Artificial intelligence: IP challenges and proposed way forward

Marta Duque Lizarralde, LL.M, TUM, and Dr Claudia Tapia, LL.M, 4iP Council, examine the recent developments in AI and their impact on the industry.

I.   Introduction

From the technical point of view, it is nowadays possible to produce 'patentable inventions' with AI. For example, by exploring and mixing large and complex bodies of data of technical compounds, AI can create new technical compounds which treat a certain disease. On the downside, these AI 'tools' focus on a very specific field and still demand significant human intervention. In other words, we are far from the so-called "artificial general intelligence", where the AI system is so independent that it requires very little to no human supervision to create inventions in several different technical fields. From the legal perspective, AI faces several unsolved issues, some of which will be analysed in the following section.

II.   Legal challenges related to AI

a.   Inventorship claims

Some believe AI systems complete the entire inventive and patenting process autonomously, thus deserving to be acknowledged as the inventor. They support this view with the fact that the most sophisticated AI systems are generating inventions and that AI is already being used to draft patent applications.[1] Others strongly disagree with that position.[2]

The heated debate became even more popular when, in 2019, Dr Thaler filed two patent applications designating an AI system named DABUS as the inventor in several offices worldwide, which were rejected by several of them (the UKIPO, the EPO, and the USPTO). Their main argument (later on affirmed by the UK High Court[3], the UK Court of Appeals,[4] and the US District Court for the Eastern District of Virginia[5]) was that, in the respective patent statutes, the inventor is addressed with pronouns

Marta Duque Lizarralde

Dr Claudia Tapia

that are only used for natural persons, such as "him" and "her". Therefore, interpreting the term inventor so broadly as to include an AI system would go against the principle of plain reading. In addition, the offices pointed out, DABUS lacks legal personality and, consequently, the capacity to own IP rights and to transfer them to Dr Thaler.

In contrast, the South Africa patent office issued in July 2021 a patent listing DABUS as the inventor, and the Federal Court of Australia ruled that AI systems can be recognised as inventors under the Australian Patent Act.[6] Thus, the debate is far from being over.

A common mistake within the inventorship discussion is to confuse automation with autonomy. The use of AI in the inventive process allows to automate the performance of different tasks. However, the conception of "invention" remains attributable to the natural persons employing it.[7] Yet, identifying the inventor of "AI-assisted works", which WIPO defines as works "generated with material human intervention and/or direction",[8] can on some occasions be challenging. This is because the degree and number of contributions from different actors vary depending on the project and the application of the AI technology for that particular case.

b.   Authorship

For a work to be eligible for copyright protection, it must be original. A work is considered original if it is "the author's own intellectual creation" manifested by their "free and creative choices".[9] On the other hand, even if not explicitly stated, it could follow from the provisions of the Berne Convention and the EU copyright directives that the author must be a natural person.[10]

A distinction must be made here between the

above-mentioned "AI-assisted works" that would be protectable if they met the originality requirement, and "AI-generated works", defined by WIPO as those created by "AI without human intervention".[11]

Many of the results that are referred to as "AI-generated", including "the next Rembrandt", are actually AI-assisted, because human involvement in the different phases that predetermine the outcome is still decisive. Since AI systems are not capable yet of generating results autonomously, the definitions adopted by WIPO do not reflect the state of the current debate. A more accurate term for this type of existing creations is that of "Authorless AI-assisted work ", adopted in the 'Trends and Developments' in the AI report.[12] Examples of this type of works would be the initial translations performed by DeepL, some reports generated in the field of automated journalism, or texts created with sophisticated language models, such as GPT-3. These results, created using advanced training methods, are still tied to pre-existing data and parameters provided by the AI developers. Thus, the space for the creative freedom needed to meet the originality requirement is too limited.

