

INTELLECTUAL PROPERTY RIGHTS, ARTIFICIAL INTELLIGENCE AND BIG DATA: FUTURE PERSPECTIVES

APLICACIONES DE LA PROPIEDAD INTELECTUAL EN LA INTELIGENCIA ARTIFICIAL Y BIG DATA: PERSPECTIVAS FUTURAS

PILAR ÍÑIGUEZ*

ABSTRACT

The purpose of this paper is to analyze the challenges that the development of Artificial Intelligence implies for the IP rights (particularly in patent, sui generis database right, copyrights and trade secret), particularly, in relation with Big Data, that increasingly available in all areas of manufacturing and operations which presents an opportunity for better decision making and discovery of the next generation of innovative technologies. However, around them, there are serious doctrinal and normative challenges that we will try to analyze in this study, as well as addressing the recent jurisprudence that circumscribes the subject currently under discussion.

Key words: artificial intelligence, Big Data, patents, copyright law, database rights, trade secrets.

RESUMEN

La finalidad de este artículo es analizar los retos que el desarrollo de la Inteligencia Artificial implica en los derechos de Propiedad Intelectual (particularmente en patentes, derecho sui generis de base de datos, derechos de autor y secreto empresarial), de manera particular, con relación al Big Data, el cual se encuentra cada vez más disponible en todas las áreas de fabricación y operaciones, presentando una oportunidad no solo para una adecuada toma de decisiones, sino también, para valorar descubrimientos de la próxima generación de tecnologías innovadoras. No obstante, en torno a ellos, son planteados serios desafíos doctrinales y normativos que trataremos de analizar en el presente estudio, así como abordar la reciente jurisprudencia que circunscribe el tema objeto de discusión.

Palabras clave: inteligencia artificial, Big Data, patentes, bases de datos, derechos de autor, secreto empresarial.

* Professor of Commercial Law (University of Alicante). Dirección de correo electrónico: pilar.i@ua.es. Contribution that is part of the Research Project DER2017-87675-R of the Ministry of Economy, Industry and Competitiveness entitled: «Legal Regime of Plant Breeding and Distinctive Signs of Quality», directed by Professor Esperanza GALLEGO SÁNCHEZ.

CONTENTS: I. INTRODUCTION.—II. CONCEPT OF ARTIFICIAL INTELLIGENCE.—1. Special reference to patentability of Artificial Intelligence.—III. DEFINING BIG DATA.—IV. INTERFACE BETWEEN INTELLECTUAL PROPERTY LAW AND BIG DATA.—V. RELATION WITH ANOTHER SYSTEM PROTECTION: COPYRIGHT, DATABASES AND TRADE SECRETS.—1. Databases *sui generis* right.—2. Big Data protected as trade secrets and confidential information.—VI. CONCLUSIONS.—VII. BIBLIOGRAPHY.

SUMARIO: I. INTRODUCCIÓN.—II. CONCEPTO DE INTELIGENCIA ARTIFICIAL.—1. Especial referencia a la patentabilidad de la Inteligencia Artificial.—III. CONCEPTO DE BIG DATA.—IV. RELACIONES ENTRE PROPIEDAD INTELLECTUAL Y BIG DATA.—V. RELACIÓN CON OTROS DE DERECHOS DE PROTECCIÓN: COPYRIGHT, BASES DE DATOS Y SECRETOS EMPRESARIAL.—1. Bases de datos y derechos *sui generis*.—2. La protección de Big Data como secretos empresariales e información confidencial.—VI. CONCLUSIONES.—VI. BIBLIOGRAFÍA.

I. INTRODUCTION

Big data and its use by Artificial Intelligence (AI) is changing the way Intellectual Property¹ is developed and granted. For decades, machines have been autonomously generating works which have traditionally been eligible for copyright and patent protection. Innovations derived from AI, represent a new paradigm of knowledge production in the economy and society, which decisively influences the ability of firms to compete in the market and redefines the limits of knowledge concentration². Given the breadth of the chosen topic³, in this article we will look into the particular applications in big data environment of patents, copyrights⁴, databases and trade secrets legislation.

II. CONCEPT OF ARTIFICIAL INTELLIGENCE

The digital economy is perhaps the main driver of innovation. Data⁵ is the fuel that drives it, including in the field of Artificial Intelligence (AI)⁶, which plays an increasingly prominent role in it⁷. In this regard, it is enough to note that one of the essential pillars of machine learning, one of the most successful types of AI in recent years, is precisely data⁸. Thanks to it, when referring to AI⁹, we allude to a specificity that some algorithms have or may have, this being

¹ Intellectual Property is defined as «intangible property the result of creativity». *Oxford English Dictionary*. Available at: <http://enoxforddictionaries.com> (accessed 17 February 2021). It is very interesting, DRAFT REPORT on Intellectual Property Rights for developments on Artificial Intelligence Technologies [2020/2015 (INI)], 20.04.2020, European Parliament, 2019.2024.

² The WIPO Conversation on Intellectual Property and Artificial Intelligence was established in 2019 to bring together diverse stakeholders in a policy dialogue to discuss the impact of AI on IP. Additionally, WIPO Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence, 21 May 2020. Available at: https://www.wipo.int/meetings/es/doc_details.jsp?doc_id=499504. In addition to the above, European Parliament Resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of Artificial intelligence, robotics and related technologies. Available at: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html.

³ In relation with design and trademark, *vid.* WIPO. Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence, second session, 21 May 2020, Issue 12 and 13. Available at: https://www.wipo.int/edocs/mdocs/mdocs/en/wipo_ip_ai_2_ge_20/wipo_ip_ai_2_ge_20_1_rev.pdf.

⁴ FOSS-SOLBREKK (2021), pp. 4-7; RANTOU (2012) pp. 1-24.

⁵ KUNER, BYGRAVE and DOCKSEY (2020), pp. 1-3.

⁶ EUROPEAN COMMISSION, «A definition of AI: Main capabilities and scientific disciplines», High-Level Expert Group on Artificial Intelligence, 8 April 2019.

⁷ *Economic impacts of Artificial Intelligence*, European Parliament Research Services, 2019. Available at: [https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_BRI\(2019\)637967](https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_BRI(2019)637967).

⁸ GIL GONZÁLEZ (2016), pp. 17-28.

⁹ Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions a European strategy for data, COM (2020) 66 final. In this aspect, it is interesting DREXL and HILTY (2019), p. 1.

a sequence of instructions that produce a result to solve a problem. Nevertheless, we have known and used algorithms for a long time, but when it is capable of learning from its own experiences, and makes its own decisions (self-learning), we are faced with AI¹⁰.

Recognizing this premise, in technical standards, the International Organization for Standardization (ISO) defines AI as an: «interdisciplinary field [...] dealing with models and systems for the performance of functions generally associated with human intelligence, such as reasoning and learning». Most recently, in its January 2018 book, «*The Future: Computed*», Microsoft describes of AI as: «a set of technologies that enable computers to perceive, learn, reason and assist in decision-making to solve problems in ways that are similar to what people do»¹¹. Likewise, the Communication from the Commission to the European Parliament, the European Council, the European Economic and Social Committee and the Committee of the Regions, Artificial Intelligence for Europe, published on April 25, 2018¹², defines AI as «that term that is applied to systems that manifest intelligent behavior, as they are capable of analyzing their environment and taking action- with a certain degree of autonomy- in order to achieve it specifically». Between this organization, there is no universal definition of AI¹³. Applying these concepts, AI is generally considered to be a discipline of computer science that is aimed at developing machines and systems that can carry out tasks considered to require human intelligence. Machine learning and deep learning are two subsets of the same. In recent years, with the development of new neural networks techniques and hardware, AI is usually perceived as a synonym for «deep supervised machine learning»¹⁴.

Is it also notable that AI systems can also improve hardware execution. There is, therefore, a reciprocity between them¹⁵. However, even though artificial intelligence is referred to as AI in the media, there are 3 types of AI: narrow or weak AI, general or strong AI, and artificial superintelligence. For this reason, we have currently only achieved narrow AI¹⁶. Continuing this line of thought, as machine learning capabilities continue to evolve, and scientists get closer to achieving general AI, theories and speculations regarding the future of AI are circulating¹⁷. Notwithstanding the above, there are two main theories: one theory, is based on fear of a dystopian future, where super intelligent killer robots take over the world, either wiping out the human race or enslaving all of human-

¹⁰ GALLEGO SÁNCHEZ (2019), pp. 1-24; FERNÁNDEZ CARBALLO-CALERO (2021), pp. 19-37; XALABARDER PLANTADA (2020), pp. 205-223; GUADAMUZ (2017), p. 1; NAVAS NAVARRO, GÓRRIZ LÓPEZ, CAMACHO CLAVIJO, ROBERT GUILLÉN, CASTELLS I MARQUÉS y MATEO BORGE (2017), pp. 25-35.

¹¹ MICROSOFT. «The Future Computed: Artificial Intelligence and its role in society», Microsoft, January 2018, p. 28. Available at: <https://news.microsoft.com/uploads/2018/01/The-Future-Computed.pdf>. In Spain, «Sistemas de Inteligencia Artificial en la Cloud Pública», produced by «IDC Research España» and «T-System» (2019).

¹² Vid. White Paper on Artificial Intelligence. A European Approach to Excellence and Trust of February 19, 2020 [(COM 2020), 65 final].

¹³ Vid., in Spanish doctrine, FERNÁNDEZ CARBALLO-CALERO (2021) pp. 38-39.

¹⁴ KAPLAN and HAENLEINM (2019), pages. 15-25; PORTELLANO (2020), pp. 8-15.

¹⁵ Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions a European Strategy for Data, COM (2020) 66 final.

¹⁶ Vid. *Artificial Intelligence: How does it work, why does it matter, and what can we do about it?* Panel for the Future of Science and Technology EPRs, European Parliamentary Research Service, 2020. Available: https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641547/EPRS_STU2020641547_EN.pdf.

