notice that $0 \leq G_i \leq 1$. Higher values represent a high variability in the technology classes of the citing patents and hence an increased generality.

Since the generality index is a form of the Herfindahl concentration index, it presents the same disadvantages. As Hall (2005) affirmed, this kind of indexes is in general biased downwards with a larger effect for small $NClass$. Thus, Hall (2005) suggests to correct the bias by computing an unbiased estimator:

$$\tilde{G}_i = \frac{NClass_J}{NClass_J - 1} G_i \tag{4.4}$$

Note that the bias correction is valid only for small N and, adopting Hall and Trajtenberg (2004) methodology, it is computed only for $NClass_J \geq 2$, to avoid infinite and null values. As a consequence, the formula for computing Generality of a patent i can be described as:

$$\tilde{G}_i = \frac{NClass_J}{NClass_J - 1}\left(1 - \sum_{k=1}^{N_i}\left(\frac{NClass_{ij}}{NClass_J}\right)^2\right) \quad \text{if} \quad NClass_J \geq 2 \tag{4.5}$$

$$G_i = 1 - \sum_{k=1}^{N_i}\left(\frac{NClass_{ij}}{NClass_J}\right)^2 \quad \text{if} \quad NClass_J < 2 \tag{4.6}$$

Moreover, as suggested by Hall and Trajtenberg (2004), the generality was not defined for patents with no citations and that for patents receiving only one forward citation, $G_i = 0$.

For the purpose of this thesis, the generality index was applied to the subset of AI-related patents, using the Cooperative Patent Classification system, the Hall-Jaffe-Trajtenberg 36 technology subcategories and the NACE industry classification v. 2. This choice was motivated by the fact that, recalling the definition of GPT, it is fundamental to analyze the generality of technologies developed by firms in different industries. Thus Hall and Trajtenberg (2004) suggest computing the generality index based on the industry, by associating an technological classes with industrial sectors using a an industry-patent class concordance.[9]

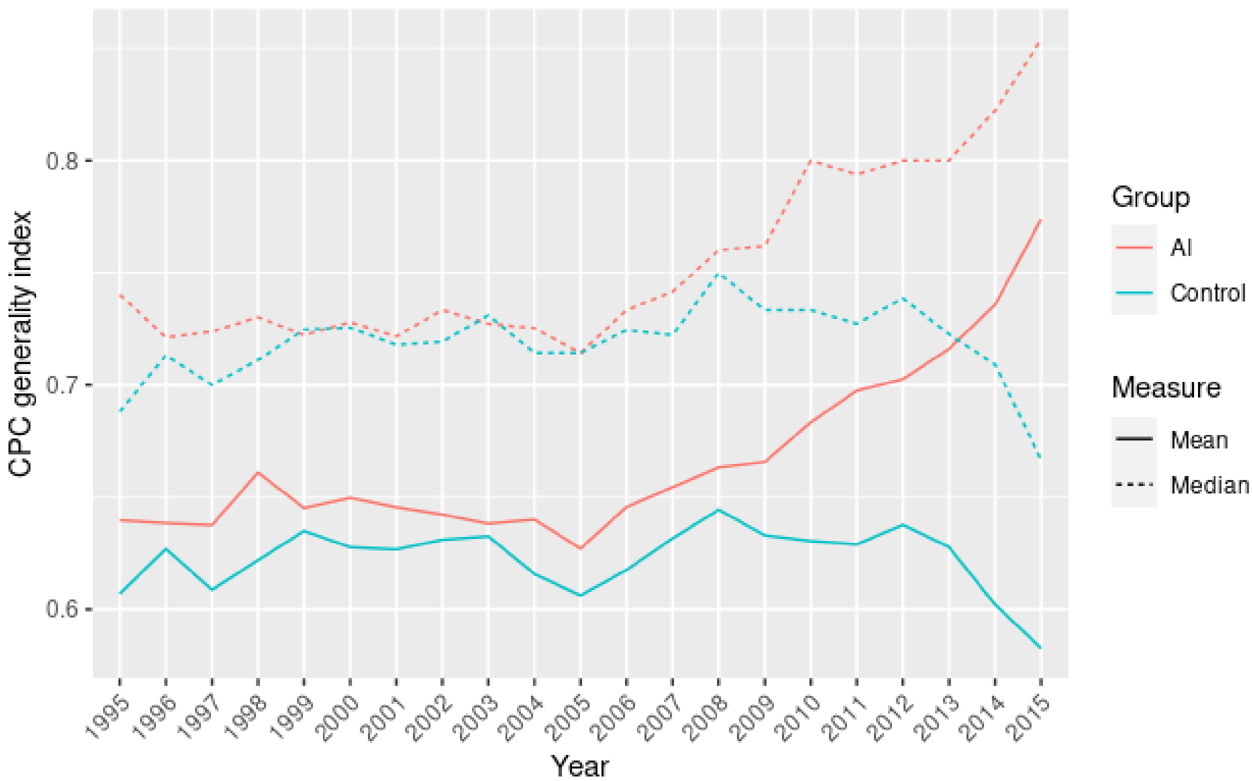


Figure 4.3: CPC generality index over time

In figure 4.3 the generality index is computed using the CPC technology classes over time of both the AI-related patents and a control sample of patents filed the same year and with the same number of forward citations[10]. It can be observed that, during the period took into consideration, AI-related patents have a higher generality. However, starting from 2008, they start to diverge, with the generality of AI-related patents

[9]In figures 4.3, 4.4, 4.5 and 4.6 the x axis is truncated in 2015. As it was previously mentioned, this choice is motivated by the fact that recent patents have a low number of forward citations, thus invalidating the results of the generality computation for the years 2016 and 2017.

[10]Details on the construction of the control sample are provided in appendix B

increasing up to reach a maximum mean value of 0.77 and the control sample decreasing. This may be caused by the fact that AI-related patents filed after 2008 are being cited by other patents containing CPC symbols of a higher variety of different technological fields, while the technological variety of the forward citations of the control sample becomes lower and lower, indicating a restricted scope of application. This is consistent with both the historical evidence that regards the renewed commercial discovery of AI technologies and previous findings related to the expansion of the fraction of IPC and CPC classification symbols (figure 4.2).

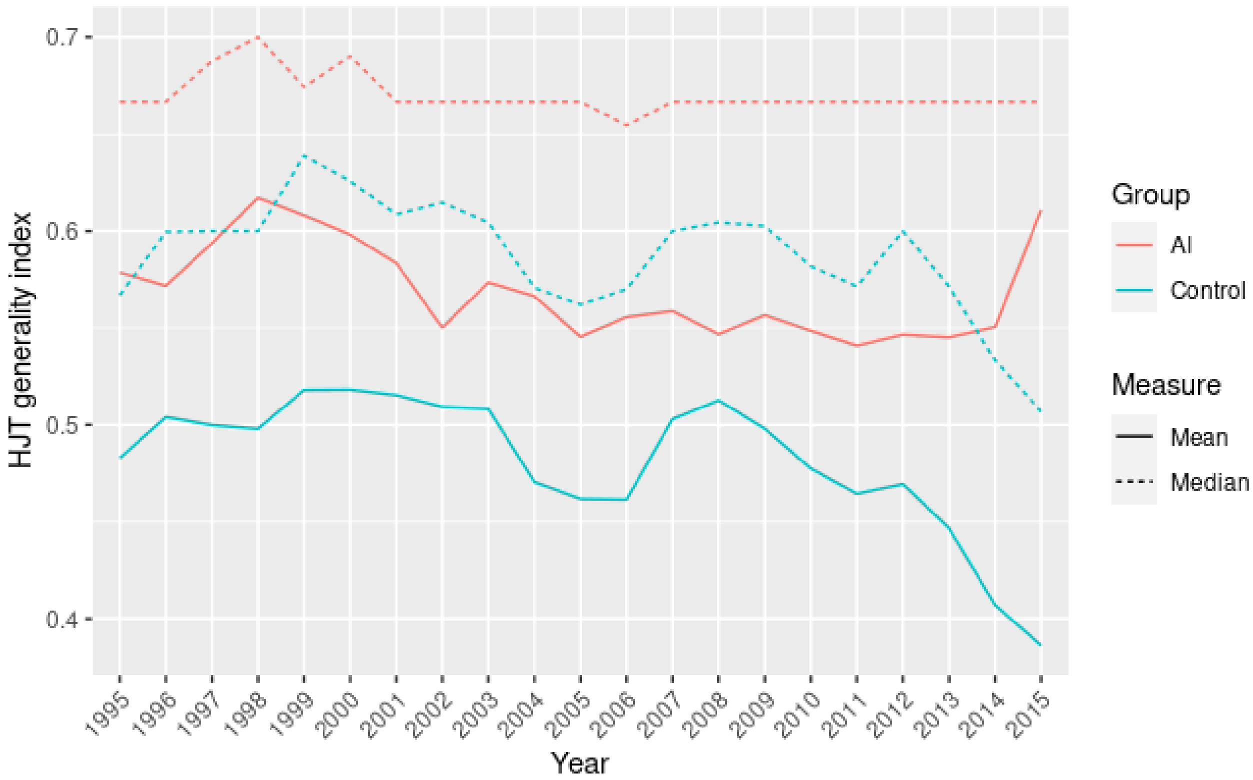


Figure 4.4: Hall-Jaffe-Trajtenberg generality index over time

To a lower extent, the same trend can be seen in figure 4.4, that shows the generality index computed using the Hall-Jaffe-Trajtenberg technology subcategories. Considering that the number of possible categorization is limited to 36, it is particularly significant that AI-related patents rank consistently higher and that there is a divergence in more recent years.

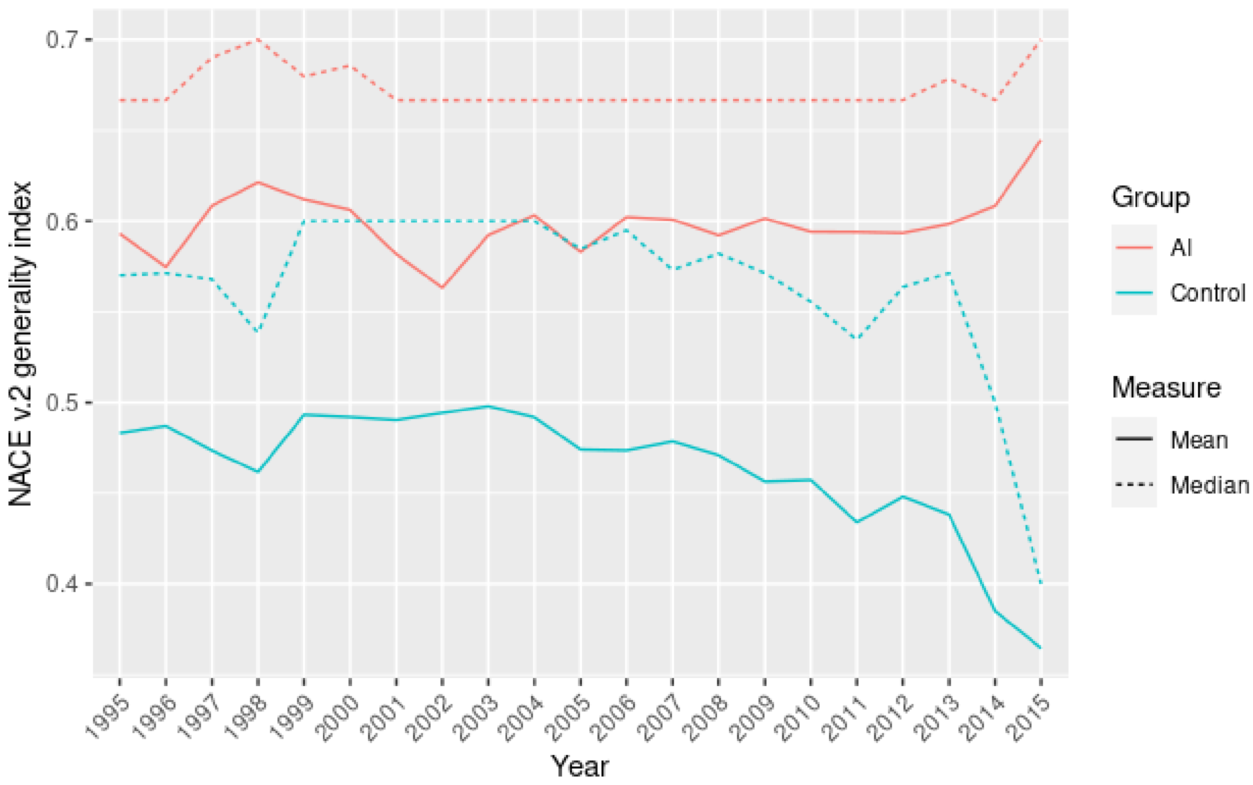


Figure 4.5: NACE v.2 generality index over time

Finally, in figure 4.5 is showed the generality index computed based on industrial categorization. Again, AI-related patents score higher than the control sample, with a hint of diverging behavior between 2013 and 2015, while in the end we assist to the usual truncation phenomenon, and thus a decrease in average generality.

Overall, we can affirm that AI-related patents have a high generality index, suggesting that, at least, AI technologies are applied in products that span in a variety of different technological and industrial areas. We also observe that, when generality is computed for patents filed in recent years, the average generality index decreases. As it was previously mentioned, this a known issue of forward-based patent measures and it is commonly referred to as the truncation problem. As the patent examined are closer to the present, the number of cumulative inventions decreases and thus we assist a reduction in forward citations. Since patents with only one forward citation were assigned $G_i = 0$ by default, with the increase of the number of such patents, it is normal to assist a reduction in generality. To compensate for this phenomenon, in the graphs and tables of this analysis patents with only one forward citation were removed.[11].

The results of the generality index computations are coherent for all metrics, suggesting that there is a correlation between the generality index in technological classification system and application industry. The CPC-based generality index has a median value comprised between 0.85 and 0.72 (figure 4.3), the Hall-Jaffe-Trajtenberg generality index has a median value comprised between 0.70 and 0.65 (figure 4.4, and the NACE v.2 generality index has a median value between 0.71 and 0.67 (figure 4.5).

[11] Along the lines of previous work by Hall and Trajtenberg (2004)

Figure 4.6: Standard deviation of generality index over time

The yearly variability of the patents generality index is particularly high for all indicators, ranging from 0.25 to 0.35, suggesting that there may be a polarization between generality indexes, with the values that are usually far from the mean (figure 4.6). Concluding, we can safely affirm that, based on these well-established indicators, AI technologies have an high degree of pervasiveness and technology complementarity.