¹⁷ Vid. <https://codebots.com/artificial-intelligence/the-3-types-of-ai-is-the-third-even-possible>.

ity, as depicted in many science fiction narratives. Thus, the other theory predicts a more optimistic future, where humans and bots work together, humans using artificial intelligence as a tool to enhance their life experience. In addition to the fact that in this sphere, AI tools are already having a significant impact on the way we conduct business worldwide, completing tasks with a speed and efficiency that would not be possible for humans¹⁸. In addition to the above, its operation is completely written in the source code of the program. However, this first stage is not directly executable by the computer. It must be translated into machine language or object code, also called binary code, because it works with lists of «1» and «0» of which there are two classes. The first is made up of lists that serve to give instructions to the machine software, while the second class of lists is made up of the data that is processed by executing the instructions. In this way, for example, a list of «1» and «0» will constitute a text editing program, while another list of «1» and «0» will be a letter written using the previous program. Similarly, a list of «1» and «0» can be an internet browser and the other list of «1» and «0» the page being viewed, etc¹⁹. However, it is necessary to clarify from now on that the same algorithm can be expressed in different ways through a computer code, so it is possible that the same sequence of instructions is translated into several computer programs with similar functionalities and one code different²⁰. On the other hand, it is necessary to take into account the peculiar relationship that exists between the source code and the binary code, since from a source code you can obtain the binary code but from a binary code you cannot obtain the source code except in a few exceptions. Therefore, it is also essential to consider the existence of related sectors, but, does it really work to take into account the existence of related sectors, which, although they are essential for the functioning of AI, are not part of it in the strict sense? First of all, it is worth mentioning the data²¹, because it is essential to computing capacity.

In relation to the above, it is paradoxical that algorithms, which constitute one of the elements that can contribute the most creativity to a Big Data project, as will be explained below, do not have specific legal protection. Thus, algorithms cannot be considered computer programs, since they are not expressed in sequences of instructions in programming language²². The algorithms constitute the sequence ordered operations to be carried out to achieve a specific result or solve a specific problem, and therefore, constitute the ideas or principles on which programs are based associated with Big Data projects²³. These ideas and principles are expressly excluded from being considered as Intellectual Property (IP)²⁴. On the other hand, and subject to various decisions adopted by the European Patent Office²⁵,

¹⁸ GERVAIS (2020), p. 107; DEVINS, TEPPON, KAUFFMAN and KOPPL (2017), p. 306; FERNÁNDEZ CARBALLO-CALERO (2021), pp. 115-124; FERNÁNDEZ MASÍA (2018), pp. 1183-1187.

¹⁹ In Spanish doctrine, GALLEGO SANCHEZ (2019), p. 3.

²⁰ GALLEGO SANCHEZ (2019), pág. 4; MASTRELIA (2018), pp. 460-463.

²¹ Data is the foundation of AI, according to the Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions-Coordinated Plan on Artificial Intelligence of 7 December. Available at: <http://www.ipex.eu/IPEXL-WEB/dossier/files/.../082dbcc5679fb7b40167a1b3f76300c1.do>.

²² KAPLAN and HAENLEIN (2019), pp. 21-22.

²³ SYNODINOU (2020), p. 181; MASTRELIA (2018), p. 462; GUADAMUZ (2017), p. 1.

²⁴ SYNODINOU (2020), pp. 182-183. *Vid. Artificial Intelligence and Robotics and their impact on the workplace*, IBA, Global Employment Institute, April 2017, pp. 102-106.

²⁵ For example, C.1082/14 (*Algorithm specific AKA key modification/ERICSSON*) of 27.6.2017 (ECLI:EP:BA:2017:T108214.20170627). In this aspect, it is interesting EPO. Guidelines for Examination, Mathe-

the algorithms also do not meet the conditions to be protected as a patent. Therefore, the only possible form of protection for algorithms is the one available for secrets, industrial and commercial²⁶, subjecting their knowledge to the obligation of confidentiality and resorting to civil and criminal actions available in our legal system in case of unauthorized access, appropriation or dissemination²⁷.

1. Special reference to patentability of Artificial Intelligence.

Technology is the backbone of the digital economy and much of its value lies in software. Indeed, all economic sectors are becoming reliant on software to leverage growth. «This has important implications for IP laws. Until the late 20th century, the functionality of most innovative products, particularly those relying on semiconductors, was primarily embedded in hardware. There was no doubt about their patentability. But today, increasingly sophisticated semiconductor technology and design tools mean that physical objects are no longer the sole basis of innovation. In other words, technical functionality is progressively migrating from hardware to software. And yet in many jurisdictions software-related inventions either do not qualify for patent protection or have a very limited scope of protection»²⁸. In sum, patentability of artificial intelligence has to be assessed in a double context: mathematical methods and computer-implemented inventions²⁹. In this aspect, we should remember that a computer implemented invention is one that involves the uses of a computer, a computer network or other programmable device in which one or more of its functions are carried out totally or partially thanks to a computer program. It should be noted that the European Patent Convention (EPC) [articles 2 (c) and 3] state that a computer program claimed «as such» is excluded from patentability³⁰. Accordingly, the technical character is essential for it to be considered an invention in the sense established by art. 4.1 Spanish Patent Law and art. 52 EPC 1973 [and also by article 52(2) EPC 2000]³¹.

Additionally, the object for which protection is sought must include technical teaching, that is an instruction addressed to a technical expert on how to solve a particular technical problem through the use of particular technical means. In relation with the above, the cited precepts exclude the consideration of invention in both mathematical methods and computer programs, which constitute, as we have seen, the two supports of AI³². Of course, the difficulty is knowing what is meant by «technical character»³³. Since there is no legal defi-

mathematical methods, March 2021 (available at: https://www.epo.org/law-practice/legal-texts/html/guidelines/eg_ii_3_3.htm).

²⁶ MASTRELIA (2018), p. 463.

²⁷ YAMAMOTO and LLOYD (2019), pp. 127-132.

²⁸ Vid. https://www.wipo.int/wipo_magazine/en/2017/01/article_0002.html (accessed 25 February 2021).

²⁹ RAMALHO (2018), pp. 1-32; HANEY (2020), pp. 410-420.

³⁰ Vid. Rules 42 and 43 EPC.

³¹ In relation with this aspect, *vid.* article 69 EPC. In the United States, patent protection for software-related inventions is limited to those on recordable media, not to computer programs themselves (*vid. In re Beauregard*, 53 F.3d 1583 (Federal Circuit 1995)). This protection falls short when it comes to the online distribution of software. But the Supreme Court's decision in *Alice Corp. vs. CLS Bank Int'l* (134 S. Ct. 2347, 2014) and some subsequent cases, have failed to provide clear boundaries for the patent eligibility of software-related inventions. Available at: <https://www.clfip.com/ip/blog/how-to-patent-software/> (accessed 23 February 2021).

³² GALLEGO SÁNCHEZ (2019), pp. 3-4; FERNÁNDEZ CARBALLO-CALERO (2021), p. 21.

³³ SPTO: www.oepm.es/export/sites/oepm/comun/documentosrelacionados/invenciones/DirExPat_DI-RECTRICES_Version_2_0.pdf.

tion, its scope should be defined according to the groups of cases that exist in practice. Firstly, together with the previous ideas, the technical purpose must be «specific»³⁴ and the model that serves a technical purpose is determined by the direct technical relevance of the results it provides³⁵. Secondly, this technical characteristic may derive from the fact that the computational model and the algorithms on which it is based are designed and adapted for a *specific technical implementation* or that their design is motivated by technical considerations on the internal functioning of the computer³⁶. For this reason, the technical character has to redirect in the *additional technical effects* derived from the execution (by hardware) of the instructions given by the computer program. Consequently, additional technical assumptions that confer technical character to computer programs are the control of a technical process or the internal functioning of the computer or its interfaces³⁷. Proportionality in this context, once the abandonment of the so-called «contribution approach» that required determining the technical contribution achieved by an invention with respect to the state of the art in order to appreciate the technical character, the presence of an additional effect, and therefore, the technical character, it was valued without reference to the state of the art³⁸. From all this, two questions arise: firstly, the determination of whether or not a computer program is excluded from patentability according to article 52.2 and 3 EPC³⁹ is independent of the «state of the art». This aspect can only be considered when examining the concurrence of the patentability requirements specific to novelty and inventive step. Secondly, which aspects can be considered novel if they lack technical character and do not justify by themselves the presence of an additional technical effect. Above all, the jurisprudence is ambiguous⁴⁰. Only describing «what» the features are or «what» the user experiences on screen are most likely to be deemed as an unpatentable abstract idea.

III. DEFINING BIG DATA

The analysis of the information obtained in the network allows to have a greater knowledge of the essential issues of a business and direct resources in an efficient and profitable forms. So much so, that the knowledge acquired from data analysis has been decisive in decision —making and in the design of the strategy of many of the most sophisticated operators—. Data is the oxygen in the information ecosystem and the currency in the information economy. In our

³⁴ EPO.Decision Board of Appeal, T 1358/09 (*Classification/BDGB Software empresarial*, 21.11.2014 (ECLI:EP:BA:2014:T135809.20141121)); EPO. Decision Board of Appeal, T1316/09,18.12.2012 (ECLI:EP:BA:2012:T131609.20121218).

³⁵ GALLEGO SÁNCHEZ (2019), p. 11.

³⁶ EPO. Decision G0003/08,12.05.2010 (ECLI:EP:BA:2010:G000308.20100512) and EPO. Decision Board of Appeal T1173/97, 17.7.1998 (ECLI: EP: BA:1998:T117397.19980701). In this Decision, it is possible to grant a patent not only in the case of an invention in which a piece of software manages.

³⁷ EPO.Guidelines for Examination, G-II, 3.6.1. Available at: https://www.epo.org/law-practice/legislation/texts/html/guidelines/eg_ii_3_6_1.htm (accessed 19 January 2021).

³⁸ DRAFT 202003, *European Patent Office*, Munich, pp. 8-24.

³⁹ The legal framework for patent in Europe, excludes mathematical methods and computer programs (defined under the EPC as claims referring to «a sequence of computer-executable instructions» that specify «a methods rather than the methods itself) from patentability». *Vid.* President's Reference/Patentability of Programs Computers, G 3/08 (12 May 2010) 11.2.5; T 494/07 (Currency Validator) of 13.05.2011.

⁴⁰ EPO.Decision Board of Appeal, T424/03; 23.02.2006 (ECLI:EP:BA:2006:T042403.20060223); T0258/03.12.5.2010(ECLI:EP:BA:2010:G000308.20100512);T1370/2011;11.03.2016,(ECLI:EP:BA:2016).

increasingly data-centered society, databases serve multiple purposes⁴¹. They are treasurable repositories of information and valuable source of knowledge, along with being precious information tools (tools enabling us to search for, process and retrieve information)⁴². Consequentially, databases offer a harmonious and user friendly synthesis of a main body content (works, data or other materials) with a variety of tools and other means of arranging, presenting and searching their *corpus* of contents⁴³.

In these aspects, the previous tools to analyze and process large amounts of data and the ability to interpret them constitute the basis of the business and the competitive advantage of numerous companies, and especially of a good part of startups that feed on the knowledge acquired by their tools to offer innovative services and solutions with great added value⁴⁴. So, the very importance of these intangible assets makes their protection a crucial element for the success and consolidation of these businesses. We cannot lose sight of the fact that the disclosure of an algorithm or free access to the knowledge stored in the databases generated, can put a business that is based on the exploitation of these intangibles at risk. With the development of technologies capable of generating high data volumes at a rapid pace, a global effort has been made to develop solutions for better storage, processing and analysis of Big Data⁴⁵. As our society grows more complex, interconnected, and technologically advanced, data is generated that reflects this societal change, and potentially allows us to better understand this complexity⁴⁶. Thus, with technology, every individual's movements, decisions, and purchases are captured and memorialized in the electronic realm. Due to the exponentially increasing efficiency of storage, the data is collected in centralized servers, stored, and analyzed in ways never before possible⁴⁷.

In general, Big Data and its use by AI is changing the way IP is developed and granted and is data that exceeds the processing capacity of conventional database systems. It is increasingly available in all areas of manufacturing and operations which presents an opportunity for better decision making and discovery of the next generation of innovative technologies and can be defined in a number of ways⁴⁸. We also understand that a common way to define it is to enumerate three essential features, a fourth that, though not essential, is increasingly typical and a fifth that is derived from other three (or four). Those features are: volume, veracity, velocity, variety and value. Thereby, «volume» or size is, as the term Big Data suggests the first characteristics that distinguishes, from other («small data») datasets. For instance, Big Data corpora are often generated automatically and the question of the quality or trustworthiness of the data («veracity») is essential. In relation to the above, «velocity» refers to «the speed at which corpora of data are being generated, collected and analyzed». Complementing what is indicated, the term «variety» denotes the many types of data and data sources

⁴¹ YAMAMOTO and LLOYD (2019), pp. 126-133.

⁴² DEBUSSCHE and JASMIEN (2019), p. 1.

⁴³ SYNODINOU (2020), pp. 181-187; MCAFEE and BRYNJOLFSSON (2012), pp. 61-67.

⁴⁴ MARTÍNEZ PÉREZ (2020), pp. 400-413; PUYOL MORENO (2014), pp. 471-505.

⁴⁵ DEVINS, TEPPA, KAUFFMAN and KOPPL (2017), p. 306.

⁴⁶ *Vid.* <https://www.scientificamerican.com/article/big-data-needs-big-theory/> (accessed 11 February 2021).

⁴⁷ IBM (2012). Institute for Business Value in collaboration with the Business School Said, Oxford University, «Analytics: Big data and real world», *IBM Global Business Services*.

⁴⁸ MCAFEE and BRYNJOLFSSON (2012), pp. 64-65.

from which data can be collected, including «Internet Browsers», social media sites and apps, cameras, cars⁴⁹ and host of other data-collection tools. Finally, if all previous features are present, a Big Data corpus likely has significant «value»⁵⁰. In addition, the way in which «Big Data» is generated and used can be separated into two phases: firstly, the creation of a Big Data requires processes to collect data from sources such as those mentioned in the previous paragraph. Secondly, the *corpus* is analyzed, a process that may involve Text and Data Mining (TDM)⁵¹ — that is a form of research that extracts data and patterns from large datasets⁵², and identifies word frequencies, syntactic patterns and thematic markers in a corpus of documents —⁵³. It allows that machine to learn from the corpus hence the term «machine learning» (ML) is sometimes used as synonym of AI⁵⁴.

In view of the above, «this process often requires human input to assist the machine in correcting errors or faulty correlations derived from, or decisions based on, the data. Data as such presents value for enabling a competitive data-driven economy, which is at the heart of «Internet of Things» (IoT)⁵⁵ and Industry 4.0⁵⁶. This processing of corpora of Big Data is done to find correlations and generate predictions or other valuable analytical outcomes. Thus, «these correlations and insights can be used for multiple purposes, including advertising targeting and surveillance, thorough an almost endless array of other applications is possible. Attached to the above, a subset of machine learning know a deep learning (DL) is an AI function that imitates the workings of the human brain in processing data and creating patterns for use in decision making. Also known as deep neural

⁴⁹ The list includes «cards» as car a personal vehicles are one of the main sources of (personal) data-up to Gigabytes per hour of driving. The data are fed back to the manufacturer. Available at: <http://www.adac.de/infostrat/> (accessed 25 January 2021).

⁵⁰ *Vid.* <https://www.bigdataframework.org/four-vs-of-big-data/>. In line with this concepts, it is interesting FAVARETTO, DECLERCQ and LEGER (2019), pp. 2-27.

⁵¹ If TDM falls under the reproduction right, which arguably exception to the exclusive reproduction rights must be added. Following some Member States (including UK). The European Commission proposal of a Directive on Copyright in The Digital Single market (hereafter 'draft DSM Directive') of 14 September 2016 contained a TDM exception. In relation, the Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market-Legal Policy Department for Citizens' Rights and Constitutional Affairs, Directorate General for Internal Policies of the Union, PE 604.941-February 2018, and Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on Copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC (article 3 and 4.3). Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32019L0790>. Doctrinally, ORDELIN FONT (2019), pp. 209-229; GARCÍA VIDAL (2020), pp. 99-124.

⁵² In USA, the fair use provision can function as a way to delineate the boundaries of the author's exclusive right with a view to its objectives. In the Google Book case, the US Court of Appeals for the Second Circuit decided that the various acts committed by Google in reproducing books, including for searching and TDM, are to be considered as fair use. Available: <http://law.justia.com/cases/federal/appellate-courts/ca2/13-4829/2015-16.html>, acceded 16 January 2021 (On 18 April 2016, The Supreme Court denied the petition for a writ of certiorari, leaving the Second Circuit ruling in Google's favour intact). To make available parts of the *corpus* of book, Google has scanned the digital copies and established a publicly available search function, the programs tool. *Vid.* GEIGER, FROSIO and BULAYENKO (2019), pp. 28-63.

⁵³ MASTRELIA (2018), p. 464.

⁵⁴ FAVARETTO, DECLERCQ and LEGER (2019), p. 23.

⁵⁵ BARRIO ANDRÉS (2018), pp. 18-22.

⁵⁶ Recalled, Industry 5.0. *Vid.* *Industry 5.0. Towards a sustainable, human-centric and resilient European Industry European*. Commission Directorate-General for Research and Innovation Directorate F-Prosperity, January 2021. It is interesting, Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions Fostering a European approach to Artificial Intelligence, COM (2021) 205 final. In addition, «The AI Index 2019 annual Report», AI Index Steering Committee, Human Centered AI Institute, Stanford University, Stanford CA, December 2019.

learning or deep neural network»⁵⁷. Consequently, «both artificial and biological neurons receive input from various sources, mapping information to a single output value»⁵⁸, this implies that any human contribution to the output of deep learning system is «second degree», the part of the system which analyzes the information. Therefore, Big Data is driving a trend towards behavioral optimization and «personalized law», in which legal decisions and rules are optimized for best outcomes and where law is tailored to individual consumers based on analysis of past data»⁵⁹. It is supposed «objectivity and predictive power are overstated, at least when applied to highly complex evolutionary systems such as the legal system»⁶⁰. Alternatively, while law is semantic, Big Data is syntactic. Law is abstract, values-based, and built on compromise. Big Data is empirical, algorithmic, and deterministic. Also, «Big Data is inherently a contextual and cannot interpret itself, nor can it discern the indeterminate boundaries of legal principles»⁶¹. Moreover, «Big Data cannot discern or create novelty, unlike humans, who can update their «frames» or paradigms as their environment changes. Consequently, cannot innovate beyond the paradigm imposed by its creators. Even the most sophisticated machine learning techniques cannot tell us what factors might become relevant in response to new challenges»⁶². However, Big Data has serious limitations and poses dangers when applied in the legal context, especially those of ownership in copyright, patents, database and trade secrets regulation.

IV. INTERFACE BETWEEN INTELLECTUAL PROPERTY LAW AND BIG DATA

The connection between Big Data and law can be thematized in several ways. The development of big data invites a reflection on the impact that this and other advanced technologies are having in the field of IP rights⁶³. The interfaces between Big Data (as the term is defined below) and IP matters both because of the impact of IP rights in Big Data, and because IP rights might interfere with the generation, analysis and use of data⁶⁴. In fact, information is the currency of the digital age, and data about customers is becoming a critical asset in this markets. Emerging, fast-growing business models are increasingly counting on the availability of massive amounts of data about customers and their behavioural patterns in order to collect and monetize them. A significant portion of this information is personal data relating to identifiable individuals. In this case, the exploitation of those data is subject to privacy laws. In this way, «EU privacy laws allow the collection and use of customers' data for commercial purposes. However, these activities are highly invasive, and companies must respect a series of strict rules which often imply high costs»⁶⁵. For this rea-

⁵⁷ HANEY (2020), pp. 407-412.

⁵⁸ *Vid.* USA Patent núm. 9471884 (assigned to IBM).

⁵⁹ DEVINS, FELIN, KAUFFMAN and KOPPL (2017), pp. 306-307.

⁶⁰ HANEY (2020), p. 410.

⁶¹ DEVINS, FELIN, KAUFFMAN and KOPPL (2017), pp. 306-307.

⁶² *Ibid.*, pág 360; KAPLAN and HAENLEINM (2018), p. 25.

⁶³ In this aspect, GURRY, Francis, General Director of the World Intellectual Property Organization (WIPO), it should be noted that «data-driven technology is the dominant force in economic production and distribution in the digital framework». Available: https://www.wipo.int/wipo_magazine/es/2019/05/article_0001.html (accessed 21 December 2020).

⁶⁴ YAMAMOTO and LLOYD (2019), pp. 127-129.

⁶⁵ *Vid.* https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en.

son, when processing personal data for commercial purposes, companies create sets of customer data which may differ depending on the level of complexity of the processing activity, from newsletter groups to advanced profiling and behavioral advertising programs⁶⁶. Undoubtedly, this information is an intangible asset having economic value whose creation requires economic investments. Thus, it may be the subject matter of IPRs (particularly, IPRs that are designed to protect data within the EU IP framework are copyrights, trade secrets and the database right)⁶⁷.