4.3.3 Network-based indicators

Other scholars, such as Petralia (2020), suggested measuring technological pervasiveness using the adjacency matrix of technological classes. This method was also used by Cecere et al. (2014) and Tang et al. (2020) to map the evolution of a technology over time allowing the application of network analysis tools. In this section, I will apply similar techniques to find evidence of pervasiveness and technological complementarity, with the objective of posing the basis for a structural analysis of GPTs.

Following the intuition of Tang et al. (2020), I built a technology network of AI-related patents by treating technological classes and patents as the nodes in a bipartite graph. A patent can be assigned one or more technological classes. Suppose of having a sample of three patents: Patent 1 was assigned the technological classes CPC1, CPC2, and CPC3, patent 2 was assigned the technological classes CPC2 and CPC4, and patent 3 was assigned the technological classes CPC1 and CPC2. The relationships between patents and technological classes can be treated as an edge list for a bipartite graph, where patents are nodes of type 1 and technological classes are nodes of type 2 (figure 4.7). By projecting the nodes of type 2 (the technological classes) we obtain a weighted undirected technology network, where the weight of the edges represents the number of patents where those particular two technologies are co-occurring (figure 4.8 represents the projection of the technology network, while table 4.1 represents the adjacency matrix of the projection).

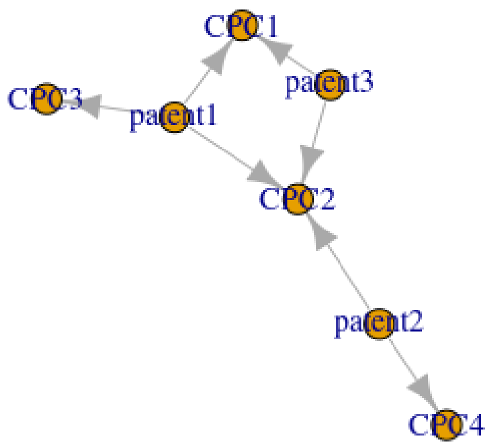


Figure 4.7: Bipartite graph representing the relations between patents and their respective technological classes

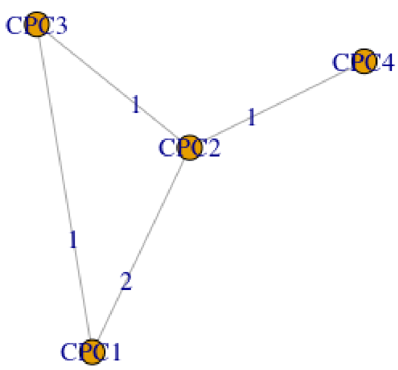


Figure 4.8: Technology projection representing the technological network formed by patents in figure 4.7

Table 4.1: Adjacency matrix of a technology network

	CPC1	CPC2	CPC3	CPC4
CPC1	-	2	1	0
CPC2	2	-	1	1
CPC3	1	1	-	0
CPC4	0	1	1	-

As it is possible to observe in figure 4.7 and 4.8, since two patents are connected to CPC1 and CPC2, then the weight of the edge linking CPC1 and CPC2 is assigned a value of 2 in the projection, while all the other edges have a weight equals to 1.

After having mapped the sample of AI patents, I examined the evolution of the technology-based projection constructed using CPC technological classes using network science techniques.

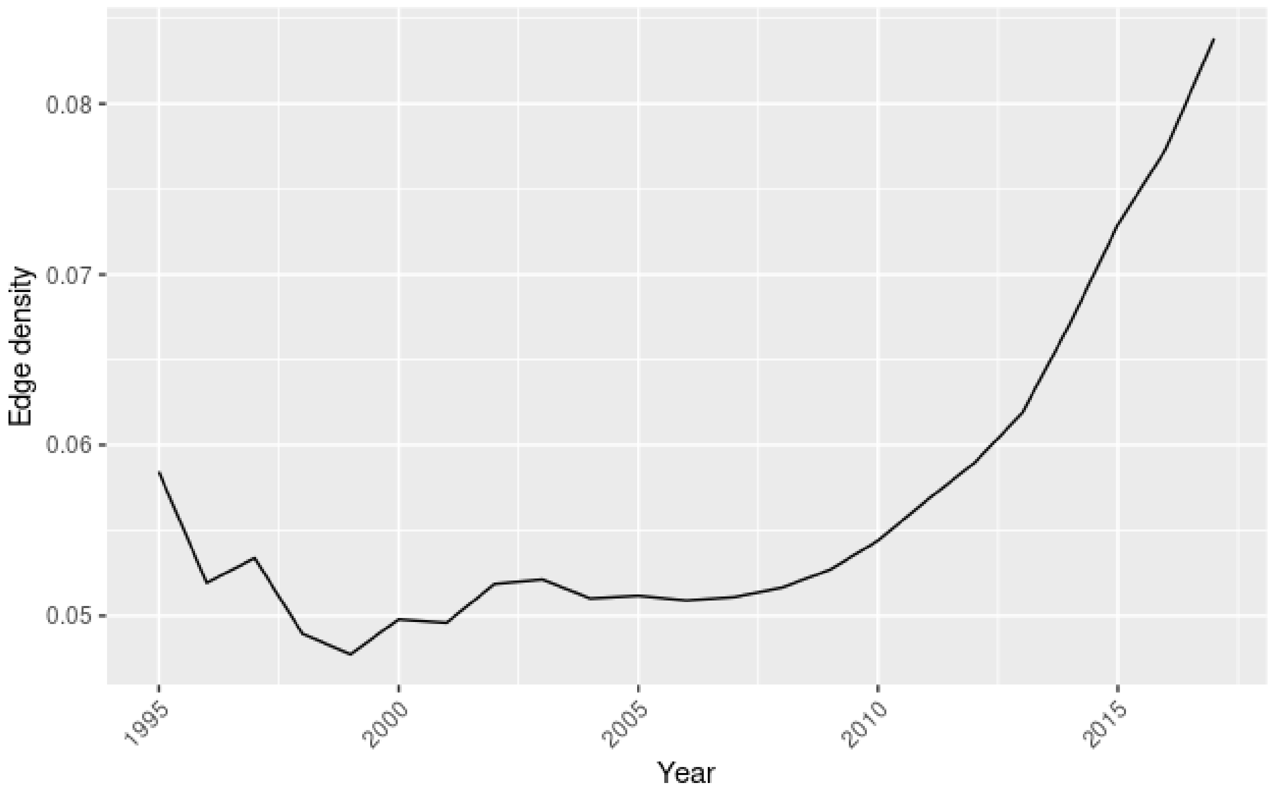


Figure 4.9: Edge density

In figure 4.9 is shown the edge density of the technology network. The edge density represents the ratio of the number of edges and the number of possible edges (Wasserman et al., 1994). In the context of a technology network it represents the ratio between the existing links between technologies and their possible links. Obviously, not all links between technologies are possible, as it was mentioned in chapter 3. However, GPTs should be able to create new non-existing links. In the AI technology network, after an initial decrease, it is possible to observe an upward trend, that shows AI technologies are decreasing the distance between different technological fields over time. Additionally, considering that the number of technological classifications increased along the years (and thus the number of possible links), the upward trend is particularly significant, since the number of links between different technological fields increased even when taking into consideration the increase in the number of possible connections.

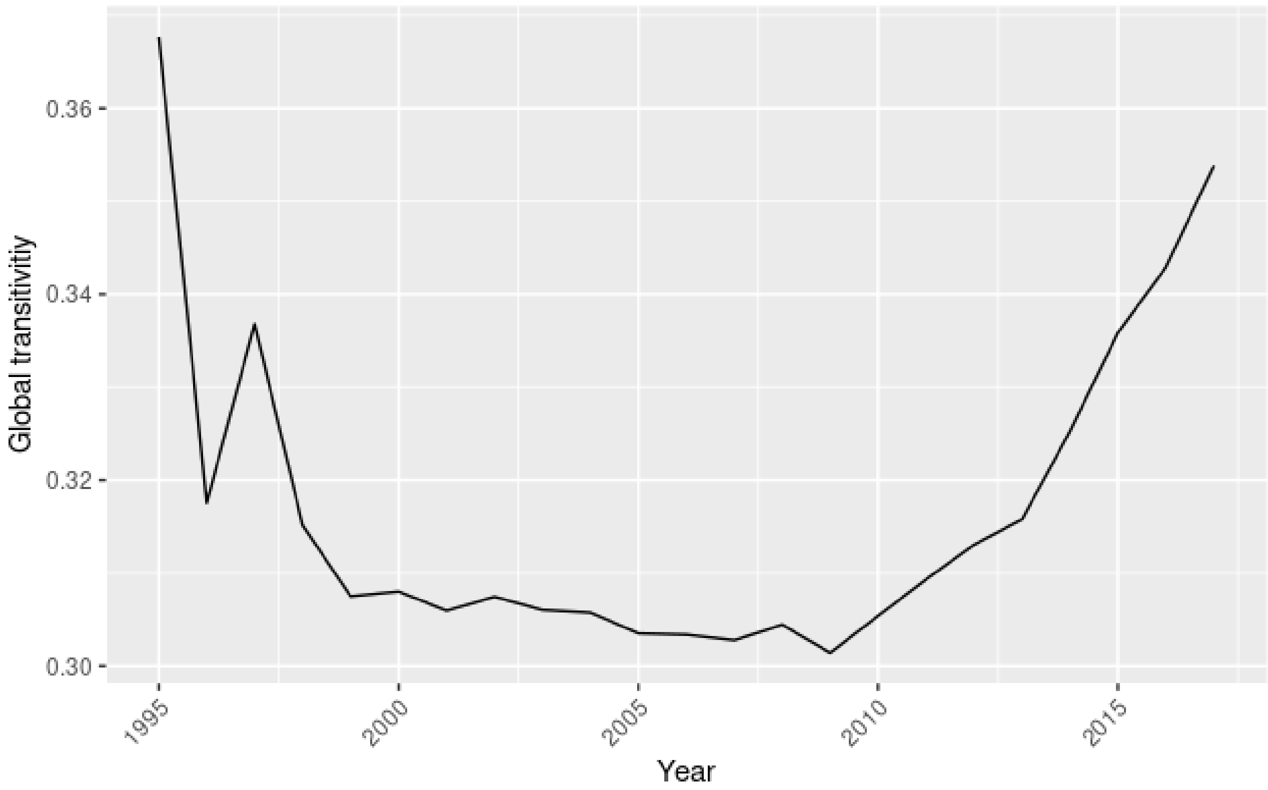


Figure 4.10: Global transitivity

Similarly, figure 4.10 shows the evolution of the global transitivity of the AI technology network. The global transitivity represents the ratio of closed triangles over the possible triangles of the network. While we initially assist to a marked decrease in transitivity, probably caused by the increase in the possible connection between technologies, global transitivity rises again starting from 2010, indicating that, after an initial introduction of new technological classes, the ties between technological classes began reinforcing. In the context of studying GPTs, edge density and global transitivity are particularly useful in determining the degree of pervasiveness of a technology. An increasing trend in edge density indicates that the technology is becoming more and more pervasive in the economy.

Centralization measures are also particularly helpful to measure pervasiveness. To compare different networks, Freeman (1978) introduced the concept of centralization measures, that aims to determine the extent to which centrality measures distributions (the positions of networks nodes in respect to the center) informs us on the structure of the network as a whole. For the purpose of this analysis, we are particularly interested in the evolution of the degree centralization and betweenness centralization, that are based on degree and betweenness centrality, respectively.

The degree of a node is determined by the number of nodes it is connected to. In weighted networks, it is normally generalized as the sum of the weights of the links between nodes, also referred to as node strength:

$$C_D^w(i) = \sum_{j}^{N} w_{ij} \tag{4.7}$$

where w_{ij} is the weight between node i and node j.

After having computed this measure for all the nodes, it can be generalized to compare graphs of different sizes, to obtain what Freeman (1978) has defined as degree centralization.

$$C_D^w(N) = \frac{\sum (max(C_D^w) - C_D^w(i))}{(N_{nodes} - 1)(N_{nodes} - 2)max(w)} \tag{4.8}$$

where $max(C_D^w)$ is the maximum weighted degree centrality of the network. Note that this measure is comprised between 0 and 1, where 0 indicates a low centralization and 1 indicates a high centralization.

Degree centralization indicates whether we assist a dominance of a particular group of nodes within the network in terms of weighted degree. Degree centralization informs us on whether there is a small number of nodes that are predominant and hold the majority of the connections with other nodes (in case of a high degree centralization) or instead the weighted degree is distributed evenly in the network (in case of a low degree centralization). In the context of the technology network, a high degree centralization suggest that a small number of technological classes are predominant in the network, thus being central for the technology took into consideration

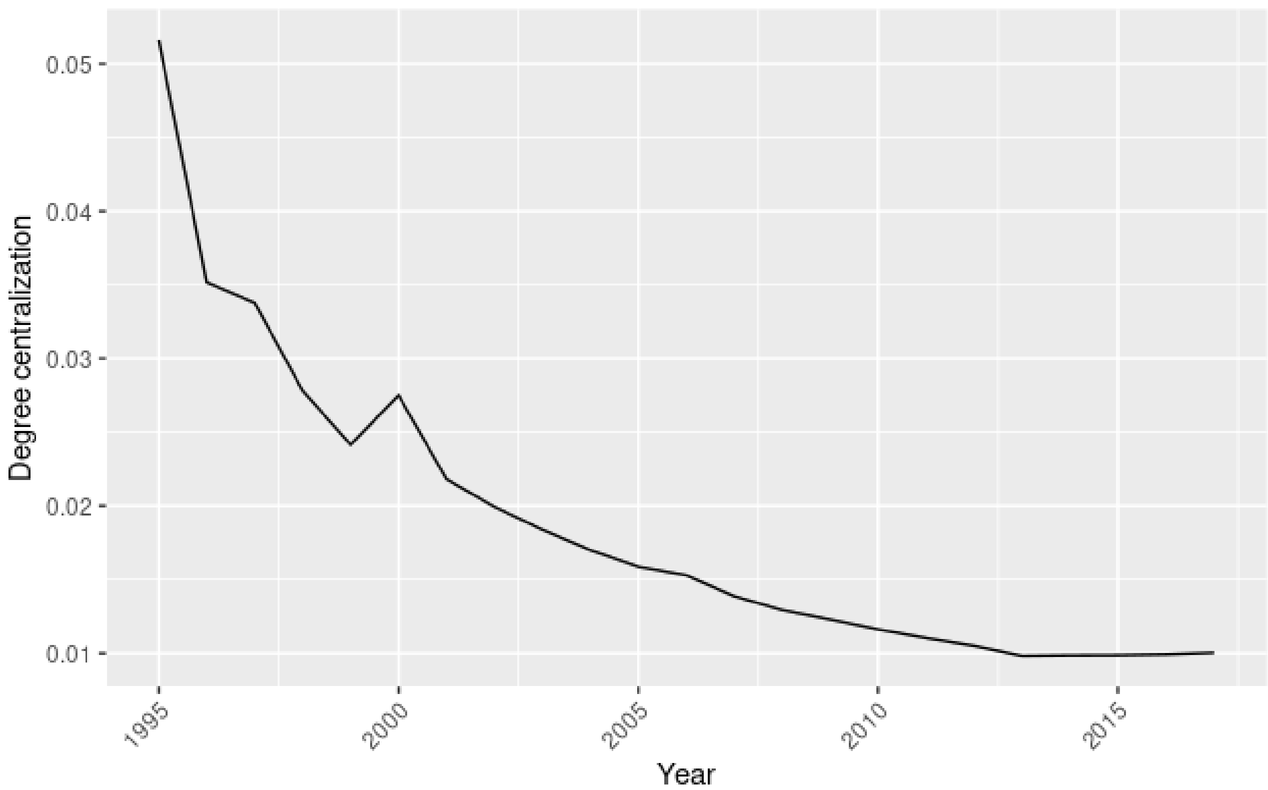


Figure 4.11: Evolution of degree centralization in the AI technology network

In figure 4.11 we observe that degree centralization is very low, with a maximum value of 0.05, and that with the passing of the years, it decreases up to become less than 0.01. This suggests that the technological classes found in AI-related patents are distributed evenly in the network, with a trend of equalizing the distribution of patents among the years. Degree centralization helps determining whether the distribution of edges and nodes weights is pervading the economy evenly or unevenly. In the case of AI technologies, we assist to a homogeneous process of technology expansion, suggesting that AI is distributing evenly in the various technological classes.