Consequently, «Big Data *corpora* are protected by secret a form of protection that relies on trade secret law combined with technological protection from hacking, and contracts»⁶⁸. Additionally, deciding which IP rights may apply should thus distinguish Big Data *corpora* that are not publicly accessible and those that are⁶⁹. So, a publicly available *corpus*, in contrast, must rely on *erga omnes* IP protection, if it deserves protection to begin with. Then, copyright protects collections of data⁷⁰, the *sui generis* database right⁷¹ (in the EU, is the concept of «substantial investment»)⁷² might apply and data exclusivity rights may be relevant. The outputs of the processing of Big Data *corpora* may contain or consist of subject matter that facially could be protected by copyright or patent law. Big Data technology can be-and in fact is-used to create and invent. Deciding whether machine-created material should be protected by copyright could thus have a profound impact on the market for creative works. If machine created material is copyright-free, machines will produce free goods that compete with paid ones, that is, those created by humans expecting a financial return. We will have to ask ourselves, if the material produced by machines is protected by copyright and its use potentially subject to payment, this might level the commercial playing field between human and machine, but then who (which natural or legal person) should be paid for the computer's work? Then, there will be border definition issues⁷³.

Overall, important legal challenges are then posed in this matter, especially in relation to IPRs, in particular, with respect to the protection of works generated by AI⁷⁴, due to the great autonomy that this technology is acquiring

⁶⁶ Vid. EASA, *Best Practice Recommendation on Online Behavioral Advertising*. Available at: <https://www.easa-alliance.org/issues/oba>.

⁶⁷ MARCO, CARLEY, JACKSON and MYERS (2015), «The USPTO Historical Patent Data Files», USPTO, Working paper, núm. 2015-1, *Office of Chief Economist U.S. Patent and Trademark Office*, págs. 1-57. Available at: https://www.uspto.gov/sites/default/files/documents/USPTO_economic_WP_2015-01_v2.pdf (accessed 16 January 2021).

⁶⁸ GERVAIS (2012), pág. 124.

⁶⁹ GÓMEZ SEGADÉ (2020), págs. 116-145.

⁷⁰ Article 2 of the Berne Convention only affords protection to «literary and artistic works», making no mention of computer program. The TRIPs Agreements assimilated the Berne articles to accommodate for technological developments [art. 10 (1)]. Thus, article 4 of WIPO Copyright Treaty reaffirms that computer programs fall within the ambit of the Berne Convention. The EU has accepted TRIPs, and in 2001 adopted Infosoc Directive (Directive 2001/29/EC of the European Parliament and the Council of 22 May 2001).

⁷¹ Vid. article 7(2) and articles 9-15 Directive Databases. Therefore, case of *Directmedia Publishing GmbH vs. Albert Ludwigs Univesitat Freiburg*, C-304/07 (ECLI:EU:C:2008:552).

⁷² The CJUE clarified the concept of «substantial» and «insubstantial» parts of database's contents. Vid., C-604/10 *Football Dataco and others* (ECLI:EU:C:2002:115). In the UK, Court of Appeal, *Decision in Football Data Co Ltd vs. Sportradar GmbH*, EWCA, Civ. 27, 2013, FSR, 30.

⁷³ DE MIGUEL ASENSIO (2015), págs. 461-463; SÁIZ GARCÍA (2019), p. 8.

⁷⁴ Vid. foodnote 3. Also, European Parliament Resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for Artificial Intelligence [2020/2014 (INL)]; IViR, Study for European Commission on AI and IP Trends and Developments in Artificial Intelligence Communications Networks,

and in which AI models strongly allows you to act and create independently of humans⁷⁵. In this form, we can ask ourselves different questions: can we apply the notion of joint authorship?⁷⁶ Or, should we consider the machine-produced portion (if separable) copyright-free, thus limiting the protection to identifiably human-authored portions? If such major doctrinal challenges—each with embedded layers of normative inquiries—emerge in the field of copyright, Big Data poses existential threats in the case of patents⁷⁷. Thus, AI tools «can be used to process thousands of published patents and patent applications and used to expand the scope of claims in patent applications. This poses normative challenges that parallel those enunciated above: who is the inventor? Is there a justification to grant an exclusive right to a machine-made invention? To whom?»⁷⁸. Then, there are doctrinal ones as well. For example, is the machine-generated «invention» disclosed in such a way that would warrant the issuance of a patent? It gets more complicated, however. In this way, if AI machines using patent-related Big Data can broaden claim scope or add claims in patent applications, then within a short horizon they could be able to predict the next incremental steps in a given field of activity by analysing innovation trajectories. Doctrinally, this raises questions about inventive step: if a future development is obvious to a machine, is it obvious for purposes of patent law? So, answering this question, poses an epistemological as well as a doctrinal challenge for European Patent Office⁷⁹. This use of patent and technological Big Data could lead to a future where machines pre-disclose incremental innovations (and their use) in such a way that they constitute publicly available prior art and thus make obtaining patents impossible on a significant part of the current patentability universe⁸⁰. Perhaps even the best AI system using a Big Data *corpus* of all published patents and technical literature will not be able to predict the next pioneer invention, but very few patents are granted on ground-breaking advances. So far, AI systems that can predict most currently patented inventions (which tend to be only incrementally different from the prior art) would wreak havoc with the patent-based incentive system⁸¹.

Content and Technology Final report Prepared by: Challenges to the Intellectual Property Rights Framework, Final Report, 20 September 2020 (available at: <https://www.ivir.nl/ivir-study-for-european-commission-on-ai-and-ip/>). In Spanish doctrine, FERNÁNDEZ CARBALLO-CALERO (2021), pp. 115-134.

⁷⁵ SANJUAN RODRÍGUEZ (2020), pp. 1-39.

⁷⁶ In the United Kingdom article 9.1 of the Copyright, Design and Patent Acts 1988. In the USA, the article 17 US Code (Copyright Act 1976 and Compendium of *Practices*, third edition, of 22 December 2014, updated 29 September 2017 (USCO Compendium) and article 15.3 of the Spanish Patent Law 25/2015.

⁷⁷ Vid. <https://www.garrigues.com/es> (accessed 20 January 2021).

⁷⁸ WIPO.Translate and other AI tool in WIPO, December 6, 2019. Available at: https://www.wipo.int/meetings/en/doc_details.jsp?doc_id=462691. Doctrinally, MASTRELIA (2018) págs. 181-184.

⁷⁹ Vid. <https://www.epo.org/searching-for-patents/data/bulk-data-sets.html>. It is interesting WHITE PAPER on Artificial Intelligence. A European approach to excellence and trust (COM/2020) 65 final (19.02.2020); the European Parliament. Framework of ethical aspects of Artificial Intelligence, robotics and related technologies, 20 October 2020 [available: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html]. So, *Digital services legislation and fundamental rights issues*, Legislative Observatory, European Parliament (2020-2022), INI (available: <https://oeil.secure.europarl.europa.eu/oeil/popups/ficheprocedure.do?lang=fr&reference=2020/2022>).

⁸⁰ HANEY (2020), pp. 411-479.

⁸¹ FOSS-SOLBREKK (2021) pp. 7-10.

V. RELATION WITH ANOTHER SYSTEM PROTECTION: COPYRIGHT DATABASES⁸² AND TRADE SECRETS

In principle, copyrights law does not confer any control of data: by application of the idea-expression dichotomy, copyrights only applies to an original form of expression, not to the ideas or simple information embedded in a creative work⁸³. Following this approach, copyright protects the form or expression of information but not the underlying information itself. It applies to software, certain databases⁸⁴, literary works, music, films, videos and broadcasts. It arises automatically by operation of law in the EU (so does not require to be registered). Ensures protection of various types of works, awarding protection to individual data as long as they are original⁸⁵ and can be expressed in a material, concrete form. The broad understanding of these protection requirements facilitates extending, in principle, protection to different types of works, including to data. It is however worth examining some of the most important characteristics of the EU copyright system in order to determine whether it may apply to (big) data. So, «Big Data *corpora* is considered a literary work eligible for copyright protection, subject to possible exclusions and limitations»⁸⁶. In the context of Big Data projects, it is crucial to understand to what extent the data used can be copyright protected⁸⁷. In all likelihood, most of the data collected and processed in a big data analytics (that provide the big picture for understanding patented technologies in a company or industry and help executives of a company make business decisions on how to build their patent portfolio and licensing programs)⁸⁸ context will not be considered original and will therefore not benefit from copyright protection. Likewise, in the context of data⁸⁹, traditional literary copyright will subsist in documentation. Having said that, «it cannot be excluded that the individual data can gain originality once they are connected with other information or presented in an original way (by means of different possible forms of expression)»⁹⁰. The protection of computer programs through copyright constitutes the traditional and ordinary form of protection. It is based

⁸² Study in support of the evaluation of Directive 96/9/EC on the legal protection of Databases, Final report, European Commission (2021). This study aimed at collecting evidence on the impacts of the Database Directive in support of the Evaluation Report on the Database Directive, published by the European Commission on the 25th of April as part of the third data package (available at: <https://digital-strategy.ec.europa.eu/>).

⁸³ SYNODINOU (2020), pp. 185-186; YAMAMOTO and LLOYD (2019), pp. 185-186. In Spanish doctrine, FERNÁNDEZ CARBALLO-CALERO (2021), pp. 176-178; DUQUE LIZARRALDE (2020), pp. 13-67.

⁸⁴ *Vid.* article 2 (5) Berne Convention. In the jurisprudence, CJUE: case *Apis-Hristovich EOOD vs. Lakod SD*, C-545/07 (ECLI:EU:C:2009:132).

⁸⁵ Case *Cofemel-Sociedade de Vestuário, S.A. vs. G-Star Raw CV*, C-683/17JEU in its recent judgment of September 12, 2019 (*vid.*, in this sense, the judgments of December 1 of 2011, *Painer vs. Standard Verlags GmbH and others*, C-145/10, (ECLI:EU:C:2011:798), paragraphs 88, 89 and 94, and of August 7, 2018, *Renckhoff*, C-161/17 (ECLI:EU:C:2018:634). In Spanish doctrine, SAIZ (2019), p. 19.

⁸⁶ YAMAMOTO and LLOYD (2019), pp. 126-136.

⁸⁷ In US Law, Copyright Office. Compendium of U.S. Copyrights Offices Practices, paragraph 306 (Third edition, 2014).

⁸⁸ *Vid.*, for example, <https://www.iptytics.com/generallbig-data-ip-analytics> (accessed 8 December 2020).

⁸⁹ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). It is interesting, in this point, European Parliament, «A comprehensive European Industrial policy on Artificial Intelligence and Robotics», 12 February 2019.

⁹⁰ Article 12 Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019, on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC.

on the claim that the software is a mere intellectual creation, not the kind of technical inventions, which, insofar as it is written in code, is similar to a literary work. Thus, article 1 of Directive 24/EC of the European Parliament and of the Council, of April 23, 2009, on the legal protection of computer programs is expressly stated when it states that: «Member States shall protect the computer programs as literary works as defined in the Berne Convention for the protection of literary and artistic works»⁹¹.

In fact, at the regional level in European Union, the Agreement on Trade-related Aspects of Intellectual Property Rights, adopted within the World Trade Organization in 1994 (hereinafter, TRIPs)⁹², whose article 10.1 states that «computer programs, whether source programs or target programs, will be protected as literary works by virtue of the Berne Convention» and the Treaty of the World Intellectual Property Organization (hereinafter, WIPO) on Copyright adopted in Geneva in 1996, article 4 of which provides that «computer programs are protected as literary works within the framework of the provisions of article 2 of the Berne Convention»⁹³. Among the advantages cited in favor of this form of protection are its flexibility and its low cost is applied to computer programs. In addition, when the Berne Convention text was last revised on substance in 1967 neither publicly available «electronic» databases nor any mass-market database software was available. In this regards, the «collections» referred to in the Convention are thus of the type mentioned by the Convention drafters: (paper-based) anthologies and encyclopedias. Additionally, the negotiators' objective was to create a separate copyright for the maker (or «arranger») of a collection, knowing that most if not all of the entries in the collection (say, an encyclopedia) were written by third parties, each an expert in her or his own field and each entitled to his or her own copyright in the entry. In a collection of this type, there are thus two layers of copyright: first, a right in each entry, and in each illustration or photograph, which is either transferred or licensed to the maker or publisher of the collection; and, second, a copyright in what one might call the «organizational layer», granted to the maker of the collection based on the «selection or arrangement» of the individual entries, photographs and illustrations. The second layer-the collection such as encyclopedia —is generally treated as a collective work —⁹⁴.

A Committee of Experts meeting under the auspices of the WIPO, which administers the Berne Convention⁹⁵, concluded that, although this is not specified expressly in the text of the Convention, the only mandatory requirement for a literary or artistic work to be protected by the Convention is that it must be «original». To arrive at this conclusion, the Committee considered both the Con-

⁹¹ GEIGER, FROSIO and BULAYENKO (2019), pp. 28-63.

⁹² The EU has accepted TRIPs, and in 2001 adopted Infosoc Directive (Directive 2001/29/EC of the European Parliament and of The Council of 22 May 2001 on the Harmonization of Certain Aspect of Copyrights and related Rights in the Information Society) to *inter alia* implement the WCT into its legal order. In 2009, the EU codified the Software Directive (Directive 2009/91/250/EEC of 14 May 1991, on the Legal protection of Computer Programs). In this aspect, it is interesting, CJUE, Case C-128/11 *UsedSoft* (ECLI:EU:C:2012:407). In reference with copyright protection under EU Software Directive, that is interesting, CJUE Case C-393/09 *Bezpečnostní Sofwarová Asociace* (ECLI:EU:C:2010:816). *Vid.* FOSS-SOLBREKK (2021), p. 3. Doctrinally, DAVISON (2020), pp. 59-65.

⁹³ In Spanish doctrine, BERCOVITZ RODRÍGUEZ-CANO (2013), pp. 3-9.

⁹⁴ GEIGER (2007), p. 487.

⁹⁵ Committee of Experts on a Possible Protocol to the Berne Convention (available at: https://www.wipo.int/meetings/en/details.jsp?meeting_id=2251). *Vid.* RICHETSON and GISNSBURG (2006), pp. 45-49.

vention's drafting history and the use of the expression «intellectual creation» in the Convention as a functional synonym of the term «work». This also means that no mandatory formality may be required to obtain copyright protection. The same statement, namely that the only applicable criterion is originality, can be made about EU law. As I anticipate what I will comment later, when «electronic» databases started to emerge in the 1990s, data generally had to be indexed and re-indexed regularly to be useable. In this line, using language meant to parallel article 2(5) of the Berne Convention⁹⁶, the data in typical (relational or «SQL») databases in existence in the 1990s generally was «structured» in some way, for example, via an index, and that structure might qualify the database for (thin) copyright protection in the database's organizational layer. Older databases also contained more limited datasets («small data»). Indeed, Big Data is sometimes defined in direct contrast to the notion of SQL database and reflected in the TRIPs Agreement⁹⁷ (and the EU database Directive discussed in the next section)⁹⁸. Big Data software is unlikely to «select or arrange» the data in a way that would meet the originality criterion and trigger copyright protection. In this context, «fixation», in a data context, would mean that the specific information needs to be saved in a tangible form⁹⁹. Thus «the form of saving the data can differ from handwritten notes (files), through photographic documentation (image) or recorded testimonies (sound) to digitalized archives (digital files), as long as it remains concrete, can be easily identified and described. Results that have not yet been produced (future data), or results that cannot yet be described (e.g. because there are no means yet to express them) cannot benefit from copyright protection for as long as they have not materialized»¹⁰⁰. This can present some difficulties in a Big Data context, given that Big Data tends to involve dynamic datasets and notably relies on cloud computing services. And the legal framework for copyright does not provide for a registration system. Accordingly, the eligibility for protection (and its scope) can only be confirmed *a posteriori* by a Court of Justice, leading to a lack of legal certainty in the meantime.

In the light of the foregoing, the copyright holder is granted several exclusive economic rights that allow controlling the protected work's use and facilitate enforcement in case a third party uses the work without authorization. Moreover, «the rights of reproduction, communication to the public and distribution are indeed a useful toolkit which, balanced by the copyright exceptions, allows for an optimal protection of right holder's interests»¹⁰¹. That is, copyright law therefore provides for a wide scope of measures securing the rights of the

⁹⁶ RICHETSON and GISNSBURG (2006), p. 47.

⁹⁷ TRIPs established global standards for Copyrights and Patent protection. The UK and all EU member states are required to adhere to the mandatory requirements in TRIPs. These requirements were modelled after the IP laws in developed nations such as the United Kingdom, the United States and Japan, so TRIPs required relatively few changes to the UK's IP laws when it came into effect on 1 January 1996 (94/800/EC Council Decision on 22 December 1994). *Vid.* further, GERVAIS (2012), pp. 123-125; RICHETSON and GISNSBURG (2006), pp. 47-48.

⁹⁸ Directive 2009/24/EC of the European Parliament and the Council of 23 April 2009 on the legal protection of computer programs, OJ L111, 5 February 2009; Directive 2006/116/EC of the European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights OJL372, 27 December 2006; Directive 96/9/EC of the European Parliament and the Council of 11 of March 1996 on the legal protection of database, OJ L77, 27 March 1996.

⁹⁹ FOSS-SOLBREKK (2021), pp. 4-6.

¹⁰⁰ RICHETSON and GISNSBURG (2006), pp.47. In Spanish doctrine, FERNÁNDEZ CARBALLO-CALERO (2021), pp. 27-29.

¹⁰¹ FOSS-SOLBREKK (2021), pp. 4-7.

author in case of dissemination of his work and the use of these works by third parties. The rules governing copyright protection aim at enabling further use of the works, securing at the same time the legitimate interests of the author. In this aspect, in a data environment, the most important hindrance resulting from copyright protection is the necessity to obtain authorization from the copyright holder of each individual data¹⁰². In the context of Big Data projects, to the extent copyright applies, it would require identifying authors of hundreds (if not hundreds of thousands) of works. In many cases, it might be difficult to identify or find the right holder and/or understand whether he has given his authorization for use of the work¹⁰³. In practice, this means that time-consuming analyses need to be performed before the data gathered can be used. Furthermore, as regards the possibility to acquire copyright in data, the exclusivity of this type of right constitutes a hindrance, since it does not allow acquiring copyright in the same data «in parallel»¹⁰⁴. Thus, the latter may however often be the case in a Big Data context, in particular where parties will be independently collecting the same or similar data, leading to the creation of convergent datasets. Consequently, regard may be had a wide range of factors when determining the exclusive economic rights, when authors are also granted so-called «moral rights»¹⁰⁵, that are related to the idea that a work is not a mere staple commercial object, but also the expression of their personality¹⁰⁶.

1. Databases *sui generis* right

Databases are generally perceived as static warehouses, storing up valuable facts and information. With the advent of e-commerce, entire businesses are being built upon databases and generating revenue streams from subscriber fees, royalties, or advertising. In order to recover the initial investment of time, money and skill put into it and to avoid parasitic competition, database manufacturers protect their compilation efforts¹⁰⁷. Therefore, without the ability to restrict access to and use of databases that it compiles, a company is effectively discouraged from participating in the information age¹⁰⁸. Although traditionally, databases have been protected under copyright laws¹⁰⁹, advances in information technology have enabled potential competitors and pirates to engage in market-destructive copying. Further, in many jurisdictions, the required level of originality has been in a state of flux in recent years, particularly in the US and Europe, after the Feist decision¹¹⁰. When considering such protection, a dis-

¹⁰² GEIGER (2007), pp. 489-491.

¹⁰³ YAMAMOTO and LLOYD (2019), pp. 127-133.

¹⁰⁴ DEBUSSCHE and JASMIEN (2019), p. 1.

¹⁰⁵ *Ibid.*

¹⁰⁶ CJUE, Case C-161/17, *Renckhoff* (ECLI:EU:C:2018:634). In relation with creative choices, Case C-145/10, *Painer* (ECLI:EU:C:2011:788). Analyzing sequences and combinations wherein authors originally express themselves, CJUE Case C-5/08, *Infopaq International A/S vs. Danske Dagblades Forening* (ECLI:EU:C:2009:465).

¹⁰⁷ SYNODINOU (2020), pp. 181-183.