Another common indicator for studying networks is betweenness centrality. Betweenness is a centrality measure based on shortest paths and it is focused on determining the number of shortest paths between two nodes that pass through a specific node in the network.

$$C_B(i) = \frac{g_{jk}(i)}{g_{jk}} \tag{4.9}$$

where g_{jk} is the number of shortest paths between the nodes j and k and $g_{jk}(i)$ is the number of those shortest paths that pass through i. This centrality measures therefore aims to determine nodes in the network that are most likely to act as intermediaries between different nodes. In the framework of this analysis, the weight of the links is treated as a distance measure, thus, since a higher weight indicates that nodes are less *distant* to each other, to compute the betweenness centrality it is used its inverse ($\frac{1}{w}$). At the network level, betweenness centrality can be used to compute the betweenness centralization, that informs on the distribution of betweenness centrality among the nodes of the network. As suggested by Freeman (1978), betweenness centralization can be computed as:

$$C_B^w(N) = \frac{\sum max(C_B^w) - C_B^w(i)}{\frac{(N_{nodes}^2)-3(N_{nodes})+2}{2}(N_{nodes}-1) * max(w)} \tag{4.10}$$

note that this value is comprised between 0 and 1. Values of betweenness centralization closer to 1 indicate that the distribution of nodes is uneven, with a relatively low number of nodes that have a high betweenness centrality while the rest of the network has a low betweenness centrality. Values of betweenness centralization closer to 0 instead indicate that the betweenness centrality is distributed evenly among nodes.

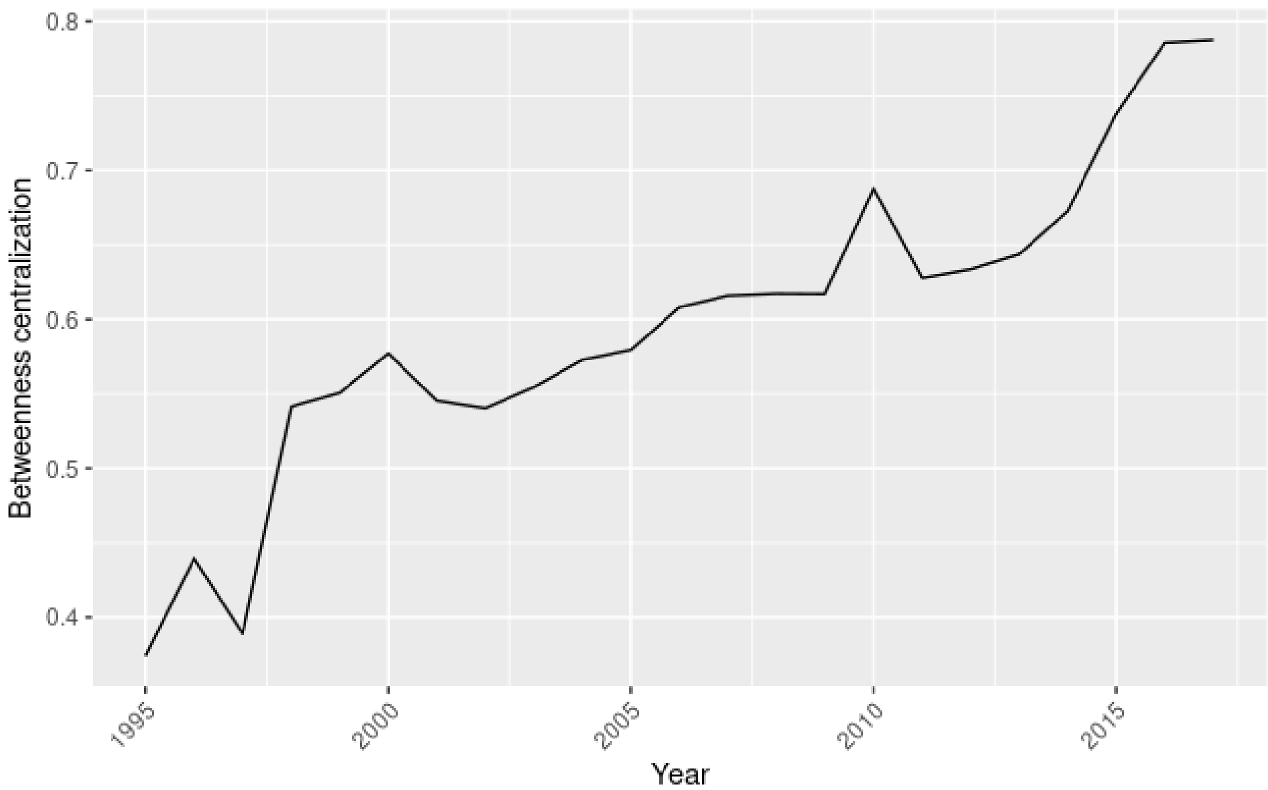


Figure 4.12: Evolution of betweenness centralization in the AI technology network

Figure 4.12 shows the evolution of the betweenness centralization of the AI technology network. We observe a progressive increase in betweenness centralization (that reaches a maximum of 0.78 in 2017), indicating that some technological classes are acting as bridges between nodes. From this can be derived that, along the years, some technological classes became essential in connecting nodes of different technological classes. For the purpose of determining whether a technology can be considered a GPT, betweenness centrality informs us on whether there is a group of essential technological classes that act as brokers in the diffusion in the economy. Technological with a high betweenness centrality are key elements for the diffusion of AI technologies, and they are likely to be the core technological classes needed for the implementation of AI in different application sectors.

Other than centralization, two measures are particularly interesting, since they explore technological complementarity: the number of new edges (so the number of new connections in the technology network) and the evolution of edge weights (how much existing edges are reinforced).

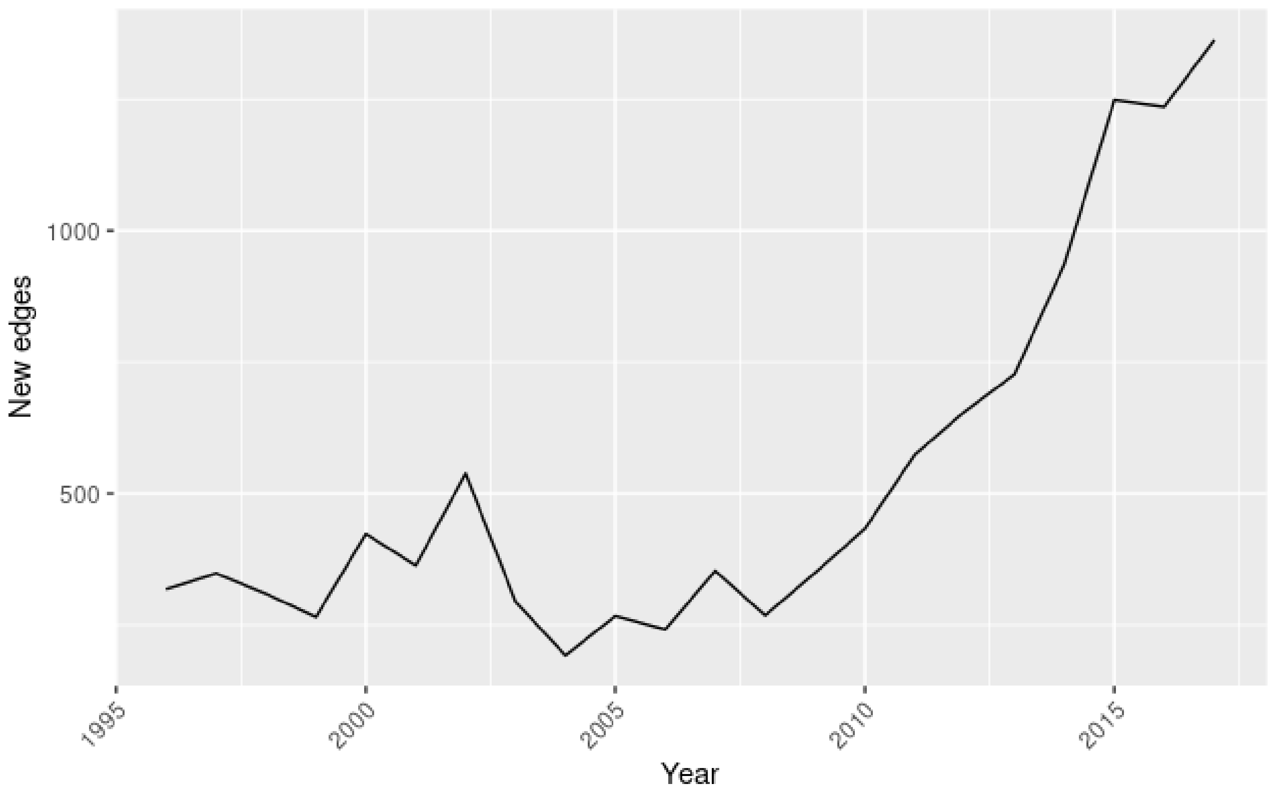


Figure 4.13: New edges in the network

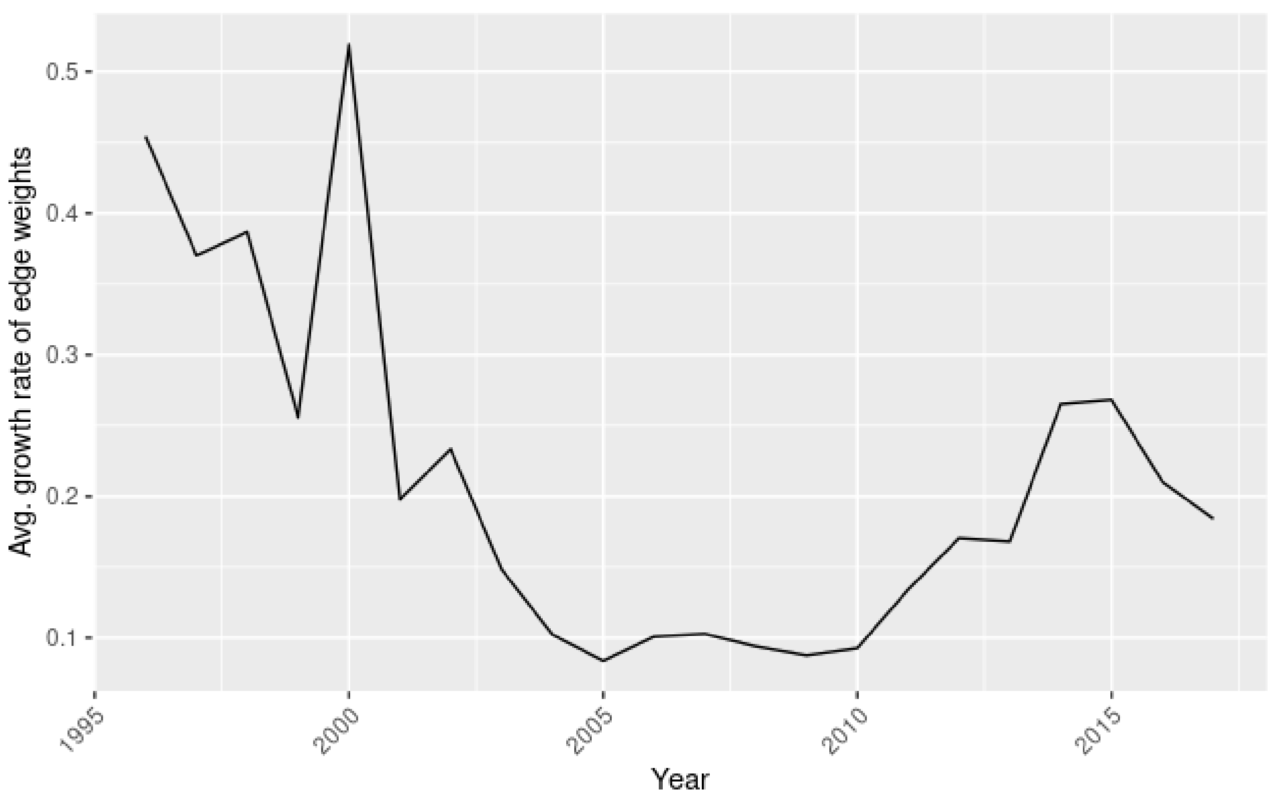


Figure 4.14: Average growth rate of the weights between existing edges

Recalling the theoretical framework proposed by Korzinov and Savin (2016), new edges can be modeled as new successful combinations of technologies. In figure 4.13 is represented the evolution of this indicator. We can observe that the number of entirely new edges increased consistently starting from 2006. As a consequence, it is possible to formulate the hypothesis that, starting from 2006, innovation complementarities derived from the introduction of AI technologies are increasing, often between different technological areas. If new edges are an important indicator to determine the inclusion of new technology recombinations in AI-related products, increases in edge weights are representative of the reinforcement of already existing links caused by the introduction of AI technologies. In figure 4.14 can be observed that, after a decrease in the average growth rate of edge weights, we are assisting to an increase in the reinforcement of already existing technology combinations.

4.3.4 Results

This analysis aimed to evaluate whether AI should be considered a General Purpose Technology. Recalling the definition of General Purpose Technology presented in chapter 3, AI should be considered as such if it respects three requirements:

1. Technological dynamism;
2. Wide variety of application fields;
3. Strong complementarities with existing and new technologies.

We found proofs of the technological dynamism of AI-related technologies in the indicators presented in section 4.3.1, where the proportion of AI-related patents following the PCT route is increasing. In section 4.3.2, traditional indicators such as the percentage of IPC and CPC technological classes used in AI-related patents showed that AI technologies are indeed spreading across society, while the more subtle generality index was used to determine the extent to which AI technologies are sparkling innovation in a wide range of fields, confirming that, within the brief time-period took into examination, AI-related patents maintained a very high generality index, even when compared with a control sample.

Finally, in section 4.3.3 original network-based indicators to identify a GPT were presented. In particular they were focused on determining the extent to which AI-related patents recombined different technological fields to create new and innovative technological combinations. Network-based indicators are particularly useful because they provide information not only on the role of each patent in the sample, but also information on the structure of the sample itself. Specifically, edge density and transitivity showed that there is an upward trend in the number of technology combinations involving by AI technologies, suggesting that AI technologies are rapidly expanding and occupying a greater space in the economy, further confirmed by the increases in the number of new edges and the growth rate of the weights of existing edges. Degree and betweenness centralization instead informed on the inner dynamics of the technology network, showing that the strength of the connections between different technologies is distributed evenly among the different nodes while a few nodes occupy a central position. When we look at the top nodes in betweenness centrality, we observe that it is mostly occupied by technological classes related to computing, and specifically those typical of AI.

This analysis leads to the conclusion that AI is a technology that is developing at an astounding rate and occupying larger fractions of the economy, and that it is complementary to existing and new technologies, thus confirming the hypothesis that should be considered a GPT.

4.4 A note on technological concentration

In chapter 3, I presented the claims made by Crémer et al. (2019) on the progressive concentration of the digital market. I also suggested that this may be favored by a a competitive advantage in innovation caused by data repurposing and know-how. This brief section aims to verify whether a concentration of technological capabilities in the patent portfolios of a relatively small number of patent applicants is occurring.

Using the AI-related patents retrieved, I analyzed the concentration of patents using applicants' names. Applicant names were subjected common preprocessing techniques to disambiguate patents formally filed under different names but that in the end belonged to the same entity [12]. During the process of data cleansing, I took into consideration that up until 2011 in the United States only by natural persons (US Congress, 2011), therefore many PCT patents filed at USPTO also indicated the inventors as applicants. In 2012 president Obama made the filing procedure possible also to legal entities. To minimize any methodological mistake, I eliminated applicants marked as individuals in patents including a legal entity. This has led me to obtaining a total of 28317 unique applicants.

From an initial look at the average number of patents for each applicant it is possible to observe that the distribution is highly skewed in favor to applicants that have filed only one patent. Thus, in order to differentiate strong patentees from weak patentees, the sample was divided in three asymmetrical classes based on the number of filed patents per year: top 1%, 90%-99%, and lower 89% [13].

Then, the yearly distribution of patents across these classes was plotted in figure 4.15. We can clearly observe that the number of AI-related patents filed by weak patentees is progressively declining in favor of the top %1 of patentees. This suggests that companies might have progressively increased their patenting rate in technologies involving AI.

Data retrieved from PATSTAT also contains information related to the legal status of applicants, specifically on whether they are companies, public institutions, hospitals, universities, or individuals. In 4.16 it is plotted the percentage of yearly patents filed by these different categories.

[12] This process is described in detail in appendix C

[13] In table D.5 it is possible to absolute number of applicants assigned to each class by year.

Figure 4.15: Yearly distribution of relative number of patents filed by group

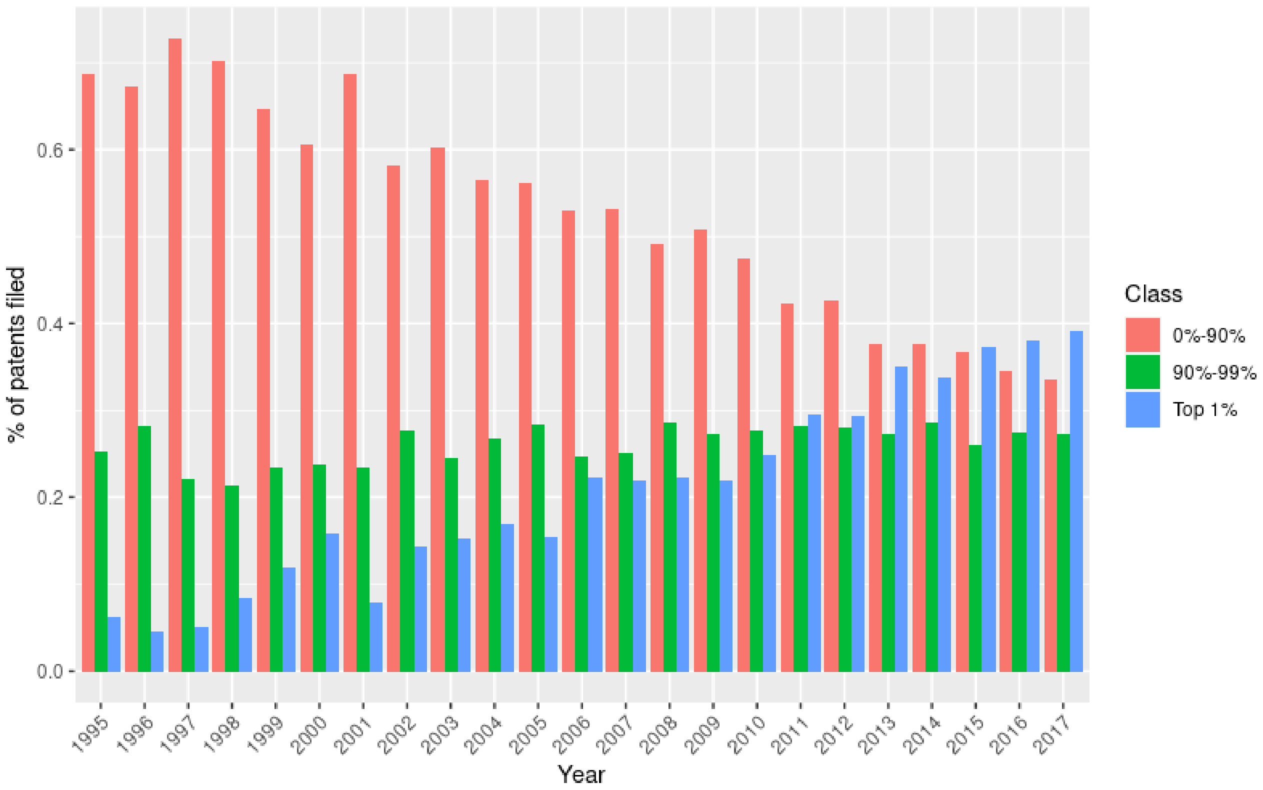
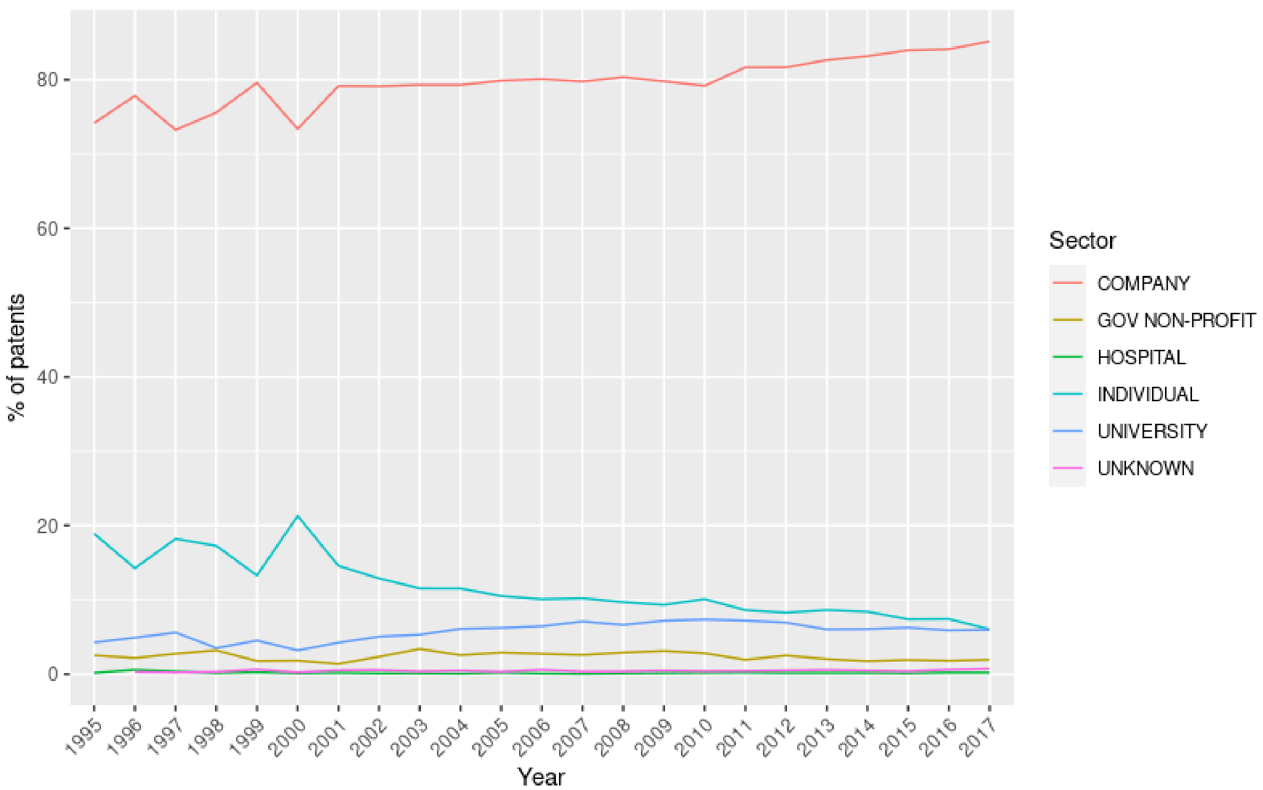


Figure 4.16: Yearly distribution of relative number of patents filed by sector



It can be observed that the majority of AI-related are filed by companies, followed by individuals and universities. During the period taken in examination, the percentage of patents filed by companies rose from 74.2% to 85.1%, while those filed by individuals diminished from being the 18.9% to 6.0%, and those filed by universities increased from 4.3% to 6.0%. This empirical evidence confirms the assumptions of Crémer et al. (2019), suggesting that today, the innovative push is mainly led by companies that are constantly increasing their innovative capabilities. From a merely innovation perspective this is not necessarily negative. In the past, big stable demanders were beneficial to the economy, since they provide a stable influx of investments in R&D that may avoid slowdowns in the rate of innovation caused by the negative horizontal externality of innovation complementarities mentioned in section 3. At the same time, under the lenses of competition policy the scenario becomes more worrying. In fact, differently from the past[14], this push is coming from private entities that can arbitrarily stop providing the fruits of their knowledge at any given moment. As a consequence, many argue in favor of a greater push towards public financing of AI technologies, especially within the European Union, that while being the location of many excellent research institutions producing high impact discoveries, lags behind in the private sector (Benzell et al., 2019), with no single firm capable to compete with American and Chinese digital giants, having no access to users' data, a fundamental ingredient for increasing competitiveness in the AI field. Recalling the ways in which AI technologies spread in the market (competition for the market), in the absence of policies that favor the development of strong European-based private firms active in the AI field, the gap is only destined to increase.

4.5 Conclusions

The evidence presented in this chapter aimed to confirm the theoretical propositions presented in chapter 3, regarding the fact that AI should be considered as a GPT. Both traditional indicators retrieved in the GPT literature (such as the generality index and the percentage of technological classes) and new network-based methods seems to point towards the conclusion that AI is indeed a GPT and that it is increasing its degree of pervasiveness in the economy at an impressive rate. Moreover, we are assisting towards a progressive concentration of AI-related patents in the portfolios of a proportionally small number of patent applicants. This confirms the theoretical basis provided by Crémer et al. (2019), that calls for regulatory intervention by antitrust authorities.

However, since this study is based on patent data, it also has some of its limitations. First, as it was already mentioned, when dealing with forward-looking measures with recent data, we inevitably face issues related to truncation. While patents are invaluable sources of information regarding a developing a technology, when we use forward citations indicators to measure contemporary cutting-edge technologies, simply not enough time has passed to find evidence of their impact on the technological scenario. Second, patent documents were not designed to measure the process of technology diffusion, but rather as a legal tool to limit the claims of the citing patent. Therefore, even if it reasonable to consider patent citations as a proxy for technological

[14]Such as the invention and development of the internet, a project that involved the collaboration between the military and public research entities. See Abbate (1999) for reference.

diffusion, it may as well be the case that the inventor of a citing patent invented a product without knowledge of the cited invention. On the other hand, citations inform us on how a technology evolved and can help us to identify trajectories and focal inventions, especially when a patent is highly cited. Third, the use of the technology network as a proxy to measure innovation complementarities is still at an early stage. As a consequence, the results coming from this technique needs to be treated with extreme caution, especially when the technological sample regards such a particular technology as AI. Regardless, it is significant that the evolution of the technology network confirms the expectations provided by more traditional indicators, such as the generality index.

Conclusion

Artificial Intelligence technologies are one of the critical elements of change in the short-medium term. The advent of a new GPT (which could very well be considered GMI[15]) will deeply change the dynamics of the economy, politics, society, and international relations. The multidimensional character of technology influences the power dynamics between those who will be able to exploit the altered scenario, and those who will not.

However, it would be fundamentally wrong to adopt a perspective in which we are merely victims of this transition. While this process of change cannot be stopped, individual and collective actors can guide and shape its direction, by establishing norms and institutions that minimize the damage and maximize the benefits. While science and technology can be used to pursue political goals, in their essence, their basic principles are politically agnostic, and only through awareness of the scope of the impact of these new technologies policymakers can aim to maximize the social welfare of the AI revolution.