¹⁰⁸ FOSS-SOLBREKK (2021), p. 6.

¹⁰⁹ Recently, CJUE Case C-833/18 *Brompton Bicycl* (ECLI:EU:C:2020:461); case *Cofemel* (ECLI:EU:C:2019:721).

¹¹⁰ In relation to the intervention of a human, being in the creation of a work for it to be eligible for copyright work for it to be protected by copyright, the US Copyright Office («USCO») as a requirement for registering a work for the purposes of copyright protection, such as for the purposes of copyright protection, as set out in section 306 of its as stated in paragraph 306 of its Compendium of Practices, Third Edition, 22 December

inction needs to be made between, on the one hand, the database's¹¹¹ contents (individual data), and, on the other hand, its structure and the investment made in its creation. While the general rules governing the protection of database are established at international level¹¹², EU law provides for a specific protection of databases which goes beyond other international legal instruments. In such respect, the EU institutions adopted the Database Directive with the objective of harmonizing the protection of databases in all Member States¹¹³. Similarly to copyright, the level of protection ensured across the Member States, especially, concerning the copyright on databases, is significantly different¹¹⁴. This particularly hinders the possibility to manage pan-european projects, since it implies the necessity to examine multiple national legislations in order to have clearance on the possibility to use data, or secure the investment made in a database containing data originating from different territories. The protection established by the Database Directive¹¹⁵ is dual, and supplements the possible protection granted to the data as such¹¹⁶. More specifically, The EU legislature decided to create a *sui generis* IP rights¹¹⁷, which would incentivize the creation of databases by offering property-type rights to the database makers according to the previous Directive shall mean a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means databases, within the broad meaning of the Database Directive, are protected in the EU by (i) copyright, where such copyright protection echoes the one recognised in the international treaties; and (ii) a *sui generis* right¹¹⁸. Therefore, while copyright protects the (original) structure of the database, the *sui generis* right aims to cover the investment made in its

2014. In this section, the USCO clarifies that copyright only protects the fruits of intellectual labour that are based on «the intellectual labour that is based on the creative powers of the mind», and specified cites to that effect, the cases *Trade-Mark Cases*, 100 U.S.82 (1879) and *Burrow-Giles Lithographic Co. vs. Sarony*, 111 U.S. 53, 58 (1884) of the U.S. Supreme Court, and *Case Feist Publications c. Telephone Service Company Inc.*, 499 U.S. 340 (1991). The artistic creations made by animals, are different. In this regard, *Case Naruto vs. Slater*, No. 16-15469 (9th Cir. 2018). Slater, núm. 16-15469 (9th Cir. 2018).

¹¹¹ *The New Shorter Oxford English Dictionary* (4th ed., Clarendon Press, 1993) defines databases as «an organized store of data for computer processing».

¹¹² There is no obligation to protect non-original databases under article 10 (2) of the TRIPs Agreement or article 5 of the WIPO Copyright Treaty date. To date, there are no international norms on the protection of non-original databases. This does not, however, imply that database rights are not valuable. Available at: <https://link.springer.com/article/10.1007/s40319-019-00873-2#Fn121> (accessed 25 January 2021).

¹¹³ *Vid.* FERRI (2020), p. 3. In Spanish Law, Royal Decree 1/1996, of April 12, which approves the consolidated text of the Intellectual Property Law, establishes a protection framework for the exploitation of databases under the so called *sui generis* right as well as under rights from author.

¹¹⁴ RICHTSON and GISNSBURG (2006), p. 48.

¹¹⁵ The most influential body of regulation protecting international data transfers is that contained in the European Union (EU) Data Protection Directive 95/462, which will be replaced by the EU General Data Protection Regulation 3 on May 25, 2018. Both instruments restrict the transfer of personal data outside the EU. In his aspect, Commission Implementing Decision 2016/1250 of 12 July 2016, Pursuant to Directive 95/46/EC of the European Parliament and of the Council on the Adequacy of the Protection Provided by the EU-US Privacy Shield, 2016, OJL 207, paragraph 137 (EU).

¹¹⁶ The ECJ (now CJEU) clarified this difference in the case law *British Horseracing Board v William Hill Organization* (ECLI:EU:C:2004:695); case C-362/14, *Schrems vs. Data Prot. Comm'r*, (ECLI:EU:C:2015:6506).

¹¹⁷ In relation with database content extraction, *vid.* footnote 59.

¹¹⁸ The «center of gravity» is always article 114 TFEU (previously, art. 95 of the Treaty of the Economic Community-EEC) which allows the European Parliament and the Council to adopt measures for the approximation of national laws, regulations or administrative acts in order to ensure establishment and functioning of the internal market. *Vid.* GERVAIS (2007), pp. 5-9; FERRI (2020), p. 10. In Spanish Law, article 12 LPI in relation with article 133.1, article 33.2 and article 133.1 and 2 LPI that reproduce article 7.1 Directive 96/9/CE. Remember, that the principal challenges, in the light of jurisprudence case law, are the proof of the infringement and the claim for compensation (*ex art.* 140 LPI).

creation. These two rights are independent, and can be applied separately. They will however apply cumulatively if the conditions for both regimes are simultaneously met. Furthermore, the term of the *sui generis* protection is much shorter than that of the copyright protection. It is limited to 15 years as from the first of January of the year following the date of completion of the database. However, such protection may in practice at may be increased for a period of similar duration for the substantial modifications of the same that are made.

In that context, according to the European legal rules, any substantial change to the contents of the database, that could be considered to be a new investment, will cause the term of protection to run anew¹¹⁹. In practice, should such protection be applied in a Big Data context, this could result in providing an indefinite protection, given that databases are usually dynamic, hence, leading in all likelihood to «substantial changes to the contents of the database» copyright protection is granted to databases which, as such, by reason of the selection or arrangement of their contents, constitute the «author's own intellectual creation»¹²⁰. For this reason, a database structure may be protected under copyright even if the elements contained therein are in the public domain or are otherwise not protected by copyright. It also follows from the previous considerations that the originality criterion might be more difficult to fulfil in case of automatically created electronic databases that contain data selected by software, without the actual involvement of an author¹²¹. In such situations, it seems more likely to award copyright protection to the underlying software, than to the database itself. This is particularly relevant in a Big Data context. Indeed, the development of technology has enabled data analytics of unstructured data. Accordingly, while protection of datasets is particularly relevant, the protection of the database structure has become less relevant and more difficult when confronted to new types of databases, unforeseen by the (over twenty-year-old) Database Directive. The second type of protection introduced by the Directive is the protection awarded on the basis of a *sui generis* rights¹²², rewarding the substantial investment of the database maker in creating the database. Involves a kind of exclusivity of the manufacturer to exploit and use the database that has been manufactured with the information obtained, which translates into the possibility of the manufacturer to prohibit the extraction and /or reuse of all or of a substantial part of the content of the same. It was developed in order to prevent free-riding on somebody else's investment in creating the database and exists in parallel to the copyright protection on the structure of the database. In order for a database to be protected by the *sui generis* right¹²³, an investment must be made in the creation of the database. Thereby, the jurisprudence of the Court of Justice of European Union («CJEU») has clarified that an investment in the creation of the data as such does not suffice to merit protection under the *sui generis* right¹²⁴. In accordance with

¹¹⁹ Article 10 (3) of the Database Directive.

¹²⁰ No other criteria shall be applied to determine the eligibility of databases for that protection (Database Directive, art. 3(1) adopts largely the same focus on processing of personal data in the context of activities of an establishment of a controller in the Union as did article 4 (1) (a) DPD.

¹²¹ MCAFEE and BRYNJOLFSSON (2012), pp. 61-67.

¹²² The term «*sui generis* right» is a generic one and means «*the right of its own kind*».

¹²³ In Spanish doctrine, FERNÁNDEZ CARBALLO-CALERO (2021), págs. 100-108; SANJUAN RODRÍGUEZ (2020), pp. 28-30.

¹²⁴ Guarantee protection against the appropriation of the results obtained from the economic and work investments made by the person who searched for and compiled the content of a database so that the whole or

the above, in order to benefit from the protection conferred by the *sui generis* right aforementioned, it is essential to prove the existence of an investment both for the constitution of the database and for its operation. This investment, as the Court of Justice of the European Union points out, «may consist of the use of resources or human, financial or technical means, but it must be substantial from a quantitative or qualitative point of view. The quantitative appreciation refers to valuable means in figures and the qualitative appreciation to non-quantifiable efforts, such as intellectual work».

Thus, «it will be essential that those who intend to benefit from the *sui generis* right register and collect all the resources- financial, technical or human- so that the investment in the generation and maintenance of the database can be credited»¹²⁵. Such reasoning would entail that the *sui generis* right does not apply to machine-generated databases, as it could be argued that the data included in such databases are «created» instead of «obtained»¹²⁶. This could have a broader effect on the data economy, which relies on digitalization processes such as «Internet of Things» (IoT) devices, big data and artificial intelligence; as it becomes increasingly difficult to distinguish between the generation and the obtainment of data in the context of such processes. That being said, there is no automatic exclusion from *sui generis* protection when the database's creation is linked to the exercise of a principal activity in which the person creating the database is also the one creating the materials that are processed in the database. It is however always the responsibility of that person to demonstrate a substantial investment (qualitative and/or quantitative) in the obtaining, verification or presentation of the content, independent from the resources used to create the content¹²⁷.

On the one hand, we foresee that it will become increasingly difficult to satisfy the *sui generis* right protection requirements in a data economy context, given that the processes of obtaining, verifying and/or presenting the data will happen more and more automatically, as they will be normally conducted using an algorithm¹²⁸. Elsewhere, it might be true that the investment in creating the raw material exceeds the investment made in segmenting and aligning that pre-existing raw material. In those cases, it might be more difficult to rely on the *sui generis* protection. It is in our view regrettable that the Database Directive, which was drafted in the 90s, does not accommodate for the technical evolution and thus everything that is possible with data and databases today. For instance, it is unclear how techniques of enrichment, partitioning, harmonization, homogenization, etc. of data would fit within the criteria of obtaining, verification or presentation of the database contents¹²⁹. Moreover, the criteria of «verification» may become less and less pertinent, especially in a big data context which allows analytics of unstructured data. About the possibility to protect

a substantial part of the database is protected against certain acts of third parties or competitors. It is very interesting case C-46/02 *Fixtures Marketing Ltd vs. Oy Veikkaus AB* (2004) (ECLI:EU:C:2004:694); case C-338/02 *Fixtures Marketing Ltd vs. Svenska Spel AB* (2004) (ECLI:EU:C:2004:696); case C-444/02, *Fixtures Marketing Ltd vs. Organismos Prognostikon Agonon Podosfairou* (2004) (ECLI:EU:C:2004:697); case C-203/02 *British Horseracing Board Ltd and others vs. William Hill Organization Ltd* (2004) (ECLI:EU:C:2004:695).

¹²⁵ SANJUAN RODRÍGUEZ (2020), p. 29.

¹²⁶ FERRI (2020), pp. 15-28.

¹²⁷ YAMAMOTO and LLOYD (2019), p. 125; XALABARDER PLANTADA (2020), pp. 205-207.

¹²⁸ FERNÁNDEZ CARBALLO-CALERO (2021), pp. 106-108.

¹²⁹ SYNODINOU (2020), p. 125.

data under database rights, in view of the rules described above it seems that it is very limited. It is true that the *sui generis* protection forbids extraction of all or a substantial part of the database contents to another medium, preventing thus also the copying of the individual data collected in a database¹³⁰. However, once the database maker renders the contents of its database accessible to the public, it cannot prevent third parties from consulting that database. Accordingly, the public is therefore aware of these data (information), and may use them without necessarily having to copy the database contents. Also, the current legal regime seems difficult to reconcile with developments in technologies such as Big Data that don't necessarily require data to be reproduced in order to perform analytics or mining processes¹³¹. In consequence, the ownership of rights to a database does not confer the rights to the individual data as such¹³². In this context, database protection (both by copyright and the *sui generis* protection) should rather be seen as a complementary measure to protection granted to individual data under other titles such as traditional copyright or trade secret protection.

2. Big Data protected as trade secrets and confidential information

The concept of trade secret traditionally refers to the secrets and confidential information laws, and contracts, can be used to enable the orderly disclosure of information¹³³. The nature of this protection is a subject of debate, however¹³⁴. That protection is reflected in the TRIPs (art. 39, para. 2)¹³⁵. This type of protection of secrets information¹³⁶ is compatible with, and often based on, legislation such as the Trade Secrets Directive¹³⁷ and a host of national laws. In this aspect, protecting undisclosed know-how and business information enables its creator to transform the effort invested in generating this know-how and information into a competitive advantage. In view of big data projects, trade secret protection may provide a safeguard as it allows for protection of individual pieces of information regardless of their originality¹³⁸. It also does not differentiate between the types of data that might be protected. Moreover, the protection

¹³⁰ GEIGER (2019), pp. 29-59.

¹³¹ MCAFEE and BRYNJOLFSSON (2012), p. 65; GERVAIS (2020), p. 107; PUYOL MORENO (2014), p. 502.

¹³² *Vid.* Recital 45 of the Database Directive (amended by Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019, on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC).

¹³³ SUÑOL (2019), p.1; FERNÁNDEZ CARBALLO-CALERO (2021) pp. 100-108.

¹³⁴ GERVAIS (2020), p. 24.

¹³⁵ Article 39 TRIPs obliges signatories to ensure that natural and legal person may protect «undisclosed information» from being shared, acquired or used without their consent on the conditions. In opinion of FOSS-SOLBREKK, «that the information is secrets, meaning it remains «generally» unknown to people commonly working with such information. Secondly, that its commercial value stems from secrecy. Thirdly, that the person managing said information has taken reasonable steps to protect secrecy. Thorough TRIPs does not define trade secrets, recital 1, 2 and Recital 14 of EUTSD elucidate that they cover know-how, as well as business and technological information, in addition to commercial data or customers. Bifurcates trade secrets from other IPRs, such copyrights and patent, which are treated as property in the EU». *Vid.* FOSS-SOLBREKK (2021), p. 11; GERVAIS (2020), pp. 23-25.

¹³⁶ In the Spanish doctrine, MASSAGUER (2019), pp. 46-70.

¹³⁷ *Vid.* Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure.

¹³⁸ It is said that the Community legislator did not have in mind the regulation of «data» when promulgating the Directive (*vid. Data Ownership and Access to Data. Position Statement of the Max Planck Institute for Innovation and Competition*). It is interesting the comment of SUÑOL (2019), p. 1.

is unlimited in time, as long as the information has not been disclosed¹³⁹. Implementing this Directive may change the way trade secrets have previously been dealt with by affording a greater level of protection to big data processes and data science techniques¹⁴⁰. The Directive aims to standardize the national laws of the Member States as regards the unlawful acquisition, disclosure and use of trade secrets. The Directive harmonizes the definition of trade secrets in accordance with existing internationally binding standards. It also defines the relevant forms of misappropriation and clarifies that reverse engineering and parallel innovation must be guaranteed (since trade secrets are not, strictly speaking, a form of exclusive IP right)¹⁴¹.

In relation with the Data protected as trade secrets, according to the definition provided in the Trade Secrets Directive, a «trade secret»¹⁴² is a piece of information which meets all of the following requirements: (i) it is secret in the sense that it is not, as a body or in the precise configuration and assembly of its components, generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question; (ii) it has commercial value because it is secret; and (iii) it has been subject to reasonable steps under the circumstances, by the person lawfully in control of the information, to keep it secret¹⁴³. Furthermore, should be seen as complementary to intellectual property rights. The protection established for trade secrets will expand to every piece of information (data), as long as it fulfils the protection requirements mentioned above. Some requirements are however difficult to fulfil, such as the need for the data to remain secret. It seems that at least in some jurisdictions it is possible to rely on confidentiality agreements to ensure that the requirement of secrecy of the data under the Trade Secrets Directive is maintained even after the transfer of data has been exercised¹⁴⁴. This is however yet to be confirmed by the Courts. Also, it may be difficult to demonstrate that an individual data has commercial value¹⁴⁵ because it is secret. Many data will be considered valuable only if they are part of a bigger dataset¹⁴⁶.

¹³⁹ LEMLEY (2008), pp. 311-332.

¹⁴⁰ In Spanish Law, L. 1/2019, 20 February, Trade secrets. The Spanish Courts have directly applied this precept for more than a decade to define the concept of business secrecy based on article 13 of the old L. 3/1991 of Competition law, modified by L. 29/2009, 30 December. The only difference that exists between both notions is the nuance that the Directive and the LSE introduce to the need to take into account the potential value that the information can have, which, moreover, has already been warned by the doctrine and the texts do.

¹⁴¹ FOSS-SOLBREKK (2021), p. 3; GERVAIS (2020), pp. 37-38; LEMLEY (2008), p. 312.

¹⁴² *Vid.* article 15 and article 20 Regulation (EU) 2016/679 of the European Parliament and of the Council, of April 27, 2016, regarding the protection of natural persons with regard to the processing of personal data; Recital 35 Trade Secrets Directive.

¹⁴³ WIPO. *Conversation on Intellectual Property and Artificial Intelligence*. Second Session, May 2021, Issue 14, ask a question: Should data and AI applications be protectable by trade secrets or is there a social or ethical interest to override existing trade secret protection? For example, in line with draft guiding principle 6, Transparency, of the UNESCO constituted, *ad hoc*, the Expert Group for the Preparation of a Draft Text of a Recommendation on the Ethics of Artificial Intelligence, 2020. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000373199>.

¹⁴⁴ SAPPÀ (2018) pp. 518-520; LEMLEY (2008), p. 322.

¹⁴⁵ DREXL (2016), «Designing Competitive Markets for Industrial Data-Between Prophetization and Access», *Max Planck Institute for Innovation and Competition. Research Paper*, núm. 16, p. 23 (accessed from <https://ssrn.com/abstract=2862975> on 28 December 2020).

¹⁴⁶ *Vid.* Commission Staff Working Document on the free flow of data and emerging issues of the European data economy accompanying the document Communication Building a European Data Economy, adopted on January 2017, available at: <https://digital-strategy.ec.europa.eu/en/library/staff-working-document-free-flow-data-and-emerging-issues-european-data-economy>.

On the other hand, trade secrets are thus different from IP rights, which are safeguarded through an exclusive right that is legally enforceable. This is notably confirmed in Recital 16 of the Trade Secrets Directive which states that «in the interest of innovation and to foster competition, the provisions of this Directive should not create any exclusive right to know-how or information protected as trade secrets»¹⁴⁷. This entails that the independent discovery of the same know-how or information remains possible¹⁴⁸. In the event that one may rely on trade secret protection, the holder of a trade secret cannot prevent competitors from copying and using the same solutions — reverse engineering —¹⁴⁹ is entirely lawful. Trade secrets are only legally protected in instances where someone has obtained the confidential information by illegitimate means¹⁵⁰.

It follows that once the dataset is published, or disclosed in any other way, the protection can no longer be claimed. This is particularly relevant in a big data context, as data used for big data analytics, and made publicly available, will not qualify as trade secrets¹⁵¹. Therefore, when considering to outsource big data analytics, any company should carefully assess whether its datasets comprise trade secrets that are valuable to the company and which cannot be disclosed for that reason.

VI. CONCLUSIONS

This article reviewed the application of several IP rights to Big Data (copyright, patent, *sui generis* database right and trade secret). Beyond the protection of software used to collect and process Big Data *corpora*, copyright's traditional role is challenged by the relatively unstructured nature of the non-relational databases typical of them, which implies that such *corpora* are unlikely to be protected by copyright or by the EU *sui generis* rights in databases. The questions concerning patents are not easy to answer. AI systems can be used to expand patent applications, but they can also be used to assess future incremental innovation and disclose them. Whether that disclosure will be interpreted by patent offices and courts as novelty-defeating is an open question. Addressing data exclusivity and trade secrets, the latter might protect correlations and insights generated by AI systems, even if those are based on deep learning including the processing of protected personal data. This might generate tension between personal data protection and IP. Consequently, legal enforcement based on data, and in-depth analysis of cases, they will set the guidelines that our Courts of Justice must follow, in the near future, in the field of AI in relation to Intellectual Property right, particularly, with the responsibility arising from its use.

VII. BIBLIOGRAPHY

BARRIO ANDRES, Moisés (2018), *Internet de las cosas*, Reus, Madrid.
 BERCOVITZ RODRÍGUEZ-CANO, Rodrigo (2013), *Comentarios al Convenio de Berna para la protección de las obras literarias y artísticas*, Tecnos, Madrid.