Unfortunately, today, one can observe an increase in the competition between the major countries in the AI industry: the US and China (Benzell et al., 2019), which have launched in a technological race. While this could be expected, since dominance in AI technologies is strictly linked to economic and political supremacy, blind competition is particularly problematic since there is the concrete risk that it would bring to the overseeing of keystone aspects of AI development, such as its ethical use. As it was presented in chapter 1, although AI systems represents a fundamental improvement in prediction technologies, they cannot be considered intelligent in the same way humans are, and they are marked by flaws and dangers that are very difficult to forecast or detect. They can safely be used only in situations where the cost of error is minimal, while for more complex situations they need to be under human oversight. With the decrease of prediction cost, the value of judgment increases, especially in complex situations, where the negative consequences of AI mistakes increase, and human intervention is needed both for security and monitoring purposes.

In chapter 2, some of the main normative constraints that shape the development of new technologies were presented: Intellectual Property Rights. The first part evaluated the main economic arguments put forward to justify their existence and concluded that the only valid descend from an utilitarian perspective, which does not presume the introduction of temporary monopoly rights without taking into account the negative consequences they generate into the economy. Then, the three IPRs most commonly applied to AI technologies were presented, with a specific focus on the economics of patents. Finally, economic analysis was applied to AI technologies,

[15] A technology widely applicable in a variety of fields, with strong innovation complementarities and that can potentially be used during the R&D process, shifting entire scientific paradigms and drastically increasing the innovation rate.

with the objective to verify whether the claims made by legal scholars on the need to introduce additional IPRs applicable to AI are justified or not. However, I found that current evidence does not justify this approach, since no additional incentives are needed to promote innovation in AI, and since there is a strong tendency towards self-regulation using patentleft or copyleft models, this calls towards a reduction in IPRs, rather than the opposite.

Chapter 3 focused on the effects of AI technologies on innovation and the market. The first part centered on some technology categorization, namely General Purpose Technologies, Inventions of a Method of Inventing, and General Methods of Invention. Particular attention was given to GPTs, since they represent the focus of chapter 4 and there is a rich literature on their effects on the economy. A crucial point regards the phenomenon of innovation complementarities, which describes why GPTs represent a game-changer in terms of economic growth. Managing the virtuous cycles of innovation by minimizing the negative impact on employment will represent a crucial challenge for XXI century policymakers. The second part of the chapter focused on determining whether AI technologies possess some characteristics of these technological classes. Using a qualitative perspective, it is safe to affirm that AI presents features of both GPTs and IMIs, a combination that the literature could only find in digital computing. AI is configuring itself as a rapidly evolving, widely applicable, and complementary technology while at the same time capable of expanding the knowledge frontier of science and increasing the rate of scientific discovery. The AI revolution will invest a large fraction of economic sectors, incentivizing investments in AI research and development, with effects that will bounce from one sector to another. Although the invention of multi-function closed-loop Autonomous Discovery Systems is far from being a reality, AI technologies are likely to cause an unprecedented increase in productivity in academia, opening up entire new paradigms of science. Finally, the chapter concludes by presenting some of the consequences that AI is having on market structure. In particular, the presence of economies of scope is leading to a process of market concentrations between a small number of firms that compete with each other in innovation, conquering new markets and having higher and higher barriers of entry, represented either by extreme network effects or the exclusive possession of fundamental assets such as users' data. Considering that the AI market structure leaves little space for new entrants, the competition in innovation is likely to increase, especially now that the COVID-19 pandemic has prompted world economy in a deep economic crisis, opening opportunities for a redistribution of global economic power. At the same time, state and regional actors are trying to manage the digitization of the economy, with the European Union at the forefront. In particular, two pieces of legislation are in the process of being evaluated by the European Parliament: the Digital Market Act (DMA) and the Digital Services Act (DSA), while a third piece of legislation specifically targeted to AI technologies is expected in the first quarter of 2021. The DMA aims to reduce the market power of gatekeepers, multi-products platforms that have extreme prescriptive and indiscriminate power on what it can and cannot be done in their platforms, while at the same time potentially competing with actors using their platforms. The DSA is instead focused on establishing obligations and accountability for intermediaries' platforms that provide digital services, and enhance consumer protection, whether the services are located in or outside the EU. These sets of norms are characterized by having extraterritorial character, since their effectiveness is tied to the individual, and not to the legal site of the firm, and it is conforming to the trend of EU law launched by the GDPR. Differently from the imposition of additional IPRs, these sets of norms aim to improve competition and consumer protection, possibly minimizing the negative consequences of a widespread introduction of AI technologies.

Finally, chapter 4 objective's was to provide quantitative evidence to the claims of chapter 3: whether AI is a General Purpose technology, and whether we are assisting to a market concentration process of the AI market. The data used to perform the analysis was formed by a subset of PCT patents issued from 1995 to 2017. Most of the chapter focused on using traditional indicators, such as the generality index, and developing new indexes based on the evolution of the technology network formed by AI-related patents. The empirical evidence shows that AI is indeed a GPT, and that the number of AI-related patents is increasing at high rates, pointing towards the fact that it is rapidly pervading our economy in a variety of different application sectors. However, patent data is intrinsically limited, and the results of the analysis on market concentration must be treated carefully, as it can only be deducted that the patent portfolios of strong patentees is increasing when compared to those of weak patentees, thus suggesting that some companies are heavily investing in AI technologies.

Being at the dawn of the economic explosion of AI, treating it as a GPT opens up new prescriptions for policy intervention. From a purely economic standpoint, we are likely to assist in increases in the innovation rate in the short term, causing disruptions in the labor market. Some jobs will cease to exist because human labor will be substituted by artificial labor (Abrardi et al., 2019), while new professional figures capable of building, managing, or operating AI systems will be needed, creating a mismatch in the labor market. If policymakers will not draft effective redistribution policies, there is a substantial risk of deepening existing inequalities between social groups, leading to unrest and discontent among the "losers" of the AI revolution. Acosta et al. (2019) studied the relationship between income inequality and political polarization, identifying how economic factors have a deep impact on voting behavior. In particular, they confirmed the assumption that governments parties get penalized in periods of poor economic performance, and proved that income inequality has become a main driver for the success of far-right political parties, while structural unemployment favors far-left parties. More interestingly, they also affirmed that political polarization was reduced in contexts where income inequality was countered by specific policies, either in the form of progressive taxation or by addressing the job market mismatch. Technological advancement can be considered one of the factors at the root of political polarization but, at the same time, while its direction and pace can be guided and influenced, it cannot be stopped. The Luddite fallacy is an ever-present threat when dealing with technological innovation, however economic and political analysis show us that states where short-sighted emergency policies have not solved structural problems, and we assist to a rise of political extremism, while at the same time they lag behind other countries in terms of economic performance, up to a point where the resources to make structural reforms lack. At the same time, excessive regulation in the digital arena may lead to a reduction in the economic competitiveness in the international scenario. The extent to which a country will be successful in dealing with the AI revolution thus will be largely depend on its ability to balance incentives for technological innovation and, the implementation of labour policies that prepare the workforce to switch occupation towards the new opportunities offered by AI technologies, and ensuring competitiveness in the digital market, that has now reached a prominent place in the economy.

AI technologies are configuring as a key factor of change in the international and national economy. They are expanding the borders of what is technically possible and conceivable by human societies, but with great power comes great responsibility. If not properly guided, mismanaged innovation can have devastating effects on social cohesion, potentially paving the way for increasing inequality and authoritarian regimes. Ethical use of AI technologies is key to ensure a future that is sustainable, rich, and fair. While regulation of what can or cannot be produced may be seen as a cost in the short term, if properly designed it may lead to enhanced social welfare in the medium-long term.

Bibliography

Abbate, J. (1999). *Inventing the Internet.* MIT Press.

Abbott, R. (2016). I think, therefore i invent: creative computers and the future of patent law. *BCL Rev. 57*, 1079.

Abbott, R. (2017). Patenting the output of autonomously inventive machines. *Landslide 10*(1), 16–22.

Abrardi, L., C. Cambini, and L. Rondi (2019). The economics of artificial intelligence: A survey. Technical report, Robert Schuman Centre for Advanced Studies Research Paper No. RSCAS.

Acosta, C. P., J. C. M. Peña, and T. Saalfeld (2019). Inequality, macroeconomic performance and political polarization: An empirical analysis. Technical report, Bamberg University, Bamberg Economic Research Group.

Agrawal, A., J. Gans, and A. Goldfarb (2018a). Prediction, judgment and complexity: A theory of decision making and artificial intelligence. Technical report, National Bureau of Economic Research.

Agrawal, A., J. Gans, and A. Goldfarb (2018b). *Prediction machines: the simple economics of artificial intelligence.* Harvard Business Press.

Agrawal, A., J. McHale, and A. Oettl (2018). Finding needles in haystacks: Artificial intelligence and recombinant growth. Technical report, National Bureau of Economic Research.

Agrawal, A., J. McHale, and A. Oettl (2019). Artificial intelligence, scientific discovery, and commercial innovation. Working Paper.

Alderete, M. V. (2020). Towards measuring the economic impact of open data by innovating and doing business. *International Journal of Innovation and Technology Management 17*, 1–20.

Alpaydin, E. (2004). *Introduction to Machine Learning.* MIT Press.

Archibugi, D. and A. Filippetti (2015). Knowledge as global public good. In *The Handbook of Global Science, Technology and Innovation. Wiley, Oxford*, pp. 479–503. Wiley Online Library.

Arrow, K. (1962, December). Economic Welfare and the Allocation of Resources for Invention. In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, NBER Chapters, pp. 609–626. National Bureau of Economic Research, Inc.

Arthur, B. and W. Polak (2006). The evolution of technology within a simple computer model. *Complex. 11*, 23–31.

Ashri, R. (2020). Core ai techniques. In *The AI-Powered Workplace*, pp. 41–57. Springer.

Aspuru-Guzik, A. and K. Persson (2018). Materials acceleration platform: Accelerating advanced energy materials discovery by integrating high-throughput methods and artificial intelligence. Technical report, Canadian Institute for Advanced Research.

Association for Computing Machinery, I. (2012). ACM computing classification system [Data set]. Retrieved from `https://dl.acm.org/ccs`. Last accessed: 2021-01-05.

Banham, M. and A. Katsaggelos (1997). Digital image restoration. *IEEE signal processing magazine 14*(2), 24–41.

Benzell, S., N. Bostrom, E. Brynjolfsson, Y. Chae, F. Chen, M. Côté, B. Faltings, K. Firth-Butterfield, J. Flaim, and D. Floreano (2019). Technology trends 2019: Artificial intelligence. Technical report, WIPO.

Bessen, J. and E. Maskin (2009). Sequential innovation, patents, and imitation. *The RAND Journal of Economics 40*(4), 611–635.

Bianchini, S., M. Müller, and P. Pelletier (2020). Deep learning in science. arXiv preprint arXiv:2009.01575.

Bloom, N., L. Garicano, R. Sadun, and J. Van Reenen (2014). The distinct effects of information technology and communication technology on firm organization. *Management Science 60*(12), 2859–2885.

Bloom, N., C. Jones, J. Van Reenen, and M. Webb (2020). Are ideas getting harder to find. *The American Economic Review 110*, 1104–1144.

Boldrin, M. and D. Levine (2013). The case against patents. *Journal of Economic Perspectives 27*(1), 3–22.

Bornmann, L. and R. Mutz (2015). Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *Journal of the Association for Information Science and Technology 66*(11), 2215–2222.

Borowiec, S. (2016, Mar). Alphago seals 4-1 victory over go grandmaster lee sedol. Last access: 20 March 2021.

Brady, M. (1985). Artificial intelligence and robotics. *Artificial intelligence 26*(1), 79–121.

Bresnahan, T. (2010). General purpose technologies. In *Handbook of the Economics of Innovation*, pp. 761–791. Elsevier.

Bresnahan, T. and M. Trajtenberg (1995). General purpose technologies 'engines of growth'? *Journal of econometrics 65*(1), 83–108.

Brynjolfsson, E., D. Rock, and C. Syverson (2017). Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. Technical report, National Bureau of Economic Research.

Brüggemann, J., P. Crosetto, L. Meub, and K. Bizer (2015). Intellectual property rights hinder sequential innovation – experimental evidence. *Research Policy 45*(10), 2054–2068.

Callon, M. (1994). Is science a public good? fifth mullins lecture, virginia polytechnic institute, 23 march 1993. *Science, Technology, & Human Values 19*(4), 395–424.

Calvano, E., G. Calzolari, V. Denicolò, and S. Pastorello (2020). Artificial intelligence, algorithmic pricing and collusion. *American Economic Review 110*(10), 3267–3297.

Campbell, M., J. Hoane, and F.-h. Hsu (2002). Deep blue. *Artificial intelligence 134*(1-2), 57–83.

Cantner, U. and S. Vannuccini (2012). A new view of general purpose technologies. Technical report, Jena Economic Research Papers.

Castelvecchi, D. (2016). Can we open the black box of ai? *Nature News 538*(7623), 20.

Catania, F., P. Crovari, M. Spitale, and F. Garzotto (2019). Automatic speech recognition: Do emotions matter? In *2019 IEEE International Conference on Conversational Data & Knowledge Engineering (CDKE)*, pp. 9–16. IEEE.

Cavallo, A. (2018). More amazon effects: online competition and pricing behaviors. Technical report, National Bureau of Economic Research.

Cecere, G., N. Corrocher, C. Gossart, and M. Ozman (2014). Technological pervasiveness and variety of innovators in green ict: A patent-based analysis. *Research Policy 43*(10), 1827–1839.

Chai, J., J. Lin, W. Zadrozny, Y. Ye, M. Stys-Budzikowska, V. Horvath, N. Kambhatla, and C. Wolf (2001). The role of a natural language conversational interface in online sales: a case study. *International Journal of Speech Technology 4*(3-4), 285–295.

Charnovitz, S. (1998). Patent harmonization under world trade rules. *J. World Intell. Prop. 1*, 127.

Chowdhury, K. (2020). Natural language processing. In *Fundamentals of Artificial Intelligence*, pp. 603–649. Springer.

Clark, A., C. Fox, and S. Lappin (2013). *The handbook of computational linguistics and natural language processing*. John Wiley & Sons.

Coase, R. (1960). The problem of social cost. In *Classic papers in natural resource economics*, pp. 87–137. Springer.

Cockburn, I., R. Henderson, and S. Stern (2018). The impact of artificial intelligence on innovation. Technical report, National bureau of economic research.

Crémer, J., Y.-A. de Montjoye, and H. Schweitzer (2019). Competition policy for the digital era. Technical report, European Commission.

Darby, M. and L. Zucker (2003). Grilichesian breakthroughs: Inventions of methods of inventing and firm entry in nanotechnology. Working Paper 9825, National Bureau of Economic Research.

David, P. (1993). Intellectual property institutions and the panda's thumb: patents, copyrights, and trade secrets in economic theory and history. *Global dimensions of intellectual property rights in science and technology 19*, 29.

De Rassenfosse, G. and B. van Pottelsberghe de la Potterie (2013). The role of fees in patent systems: Theory and evidence. *Journal of Economic Surveys 27*(4), 696–716.

Depoorter, B. and F. Parisi (2002). Fair use and copyright protection: a price theory explanation. *International Review of Law and Economics 21*(4), 453–473.

Dey, A. (2016). Machine learning algorithms: a review. *International Journal of Computer Science and Information Technologies 7*(3), 1174–1179.

Diderot, D. (1767). *Lettre sur le commerce de la librairie.* publie. net.

Drexl, J. (2010). *Research handbook on Intellectual property and Competition law.* Edward Elgar Publishing.

Dreyfus, H. (1972). *What computers can't do: The limits of artificial intelligence.* MIT Press.

Drezner, D. (2019). Technological change and international relations. *International Relations 33*, 286–303.

Dusollier, S. (2007). Sharing access to intellectual property through private ordering. *Kent Law Review 82*, 1391.

Eskenazi, M. (1999). Using automatic speech processing for foreign language pronunciation tutoring: Some issues and a prototype. *Language learning & technology 2*(2), 62–76.

European Patent Office, E. (2019). 4.1.3 economic effects; part g patentability; chapter 2.

Feldman, M. and J. W. Yoon (2012). An empirical test for general purpose technology: an examination of the cohen–boyer rdna technology. *Industrial and Corporate Change 21*(2), 249–275.

Ferrucci, D. (2012). Introduction to "this is watson". *IBM Journal of Research and Development 56*(3.4), 1–1.

Flowers, J. C. (2019). Strong and weak ai: Deweyan considerations. In *AAAI Spring Symposium: Towards Conscious AI Systems.*

Forsyth, D. and J. Ponce (2012). Computer vision: A modern approach. always learning.

Freeman, L. (1978). Centrality in social networks conceptual clarification. *Social networks 1*(3), 215–239.

Gallini, N. (1992). Patent policy and costly imitation. *The RAND Journal of Economics 23*(1), 52–63.

Gautier, A., A. Ittoo, and P. Van Cleynenbreugel (2020). Ai algorithms, price discrimination and collusion: a technological, economic and legal perspective. *European Journal of Law and Economics 50*, 405–435.

Geradin, D., C. Ahlborn, V. Denicolò, and J. Padilla (2006). Implications of the proposed framework and antitrust rules for dynamically competitive industries. Technical report, DG Comp.

Gilbert, R. and C. Shapiro (1990). Optimal patent length and breadth. *The RAND Journal of Economics 21*(1), 106–112.

Goldfarb, A. and D. Trefler (2019). Artificial intelligence and international trade. In *The Economics of Artificial Intelligence: An Agenda*, pp. 463–492. National Bureau of Economic Re.

Goldfarb, A. and C. Tucker (2017). Digital economics. *Microeconomics: Search; Learning; Information Costs & Specific Knowledge; Expectation & Speculation eJournal 57*(1), 3–43.

Goldstein, P. (2001). *International copyright: principles, law, and practice.* Oxford University Press, USA.

Gonsalves, T. (2019). The summers and winters of artificial intelligence. In *Advanced Methodologies and Technologies in Artificial Intelligence, Computer Simulation, and Human-Computer Interaction*, pp. 168–179. IGI Global.

Gonzalez, D., T. Zimmermann, and N. Nagappan (2020). The state of the ml-universe: 10 years of artificial intelligence & machine learning software development on github. In *Proceedings of the 17th International Conference on Mining Software Repositories*, New York, pp. 431–442. Association for Computing Machinery.

Goodfellow, I., Y. Bengio, A. Courville, and Y. Bengio (2016). *Deep learning*, Volume 1. MIT Press.

Grace, K., J. Salvatier, A. Dafoe, B. Zhang, and O. Evans (2018). When will ai exceed human performance? evidence from ai experts. *Journal of Artificial Intelligence Research 62*, 729–754.

Greene, D. P. (1987). Automated knowledge acquisition: Overcoming the expert system bottleneck. In *ICIS*, pp. 33.

Greenhalgh, C. and M. Rogers (2010). *Innovation, intellectual property, and economic growth.* Princeton University Press.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica 25*(4), 501–522.

Grosheide, F. W. (2002). Database protection–the european way. *Wash. UJL & Pol'y 8*, 39.

Guellec, D. and B. V. P. de La Potterie (2007). *The economics of the European patent system: IP policy for innovation and competition.* Oxford University Press on Demand.

Gupta, S., K. Bhurchandi, and A. Keskar (2016). An efficient noise-robust automatic speech recognition system using artificial neural networks. In *2016 International Conference on Communication and Signal Processing (ICCSP)*, pp. 1873–1877. IEEE.

Haenlein, M. and A. Kaplan (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California management review 61*(4), 5–14.

Hall, B. (2005). A note on the bias in herfindahl-type measures based on count data. *Revue d'économie industrielle 110*(1), 149–156.

Hall, B. and M. Trajtenberg (2004). Uncovering gpts with patent data. Technical report, National Bureau of Economic Research.

Helpman, E. and M. Trajtenberg (1994). A time to sow and a time to reap: growth based on general purpose technologies. Technical report, National Bureau of Economic Research.

Helpman, E. and M. Trajtenberg (1996). Diffusion of general purpose technologies. Working Paper 5773, National Bureau of Economic Research.

Henderson, J. (2010). Artificial neural networks. In *The Handbook of Computational Linguistics and Natural Language Processing*, pp. 221. Wiley Online Library.

Henderson, R., A. Jaffe, and M. Trajtenberg (1995). Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988. Working Paper 5068, National Bureau of Economic Research.

Hentschel, K. (2015). A periodization of research technologies and of the emergency of genericity. *Studies in History and Philosophy of Modern Physics 52*, 223–233.

Herrera, G. (2003). Technology and international systems. *Millennium 32*(3), 559–593.

Herrera, G. (2012). *Technology and international transformation: The railroad, the atom bomb, and the politics of technological change.* SUNY Press.

Hettinger, E. (1989). Justifying intellectual property. *Philosophy & Public Affairs 18*(1), 31–52.

Hilty, R., J. Hoffmann, and S. Scheuerer (2020). Intellectual property justification for artificial intelligence. In *Artificial Intelligence & Intellectual Property.* Oxford University Press. Draft chapter.

Howells, J. and R. Katznelson (2016). The 'overly-broad' selden patent, henry ford and development in the early us automobile industry.

Hunt, R. (2004). Patentability, industry structure, and innovation. *The Journal of Industrial Economics 52*(3), 401–425.

Imagination Engines Inc. (2021). What is dabus? `http://imagination-engines.com/iei_dabus.php`. Accessed: 2021-02-14.

Jaffe, A., M. Trajtenberg, and R. Henderson (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly journal of Economics 108*(3), 577–598.

Jiménez, S., T. De La Rosa, S. Fernández, F. Fernández, and D. Borrajo (2012). A review of machine learning for automated planning. *The Knowledge Engineering Review 27*(4), 433–467.

Jones, B. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies 76*(1), 283–317.

Jordan, M. and T. Mitchell (2015). Machine learning: Trends, perspectives, and prospects. *Science 349*(6245), 255–260.

Jovanovic, B. and P. Rousseau (2005). General purpose technologies. In *Handbook of economic growth*, Volume 1, pp. 1181–1224. Elsevier.

Kenney, M. and J. Zysman (2016). The rise of the platform economy. *Issues in science and technology 32*(3), 61.

Kitch, E. (1977). The nature and function of the patent system. *The Journal of Law and Economics 20*(2), 265–290.

Klein, T. (2019). Autonomous algorithmic collusion: Q-learning under sequential pricing.

Klemperer, P. (1990). How broad should the scope of patent protection be? *The RAND Journal of Economics 21*(1), 113–130.

Klinger, J., J. Mateos-Garcia, and K. Stathoulopoulos (2018). Deep learning, deep change? mapping the development of the artificial intelligence general purpose technology. Working Paper.

Koboldt, C. (1996). The eu-directive on the legal protection of databases and the incentives to upyear: An economic analysis. Technical report, CSLE Discussion Paper.

Korinek, A. and N. Ding Xuan (2018). Digitization and the macroeconomics of superstars. Working Paper.

Korzinov, V. and I. Savin (2016). Pervasive enough? general purpose technologies as an emergent property. Working paper series in economics, Karlsruhe Institute of Technology (KIT), Department of Economics and Management.

Kotsiantis, S. (2007). Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering 160*(1), 3–24.

Landes, W. and R. Posner (1989). An economic analysis of copyright law. *The Journal of Legal Studies 18*(2), 325–363.

Landes, W. and R. Posner (2003). Indefinitely renewable copyright. *The University of Chicago Law Review 70*(2), 471–518.

Langinier, C. and G. Moschini (2002). The economics of patents. In *Intellectual Property Rights in Animal Breeding and Genetics*, pp. 31–50. CABI.

Lee, P. (2010). Patent law and the two cultures. *Yale LJ 120*, 2.

Lee, Y.-L., P.-K. Tsung, and M. Wu (2018). Techology trend of edge ai. In *2018 International Symposium on VLSI Design, Automation and Test (VLSI-DAT)*, pp. 1–2. IEEE.

Leonetti, M., L. Iocchi, and P. Stone (2016). A synthesis of automated planning and reinforcement learning for efficient, robust decision-making. *Artificial Intelligence 241*, 103–130.

Lin, P. (2011). Introduction. In *Robot Ethics: The Ethical and Social Implications of Robotics.*

Lipsey, R., K. Carlaw, and C. Bekar (2005). *Economic transformations: general purpose technologies and long-term economic growth.* OUP Oxford.

Locke, J. (1689). Chapter 5: Property.

Lohn, J., G. Hornby, and D. Linden (2005). An evolved antenna for deployment on nasa's space technology 5 mission. In *Genetic Programming Theory and Practice II*, pp. 301–315. Springer.

Lévêque, F. and Y. Ménière (2004). *The economics of patents and copyright.* Berkeley Electronic Press, July.

Machlup, F. and E. Penrose (1950). The patent controversy in the nineteenth century. *The Journal of Economic History 10*(1), 1–29.

Martens, B. (2018). The impact of data access regimes on artificial intelligence and machine learning. Technical report, European Commission.

Matutes, C., P. Regibeau, and K. Rockett (1996). Optimal patent design and the diffusion of innovations. *The RAND Journal of Economics 27*(1), 60–83.

McCulloch, W. and W. Pitts (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology 5*, 115–133.

McGowan, R. and N. Corrado (2019). Colliding worlds: An aging workforce, artificial intelligence, and the impact on the future of work. Technical report, AARP International.

Menell, P. (1999). Intellectual property: General theories. *Encyclopedia of law and economics 2*, 129–188.

Merges, R. (2009). The trouble with trolls: innovation, rent-seeking, and patent law reform. *Berkeley Tech. LJ 24*, 1583.

Meurer, M. and C. A. Nard (2004). Invention, refinement and patent claim scope: A new perspective on the doctrine of equivalents. *Geo. LJ 93*, 1947.

Milgrom, P. and S. Tadelis (2018). How artificial intelligence and machine learning can impact market design. Working Paper 24282, National Bureau of Economic Research.

Mohler, J. (1999). Toward a better understanding of substantial similarity in copyright infringement cases. *U. Cin. L. Rev. 68*, 971.

Mokyr, J. (1990). Punctuated equilibria and technological progress. *The American Economic Review 80*, 350–354.

Mokyr, J. (2018). The past and the future of innovation: Some lessons from economic history. *Explorations in Economic History 69*, 13–26.

Moser, P. and T. Nicholas (2004). Was electricity a general purpose technology? evidence from historical patent citations. *American Economic Review 94*(2), 388–394.

Murphy, R. (2000). *Introduction to AI Robotics*. MIT Press.

Nardin, A. (2020). Rops: an r package to access ops api by epo. `https://github.com/AlessioNar/Rops`.

Nardin, A. (2021). Mythesis. `https://github.com/AlessioNar/MyThesis`.

Nassif, A. B., I. Shahin, I. Attili, M. Azzeh, and K. Shaalan (2019). Speech recognition using deep neural networks: A systematic review. *IEEE Access 7*, 19143–19165.

Nelson, R. (1959). The simple economics of basic scientific research. *Journal of political economy 67*(3), 297–306.

Neuhäusler, P. and R. Frietsch (2019). Computer-implemented inventions in europe. In *Springer Handbook of Science and Technology Indicators*, pp. 1007–1022. Springer.

Newell, A. (1980). Physical symbol systems. *Cognitive science 4*(2), 135–183.

Newell, A. and H. Simon (2007). Computer science as empirical inquiry: Symbols and search. In *ACM Turing award lectures*, pp. 1975.

Nilsson, N. (2009). *The quest for artificial intelligence*. Cambridge University Press.

Nordhaus, W. (1969a). An economic theory of technological change. *The American Economic Review 59*(2), 18–28.

Nordhaus, W. (1969b). *Invention, growth and welfare: a theoretical treatment of technological change*. MIT Press.

Nozick, R. (1974). *Anarchy, state, and utopia*, Volume 5038. New York: Basic Books.

Nuccio, M. and M. Guerzoni (2019). Big data: Hell or heaven? digital platforms and market power in the data-driven economy. *Competition & Change 23*(3), 312–328.

O'Donoghue, T. (1998). A patentability requirement for sequential innovation. *The RAND Journal of Economics 29*(4), 654–679.

OECD (2019). Oecd han database. `http://www.oecd.org/science/inno/intellectual-property-statistics-and-analysis.htm#ip-data`. Accessed: 2020-10-01.

Park, W. (2008). International patent protection: 1960–2005. *Research policy 37*(4), 761–766.

Partridge, D. (1987). The scope and limitations of first generation expert systems. *Future Generation Computer Systems 3*(1), 1–10.

Pavitt, K. (1987). On the nature of technology. In *Technology, Management, and Systems of Innovation*, pp. 3–14. Edward Elgar.