¹⁴⁷ GÓMEZ SEGADE (2020), pp. 116-135.

¹⁴⁸ Trade Secrets Directive, Recital 16.

¹⁴⁹ SUNOL (2019), p. 1.

¹⁵⁰ SAPPÀ (2018), p. 523.

¹⁵¹ European Commission, IPR Helpdesk-Guide to IP in Europe, 30 April 2021, available at: https://intellectual-property-helpdesk.ec.europa.eu/trade-secrets_en.

- DEBUSSCHE, Julien, and JASMIEN, César, (2019), «Big Data & Issues & Opportunities: Intellectual Property Rights», available at: <https://www.twobirds.com/en/news/articles/2019/global/big-data-and-issues-and-opportunities-ip-rights>.
- DEVINS, Caryn; TEPPON, Felin; KAUFFMAN, Stuart, y KOPPL, Roger (2017), «The Law and Big Data», *Journal of Law and Public Policy* 27, pág. 306.
- DREXL (2016), «Designing Competitive Markets for Industrial Data-Between Propheticization and Access», Max Planck Institute for Innovation and Competition. Research Paper, núm. 16, p. 23 (accessed from <https://ssrn.com/abstract=2862975> on 28 December 2020).
- DREXL, Josef, y HILTY, Reto (2019), *Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective*, Version 1.0, available at: <https://ssrn.com/abstract=3465577>.
- DE MIGUEL ASENSIO, Pedro (2015), *Derecho privado de Internet*, Civitas, Madrid.
- DUQUE LIZARRALDE, Marta (2020), «Las obras creadas por Inteligencia Artificial: un nuevo reto para la Propiedad Intelectual», *Revista de Propiedad intelectual* 64, enero-abril, págs. 13-67.
- FAVARETTO, Magdalena; DE CLERCQ, Eva, y SIMONE LEGER, Berenice (2019), «Big Data and discrimination: perils, promises and solutions. A systematic review», *Journal of Big Data*, vol. 6, núm. 12, págs. 2-27.
- FERNÁNDEZ CARBALLO-CALERO, Pablo (2021), *La Propiedad Intelectual de las obras creadas por Inteligencia Artificial*, Aranzadi, Cizur menor.
- FERNÁNDEZ MASIÁ, Enrique (2018), «Programas de ordenador», en PALAU RAMÍREZ, F., y PALAO MORENO, G. (dirs.), *Comentarios a la Ley de Propiedad Intelectual*, Tirant lo Blanch, Valencia, págs. 1183-1197.
- FERRI, Federico (2020), «The dark side(s) of the EU Directive on Copyright and related rights in the Digital Single Market», *China-EU Law Journal*, págs. 1-18.
- FOSS-SOLBREKK, Katarina (2021), «Three routes to protecting AI systems and their algorithms under IP law: The good, the bad and the ugly», *Journal of Intellectual Property Law & Practice*, vol. 00, núm. 0, págs. 1-12.
- GALLEGO SÁNCHEZ, Esperanza (2019), «La patentabilidad de la Inteligencia Artificial. La compatibilidad con otros sistemas de protección», *La Ley Mercantil* 59, págs. 1-24.
- GARCÍA VIDAL, Ángel (2020), «Propiedad Intelectual y minería de textos y datos: estudio de los artículos 3 y 4 de la Directiva (UE) 2019-790», *ADI* 40 (2019-2020), págs. 99-124.
- GEIGER, Christopher (2007), «From Berne to National Law, via the Copyright Directive: The Dangerous Mutations of the three-step test», *European Intellectual Property Review*, vol. 29, núm. 12, págs. 486-491.
- «Text and Data Mining: article 3 and 4 of the Directive 2019/790/EU», en SAIZ GARCÍA, C., y EVANGELIO LLORCA, R. (dirs.), *Propiedad Intelectual y Mercado Único Digital*, Tirant lo Blanch, págs. 28-63.
- GEIGER, Christopher; FROSIO, Giancarlo, and BULAYENKO, Oleksandr (2019), *Text and Data Mining in the Proposed Copyrights Reform: Making the EU Ready for Age of Big Data? Legal Analysis and Policy Recommendation*, Max Planck Institute for Innovation and Competition, Munich.
- GERVAIS, Daniel (2012), *The TRIPs Agreement: Drafting History and Analysis*, Sweet & Maxwell, UK.
- (2020), «Is Intellectual Property Law Ready for Artificial Intelligence?», *GRUR International*, 69 (2), pág. 107.
- GIL GONZÁLEZ, Elena (2016), *Big Data, privacidad y protección de datos*, Agencia Española de Protección de Datos.
- GÓMEZ SEGADÉ, José Antonio (2020), «Protección de los macrodatos (*Big Data*) mediante secretos empresariales», en GARCÍA VIDAL, Á. (dir.), *Big data e Internet de las cosas. Nuevos retos para el Derecho de la competencia y de los bienes inmateriales*, Tirant lo Blanch, Valencia, págs. 116-150.

- GUADAMUZ, Andrés (2017), «Artificial Intelligence and Copyright», *Revista OMPI*, octubre, pág. 1.
- GUDÍN RODRÍGUEZ-MAGARIÑOS, Faustino (2018), *Nuevo Reglamento Europeo de Protección de Datos versus Big Data*, Tirant lo Blanch, Valencia.
- HANEY, Brian (2020), «AI Patents: A data driven approach», *Chicago Kent Journal of Intellectual Property*, vol. 19, págs. 410-484.
- KAPLAN, Andreas, and HAENLEINM, Michael (2019), «Siri, Siri, in my Hand, who's the Fairest in the Land? On the Interpretations, Illustrations and Implications of Artificial Intelligence», *Business Horizons Review*, págs. 15-25.
- KUNER, Christopher; BYGRAVE, Lee, and DOCKSEY, Christopher (2020), *The EU General Data protection regulation (GDPR). A Commentary*, Oxford University Press, Oxford, UK.
- LEMLEY, Mikel (2008), «The Surprising Virtues of Treating Trade Secrets as IP Rights», *Stanford Law Review*, núm. 61, págs. 311-332.
- MARCO, Alan; CARLEY, Michael; JACKSON, Steven, y MYERS, Amanda (2015), «The USPTO Historical Patent Data Files Two Centuries of Innovation», *USPTO. Economic Working Paper*, págs. 1-55.
- MARTÍNEZ PÉREZ, Miriam (2020), «Protección tecnológica implicada en el funcionamiento del Internet de las cosas», en GARCÍA VIDAL, A. (dir.), *Big data e internet de las cosas. Nuevos retos para el Derecho de la competencia y de los bienes inmateriales*, Tirant lo Blanch, Valencia, págs. 400-414.
- MASTRELIA, Darío (2018), «Patent information and technology transfer in the information society era: From the current scenario to new business ideas», *European Intellectual Property Right Review*, vol. 40, núm. 7, págs. 460-466.
- MASSAGUER, José (2019), «De nuevo sobre la protección jurídica de los secretos empresariales. A propósito de la Ley 1/2019, de 20 de febrero, de secretos empresariales», *Actualidad Jurídica Uría Menéndez* 51, págs. 46-70.
- MCAFFEE, Charles, y BRYNJOLFSSON, Anthony (2012), «Big Data: The Management Revolution», *Harvard Business Review*, vol. 90, núm. 10, págs. 61-67.
- NAVAS NAVARRO, Susana; GÓRRIZ LÓPEZ, Carlos; CAMACHO CLAVIJO, Sandra; ROBERT GUILLÉN, Santiago; CASTELLS I MARQUÉS, y MATEO BORGE, Iván (2017), *Inteligencia artificial*, Tirant lo Blanch, Valencia.
- ORDELIN FONT, Jorge Luis (2019), «El uso de la inteligencia artificial en la minería de datos y textos: Un análisis desde el derecho de autor», *ADI* 39 (2018-2019), págs. 2019-2029.
- PORTELLANO, Pedro (2020), «Inteligencia Artificial y responsabilidad por productos defectuosos», *RDM* 316 (abril-junio), págs. 8-25.
- PUYOL MORENO, Javier (2014), «Una aproximación al Big data», *Revista de Derecho de la Universidad Nacional de Educación a Distancia* 14, págs. 471-505.
- RAMALHO, Ana (2018), «Patentability of AI generated Inventions is a Reform of the Patent system needed?», *Institute of Intellectual Property*, Foundation for Intellectual Property of Japan, págs. 1-32.
- RANTOU (2012), «The growing tension between copyright and personal data protection», *European Journal of Law And Technology*, vol. 3, núm. 2.
- RICHETSON, Sam, y GISNSBURG, Jane (2006), *International Copyright and Neighbouring Rights: The Berne Convention and Beyond Two*, OUP Oxford, New York.
- SÁIZ GARCÍA, Concepción (2019), «Las obras creadas por sistemas de Inteligencia Artificial y su protección por el Derecho de autor», *Indret* 1, págs 8-19.
- SANJUAN RODRIGUEZ, Nerea (2020), «La Inteligencia Artificial y la creación intelectual: ¿está la Propiedad Intelectual preparada para este nuevo reto?», *La Ley Mercantil* 72, págs. 1-39.
- SAPPA, Cristiana (2018), «What Does Trade Secrecy Have To Do with the Interconnection-Based Paradigm of the Internet of Things?», *European Intellectual Property Right Review*, vol. 40, núm. 8, págs. 518-523.

- SUÑOL, Aurea (2019), «Big data, inteligencia artificial y secretos empresariales». Disponible en <https://almacenederecho.org/43480-2>.
- SYNODINOU, Tatiana (2020), «Databases producer protection: between rights and liabilities», in APLIN, T. (ed.), *Research Handbook on Intellectual Property and Digital Technologies*, Edward Elgar Publishing, Cheltenham, UK, págs. 181-187.
- XALABARDER PLANTADA, Raquel (2020), «Inteligencia Artificial y Propiedad Intelectual», en CERRILLO I MARTÍNEZ y PEGUERA POCH (coords.), *Retos jurídicos de la Inteligencia Artificial*, Aranzadi, Cizur menor, págs. 205-223.
- YAMAMOTO, Kat, and LLOYD, Robert (2019), «The Role of Big Data and Digitization in Just-In-Time (JIT) Information Feeding and Marketing», *American Journal of Management*, vol. 19, núm. 2, págs. 126-133.