Pearlman, R. (2017). Recognizing artificial intelligence (ai) as authors and investors under us intellectual property law. *Rich. JL & Tech. 24*, i.

Petralia, S. (2020). Mapping general purpose technologies with patent data. *Research Policy 49*, 104013.

Potamianos, A., S. Narayanan, and S. Lee (1997). Automatic speech recognition for children. In *Fifth European Conference on Speech Communication and Technology.*

Pouyanfar, S., S. Sadiq, Y. Yan, H. Tian, Y. Tao, M. P. Reyes, M.-L. Shyu, S.-C. Chen, and S. Iyengar (2018). A survey on deep learning: Algorithms, techniques, and applications. *ACM Computing Surveys (CSUR) 51*(5), 1–36.

Pratt, G. (2015). Is a cambrian explosion coming for robotics? *Journal of Economic Perspectives 29*(3), 51–60.

Prince, S. (2012). *Computer vision: models, learning, and inference.* Cambridge University Press.

Rajan, K. and A. Saffiotti (2017). Towards a science of integrated ai and robotics.

Ray, T. (1998). A case study of britain's response to miti's fifth generation computer initiative. *Technology and Innovation in Japan: Policy and Management for the Twenty First Century 18*, 151.

Rose, M. (1988). The author as proprietor: Donaldson v. becket and the genealogy of modern authorship. *Representations 23*, 51–85.

Rosenfeld, A. (1988). Computer vision: basic principles. *Proceedings of the IEEE 76*(8), 863–868.

Russell, S. and P. Norvig (2010). *Artificial intelligence: a modern approach.* test.

Samuelson, P. (1954). The pure theory of public expenditure. *The review of economics and statistics 36*(4), 387–389.

Sanchez-Lengeling, B. and A. Aspuru-Guzik (2018). Inverse molecular design using machine learning: Generative models for matter engineering. *Science 361*(6400), 360–365.

Schank, R. (1983). The current state of ai: One man's opinion. *AI Magazine 4*(1), 3–3.

Schumpeter, J. (1934). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle.* Transaction Publishers.

Schumpeter, J. (1942). *Capitalism, socialism and democracy.* routledge.

Scotchmer, S. (2004). *Innovation and incentives.* MIT press.

Searle, J. (1990). Minds, brains, and programs. In *The Philosophy of Artificial Intelligence.*

Shemtov, N. (2019). A study of inventorship in inventions involving ai activity.

Shulman, S. (2003). Unlocking the sky: Glenn hammond curtiss and the race to invent the airplane. *Aviation, Space, and Environmental Medicine 74*(2), 189–189.

Silverman, B. (2002). Documentation for international patent classification–us sic concordance.

Sparck Jones, K. (1999). What is the role of nlp in text retrieval? In *Natural language information retrieval*, pp. 1–24. Springer.

Sparkes, A., W. Aubrey, E. Byrne, A. Clare, M. Khan, M. Liakata, M. Markham, J. Rowland, L. Soldatova, and K. Whelan (2010). Towards robot scientists for autonomous scientific discovery. *Automated Experimentation 2*(1), 1.

Spence, M. (2007). *Intellectual Property.* Oxford University Press.

Stiglitz, J. (1999). Knowledge as a global public good. *Global public goods: International cooperation in the 21st century 308*, 308–325.

Tandon, P. (1982). Optimal patents with compulsory licensing. *Journal of Political Economy 90*(3), 470–486.

Tang, Y., X. Lou, Z. Chen, and C. Zhang (2020). A study on dynamic patterns of technology convergence with ipc co-occurrence-based analysis: The case of 3d printing. *Sustainability 12*, 2655.

Teece, D. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research policy 15*(6), 285–305.

Toole, A., N. Pairolero, A. Giczy, J. Forman, C. Pulliam, M. Such, K. Chaki, D. Orange, A. T. Homescu, J. Frumkin, Y. Y. Chen, V. Gonzales, C. Hannon, S. Melnick, E. Nilsson, and B. Rifkin (2020). Inventing ai: Tracing the diffusion of artificial intelligence with us patents.

Townley, C., E. Morrison, and K. Yeung (2017). Big data and personalized price discrimination in eu competition law. *Yearbook of European Law 36*, 683–748.

Trucco, E. and A. Verri (1998). *Introductory techniques for 3-D computer vision*, Volume 201. Prentice Hall Englewood Cliffs.

US Congress (2011). Leahy-smith america invents act - public law 112–29. `https://www.uspto.gov/sites/default/files/aia_implementation/20110916-pub-l112-29.pdf`.

VerSteeg, R. (1995). Defining author for purposes of copyright. *Am. UL Rev. 45*, 1323.

Vlasov, S. (2020). harmonizer - an r package to harmonize organizational names. `https://github.com/stasvlasov/harmonizer`.

Waltz, D. (1983). Artificial intelligence: An assessment of the state-of-the-art and recommendations for future directions. *AI Magazine 4*(3), 55–55.

Wang, Y., Y. Chen, and B. Koo (2020). Open to your rival: Competition between open source and proprietary software under indirect network effects. *Journal of Management Information Systems 37*, 1128–1154.

Wasserman, S., K. Faust, et al. (1994). *Social network analysis: Methods and applications*, Volume 8. Cambridge university press.

Wawrzyniak, R. (1990). Proving patent infringement under the doctrine of equivalents: The specific evidentiary requirements for getting to the jury. *Missouri Law Review 55*(4), 1–18.

Weiss, C. (2005). Science, technology and international relations. *Technology in Society 27*, 295–313.

Weiss, C. (2015). How do science and technology affect international affairs? *Minerva 53*(4), 411–430.

Weitzman, M. (1998). Recombinant growth. *The Quarterly Journal of Economics 113*(2), 331–360.

Weld, D. (1999). Recent advances in ai planning. *AI magazine 20*(2), 93–93.

Wells, M. (2001). Internet business method patent policy. *Virginia Law Review 87*(4), 729–780.

WIPO (2004). *WIPO intellectual property handbook: Policy, law and use*, Volume 489. WIPO.

Appendix A

A note on closed-loop systems

Recently, a lively debate has sparkled around the notion that some Artificial Intelligence systems can perform the research process autonomously, especially within the academic community that investigated the legal aspects of intellectual property rights (Abbott, 2016, 2017; Pearlman, 2017; Shemtov, 2019). I argue that the excitement around the inventing capabilities of AI systems is mainly caused by a lack of comprehension of their inner workings and that state-of-the-art Autonomous Discovery Systems (from now on, ADS) are either very far from being completely autonomous or inapplicable in fields other than the extremely specific ones they were designed for. The legal arguments are often more based on conjectures and suppositions than on concrete use-cases of ADSs. In the next paragraphs, I will examine three examples of ADS products: the NASA antenna, the Robot Scientist project, and DABUS.

The **Robot Scientist project** built two Automated Discovery Systems (ADS), ADAM and EVE, which automated the entire research pipeline. ADAM is an ADS designed to *"carry out microbial growth experiments to study functional genomics in the yeast Saccharomyces cervisiae, specifically to identify the genes encoding 'locally orphan enzymes'"* (Sparkes et al., 2010), while Eve is designed to find new drugs to cure neglected tropical diseases.

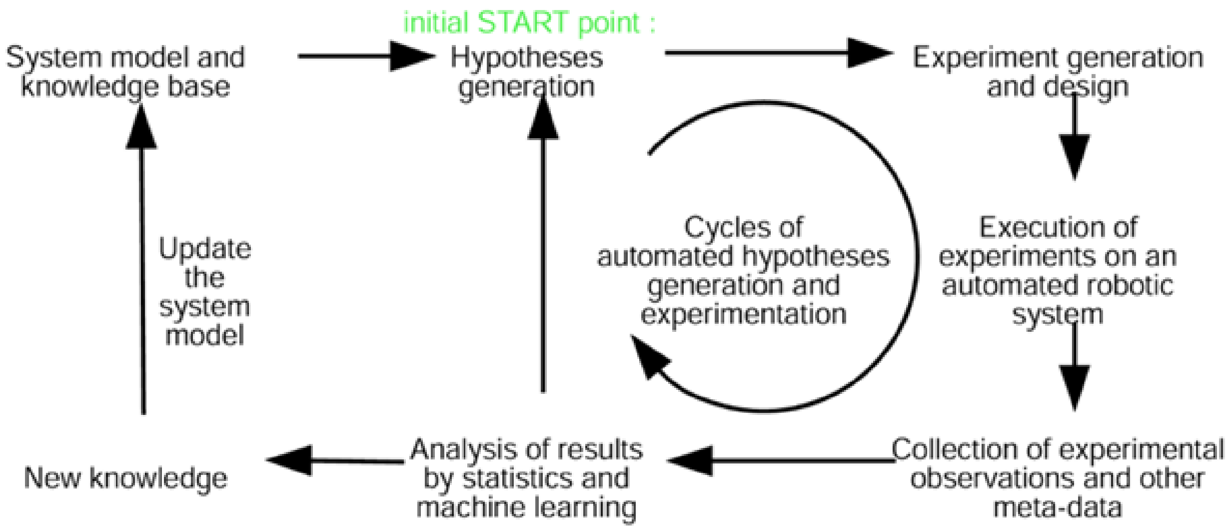


Figure A.1: The workflow of the ADAM robot scientist (Sparkes et al., 2010)

These systems use a *"combination of computational methods, automated instruments, closed-loop learning, advanced laboratory robotic systems and formal logical expression of information"* (Sparkes et al., 2010) to gain new knowledge. By tasking the AI ayarwm to conduct cycles of experimentation on a robotic laboratory, the robot scientist *"automatically generates hypothesis from the available background knowledge and models, design physical experiments to test them, carries out experiments, and then analyzes and interprets the results"* (Sparkes et al., 2010). This approach has several advantages, such as capturing every detail of the scientific discovery process: goals, hypotheses, results, and conclusions, while gathering useful meta-data such as environmental conditions, detailed content layout information, instrument settings, protocols, and runtime logs, fundamental to achieve explainability. However, despite these unquestionable advantages, ADAM is only effective in a narrow field. Its algorithms were trained on data that were first selected by a scientist, and it was designed exclusively for research on yeast. Repurposing ADAM requires considerable time and effort.

The amount of work that had to be done to build ADAM shows how it is challenging to consider it fundamentally different from any other research tool, thus inserting its by-products in the category of AI-aided inventions.

Another approach to *"automate invention"* is through the use of **genetic programming**. Genetic programming is an AI technique where computer programs are encoded as a set of genes that are then modified using an evolutionary algorithm. NASA used this technique in 2004 to develop an antenna with specific requirements (Lohn et al., 2005). After the developer provided the genetic algorithm with some parametric information regarding the final product, the software started producing random designs, and, through an evolutionary process, the AI simulated various prototypes keeping only the one closer to the given requirements. When the specifications were reached, the AI system stopped. Genetic programming is also used in intelligent manufacturing, where it tests possible improvements in the production process. Based on a series of parameters, the AI detects the most efficient way to improve production in a specific automatized plant. The developer plays a fundamental role in choosing the environment and in providing the data needed for the innovation to take place, a process that he/she needs to repeat every time the algorithm is repurposed.

Other claims regarding existing ADS are much debated since they do not provide a precise technical explanation of their computational process. A curious case is the one regarding DABUS (Device for the Autonomous Bootstrapping of Unified Sentience) that, according to its creator, Stephen Thaler, it has developed a food container and various methods capable of attracting enhanced attention in a digital setting. Differently from other ADSs, the detail information regarding the functioning of DABUS is kept secret, and Thaler himself provided a vague explanation, even in his patent application.

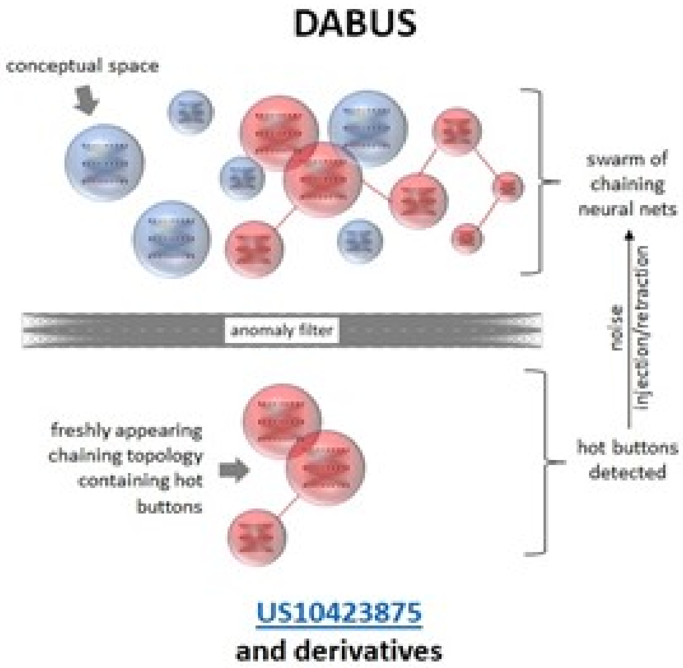


Figure A.2: DABUS, (Imagination Engines Inc., 2021)

According to what it is claimed in the project website, *"the DABUS architecture consists of a multitude of neural nets, with many ideas forming in parallel across multiple computers"* (Imagination Engines Inc., 2021). The process of idea creation is formed by a series of disconnected neural nets that contain different interrelated *"memories"* of a linguistic, visual, or auditory nature. When DABUS is activated, nets combine and detach, following a controlled chaos algorithm that operates both within and between them. In this way, *"cumulative cycles of learning and unlearning"* making a fraction of these nets *"interconnect into structures representing complex concepts"* (Imagination Engines Inc., 2021). These structures then tend to connect continuously with others that allegedly represent their *"consequences'* in cycles of creation and disruption of connections.

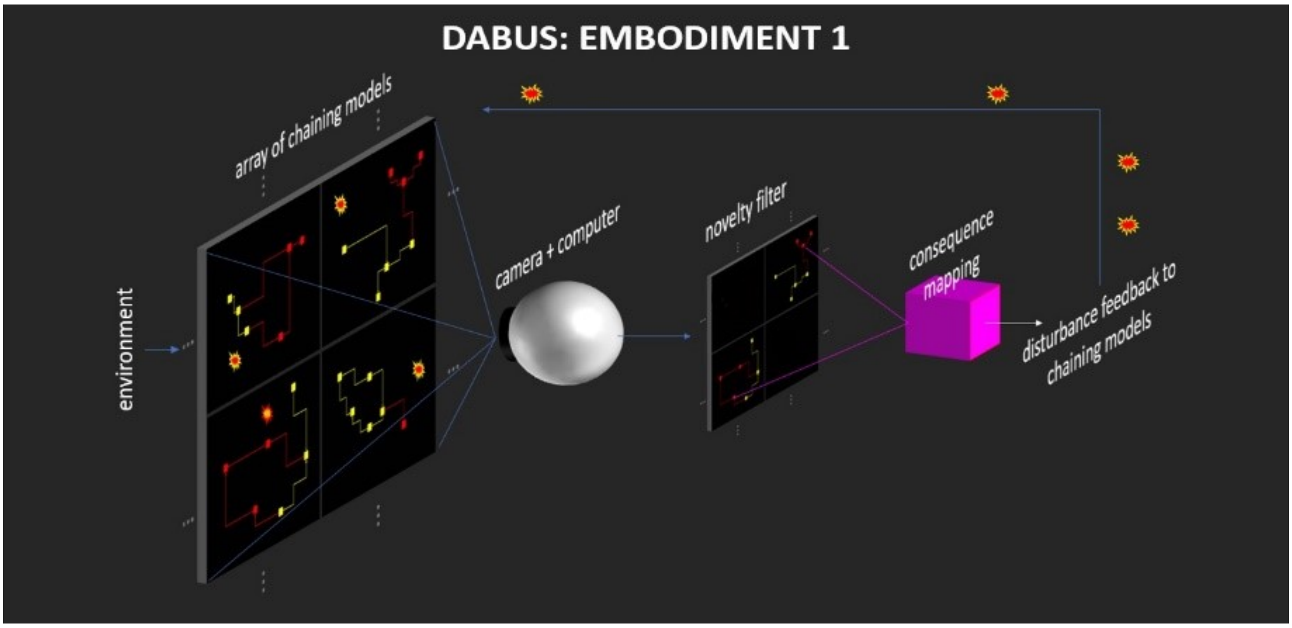


Figure A.3: Representation of DABUS "thought process" (Imagination Engines Inc., 2021)

At the same time, DABUS applies what they define *"novelty filter"* or *"anomaly filter, [...] adaptive neural nets that absorb the status quo within any environment and emphasize any departures from such normalcy,"* to detect and isolate freshly forming concepts (Imagination Engines Inc., 2021). These filters allegedly serve to identify critical neural nets (also called *"hot buttons"*) that are considered essential for achieving one or more desirable outcomes, thus triggering the release of signals that reinforce or destroy a concept chain. After a certain threshold, *"such ideas are converted into long term memories, eventually allowing DABUS to be interrogated for cumulative inventions and discoveries"* (Imagination Engines Inc., 2021).

While at first sight, DABUS may appear as an *"invention machine,"* as claimed by its creator, when what it has been disclosed so far is analyzed in detail, several observations arise. First, it seems like the basic concepts draw a lot from genetic programming, where evolutionary algorithms detect anomalies and suppress *"unfit"* combinations. However, rather than the criteria being the degree of adherence to specific parameters, it seems that the benchmark consists of eluding prior art and provide some degree of novelty. In this sense, DABUS should not be considered as an actual *"invention machine"* but instead an *"inventing-around machine,"* that looks the right combinations to elude the breadth of already existing patents without infringing other patents. However, DABUS operations are limited to automating the process of identifying novelty, and it cannot be said that it understands the *"ideas"* that allegedly produces. Recalling the Chinese Room test proposed by Searle (1990), it would seem that there is no difference between DABUS and other AIs, such as machine translation models. As even Thaler affirmed, the outcome of the allegedly generated inventions largely depends on the provided data: *"DABUS has autonomously generated two fractal-related inventions, but that's because this AI-child was raised in a 'household' wherein fractal theory was often discussed"* (Imagination Engines Inc., 2021). As a result, is it correct to affirm that DABUS generated these inventions? Should not instead be considered an instrument that, based on a selection of ideas chosen by an operator, produces something new by randomly trying out combinations? When framed like this, the control over what is being invented shifts from DABUS to whoever selects the initial data. What I suggest is instead that DABUS should not be considered an *"artificial inventor"* at all (as claimed by Thaler), but rather an extremely sophisticated tool able to enhance human capabilities. So, even if Thaler's propositions over DABUS capabilities respect reality, an operator still controls what is being invented. Again, the by-products of DABUS should be considered AI-assisted inventions.

The examples provided in this paragraph represent the state-of-the-art in Automated Discovery Systems. However, all these systems still involve in some way human intervention, that occurs either through the decision of the direction of the discovery process, by explicitly tasking an ADS towards the development of a product with specific characteristics (inverse design, genetic programming), or by providing the ADS with a selection of pre-processed data (DABUS). Moreover, the scope of application of these ADSs is still extremely narrow. At present, extensive training and/or modifications in the physical apparatus represent a bottleneck for the repurposing of closed-loop ADS towards different goals.

Appendix B

Data collection

When dealing with patent data, a fundamental step concerns identifying which patents are associated with a specific technology. This generally involves two steps:

- Defining the boundaries of a product concerning a technology;
- Determining the ways in which a that boundaries can be identified in patent data.

When we apply this procedure to AI technologies, the first step is particularly problematic. As it was already mentioned in chapter 1, the definition of AI has changed many times in the past, and it is particularly complex for the researcher to codify this definition in a structured query. For the purpose of this dissertation, I applied the broadest possible definition of AI, following the model provided in Benzell et al. (2019).

The second step (determining the way in which such boundaries can be identified in patent data), involves a different set of decisions by the researcher. Among the literature on the methodology of patent analysis, two main method are used to individuate and locate a specific technology field. One is generally based on analyzing sequential parts of text of variable length and scope (such as abstract, claims, or description) of patent documents, and may be based on text mining techniques, using either heuristic, machine learning and deep learning techniques (in most cases mixed methodologies are used). The other is based on using metadata and bibliographic information, such as technological classes. Both approaches have advantages and disadvantages, and it is common practice to use a mixed methodology. For the purposes of this thesis, I used the approach adopted by WIPO (Benzell et al., 2019), in which three different queries were used to identify AI technologies, thus leading to determining that a patent involves the use of AI technologies if is present in either one of the three queries.

The first query was structured to target patent applications that had associated the Cooperative Patent Classification technological classes retained as specific of AI technologies by experts of the field. This led to the retrieval of 54816 unique patent applications. The second query was based on the hypothesis that a relatively large amount of AI-related patents may have been classified in non-specifically AI-related classification codes and could only be captured using keywords. The second subset of unique patents that contained a selected list of keywords in their titles or abstracts returned 6651 unique patent applications.

The third query aimed to combine symbol-based and keyword-based search, retrieving unique patents that were assigned either one of the International Patent Classification technological classes or one of a second group of CPC technological classes, and which contained in their titles and abstracts at least one of the keywords present in a second selected list. This has led to the retrieval of 34269 unique applications. Differently from the WIPO query strategy, this third query was limited to CPC and IPC codes, since the sample choice was restricted to PCT patents, which have no technological classes assigned using the Japanese scheme.

After dropping duplicate patents, the total subset retrieved contained 91797 unique applications.

The classification codes and keywords used for the query are contained in the csv files `query_codes.csv` and `keywords.csv`, available in the folder `data/data_gathering` of the GitHub repository of this dissertation (Nardin, 2021).

B.1 Control sample

To compare the result of the generality index of AI patents, I built a control sample to verify whether AI-related patents had, on average, a higher generality. The control sample was built on three criteria: the patents had to be filed at the same patent authority (in our case, WIPO), in the same year, and had to have the most similar number of forward citations. The matching strategy was 1:1, with no other weights assigned. This choice was motivated by the fact that by using a matching strategy based on technological classes would have increased the risk of including in the control sample false negatives.

B.2 An alternative source of data: Open Patent Services

During the process of writing this thesis, before I gained access to the PATSTAT 2018b, I used the Open Patent Services API to search for patents related to AI. To facilitate the authentication process and perform query directly from R, I wrote various functions that I united in an R package, Rops, available on GitHub (Nardin, 2020).

Appendix C

Applicant disambiguation

The process of applicant harmonization is complex and subjected to a trade-off between precision and accuracy. One of the main issues in the analysis of patent applicants derives from the different ways their name is registered in patent documents. A firm may apply for a patent through different branches according to their geographical localization, legal status, corporate policy, or other reasons. Name differentiation may even be a deliberate choice by companies, to actively obstacle business intelligence operations of competitors. Some patent authorities compiled databases with the purpose of partially overcome this issue. Datasets compiled by trustworthy authorities are particularly useful for our research purposes. For the purposes of this thesis I used the HAN database, developed by OECD (2019), that associated to the patent identifier `appln_id` a large number of harmonized applicants' names. PATSTAT provides such feature in the table `tls206_pers`. Unfortunately, the HAN database is rich of false negatives, since it prioritizes accuracy over completeness. For the purpose of this research, a higher degree of completeness was required, thus the applicant sample data was subjected to additional harmonization steps.Another issue of applicant name harmonization derives from the fact that patents are fully transferable property items and that the PATSTAT database does not keep track of property changes after its publication thus without following it during the patent life-span. However, since we are mostly interested in determining the innovator that developed the original invention, this limit can be left out.

First, patent applicants for the AI sample were retrieved from the original database, that returned 96729 unique patent applicants. Then a disambiguation procedure was applied and the number of unique applicants was reduced to 93646 unique applicants names. The harmonized applicant names coming from the HAN database were first subjected to the harmonize function contained in the harmonizer package (Vlasov, 2020), that performs parsing operations using the following algorithm:

1. Cleaning spaces
2. Removing HTML codes
3. Translating non-ASCII to ASCII
4. Upper casing

5. Standardizing organizational names
6. Removing brackets
7. Cleaning spaces

In particular, during the fifth step, the harmonize function looks for standard company information such as 'corporation', 'company', and 'limited', or information related to the geographical position of the applicants such as 'America', or 'Europe', and transforms them in a standardized format such as CORP, IN, LT, USA, or EU that are later added to the final part of the application name.

Next, these particles were removed to increase the harmonization. The complete list of the stop-words used in the procedure can be found in the `data/applicants/stopwords.csv` file in the Github repository (Nardin, 2021). This allowed to gather under the same applicant patents that were previously associated to different applicant names, such as 'GOOGLE LL', 'GOOGLE LT' and 'GOOGLE USA'.

Further harmonization would require to group companies with different names that are part of the same conglomerate. However, it would require keeping track of mergers and acquisitions regarding companies and I had no access to such database, since they are mostly licensed on a commercial basis.

Appendix D

Further descriptive statistic

Table D.1: Yearly patents application filed

Year	AI	PCT	AI/PCT
1995	570	39059	1.459
1996	623	47097	1.323
1997	809	55905	1.447
1998	1158	65662	1.764
1999	1460	74932	1.948
2000	2273	91434	2.486
2001	2678	106215	2.521
2002	2654	108662	2.442
2003	2432	113403	2.145
2004	2458	120643	2.037
2005	2803	134605	2.082
2006	3138	149345	2.101
2007	3458	157557	2.195
2008	3484	160537	2.170
2009	3272	152745	2.142
2010	3607	161441	2.234
2011	4549	179314	2.537
2012	4874	191608	2.544
2013	6459	201233	3.210
2014	7451	210153	3.546
2015	8635	212737	4.059
2016	10434	226513	4.606
2017	12518	237452	5.272

Table D.2: Growth rate of patent applications filed using the PCT route

Interval	AI	PCT
1995-1996	9.30	20.58
1996-1997	29.86	18.70
1997-1998	43.14	17.45
1998-1999	26.08	14.12
1999-2000	55.68	22.02
2000-2001	17.82	16.17
2001-2002	-0.90	2.30
2002-2003	-8.36	4.36
2003-2004	1.07	6.38
2004-2005	14.04	11.57
2005-2006	11.95	10.95
2006-2007	10.20	5.50
2007-2008	0.75	1.89
2008-2009	-6.08	-4.85
2009-2010	10.24	5.69
2010-2011	26.12	11.07
2011-2012	7.14	6.86
2012-2013	32.52	5.02
2013-2014	15.36	4.43
2014-2015	15.89	1.23
2015-2016	20.83	6.48
2016-2017	19.97	4.83

Year	AI	PCT	ratio
1995	87	626	13.898
1996	77	628	12.261
1997	93	635	14.646
1998	111	634	17.508
1999	119	636	18.711
2000	133	634	20.978
2001	139	633	21.959
2002	157	640	24.531
2003	136	638	21.317
2004	132	639	20.657
2005	143	638	22.414
2006	141	642	21.963
2007	150	639	23.474
2008	153	641	23.869
2009	157	645	24.341
2010	160	649	24.653
2011	195	642	30.374
2012	208	648	32.099
2013	228	644	35.404
2014	249	643	38.725
2015	280	650	43.077
2016	288	652	44.172
2017	300	648	46.296

Table D.3: Variety of CPC technological classes

Year	AI	PCT	ratio
1995	88	606	14.521
1996	78	603	12.935
1997	69	574	12.021
1998	90	573	15.707
1999	97	578	16.782
2000	120	589	20.374
2001	119	585	20.342
2002	132	593	22.260
2003	139	618	22.492
2004	145	620	23.387
2005	136	619	21.971
2006	137	624	21.955
2007	132	623	21.188
2008	141	618	22.816
2009	136	616	22.078
2010	143	619	23.102
2011	159	617	25.770
2012	173	618	27.994
2013	179	616	29.058
2014	210	613	34.258
2015	225	616	36.526
2016	234	619	37.803
2017	254	623	40.770

Table D.4: Variety of IPC technological classes

Year	0%-90%	90%-99%	Top 1%
1995	0.69	0.25	0.06
1996	0.67	0.28	0.04
1997	0.73	0.22	0.05
1998	0.70	0.21	0.08
1999	0.65	0.23	0.12
2000	0.61	0.24	0.16
2001	0.69	0.23	0.08
2002	0.58	0.28	0.14
2003	0.60	0.24	0.15
2004	0.56	0.27	0.17
2005	0.56	0.28	0.16
2006	0.53	0.25	0.22
2007	0.53	0.25	0.22
2008	0.49	0.29	0.22
2009	0.51	0.27	0.22
2010	0.47	0.28	0.25
2011	0.42	0.28	0.29
2012	0.43	0.28	0.29
2013	0.38	0.27	0.35
2014	0.38	0.29	0.34
2015	0.37	0.26	0.37
2016	0.35	0.27	0.38
2017	0.34	0.27	0.39

Table D.5: Proportion of patents by applicant group