It is also debatable whether authorless creations could be protected by certain related rights, such as the rights of phonogram and film producers, broadcasting organisations, publishers of press publications, and non-original photographs, since they do not require originality or human authorship; or whether a legislative reform would be needed, since their ownership is still conceived only for humans. There has furthermore been some discussion on the desirability of creating a new sui generis right.[13]

Résumés

Marta Duque Lizarralde, LL.M, is a Doctoral Candidate and Research Associate at the Technical University of Munich (TUM). Prior to joining TUM, she worked as IPR policy researcher at Ericsson. She holds a law degree from the University of Salamanca (thesis 10/10), an LL.M specialising in Intellectual Property from the Universidad Carlos III Madrid (Special Award as the student with the highest grades of the LL.M. in Intellectual Property, edition 2018-2019), and an LL.M. in Law of Internet Technology from the Bocconi University (110/110, cum laude). Marta also won the Intellectual Property Prize organized by the LL.M. in Intellectual Property and New Technologies of the Universidad Autónoma de Madrid (UAM) in 2020.

Dr Claudia Tapia, LL.M is Chair of 4iP Council (a non-profit research council focused on IP and innovation), Director of IPR Policy & Legal Academic Research at Ericsson, and vice chair of the Patent and Technology Licensing Committee of LESI. Prior to joining Ericsson, she worked for over five years as Director IP Policy at BlackBerry. She holds a law degree from the University of Valencia, an LLM degree specialising in International Patent Law from the Ludwig-Maximilian University in Munich and a PhD degree on FRAND and Standardisation in the telecoms sector (summa cum laude) from the Faculty of Law in Augsburg. Claudia is also chair of the Accelerator IP Advisory Group at EIT Health, and an Advisory Board member of C-IP2.

"

A common mistake within the inventorship discussion is to confuse automation with autonomy.

"

1   The World Economic Forum (WEF), 'Artificial Intelligence Collides with Patent Law' (2018) <http://www3.weforum.org/docs/WEF_48540_WP_End_of_Innovation_Protecting_Patent_Law.pdf>

2   Daria Kim 'AI-Generated Inventions': Time to Get the Record Straight?', *GRUR International 69 (5)* 443.456.

3   *Thaler v Comptroller General of Patents Trade Marks and Designs* [2020] EWHC 2412 (Pat).

4   *Thaler v Comptroller General of Patents Trade Marks and Designs* [2021] EWCA Civ 1374.

5   *Stephen Thaler v. Andrew Hirshfeld, Performing the Functions & Duties of the Under Sec'y of Com. for Intell. Prop. & Dir. of the United States Pat. & Trademark Off., et al.,* No. 1:20-cv-903 (LMB/TCB), 2021 WL 3934803 (E.D. Va. Sept. 2, 2021).

6   *Thaler v Commissioner of Patents* [2021] FCA 879 (30 Jul 2021).

7   Noam Shemtov, 'A Study on Inventorship in Inventions Involving AI Activity' (2019) <http://documents.epo.org/projects/babylon/eponet.nsf/0/3918F57B010A3540C125841900280653/$File/Concept_of_Inventorship_in_Inventions_involving_AI_Activity_en.pdf>:

8   WIPO, 'Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence' (21 May 2020) <https://www.wipo.int/edocs/mdocs/mdocs/en/wipo_ip_ai_2_ge_20/wipo_ip_ai_2_ge_20_1_rev.pdf>

9   Among others, C-145/10, *Painer v. Standard VerlagsGmBH and Others (2011) ECLI:EU:C:2011:798,* para 119,120; C-604/10, *Football Dataco Ltd and Others v. Yahoo! UK Ltd and Others (2012) ECLI:EU:C:2012:115,* para 37,39; C-403/08 and C-429/08, *Football Association Premier League v. QC Leisure* and *Karen Murphy v. Media Protection Services* (2011) ECLI:EU:C:2011:631, para 97; C-310/17, Levola Hengelo (2018) ECLI: EU: C: 2018:899, para 35,36.

10   For example, Art. 7.1 Berne Convention states that *"The term of protection granted by this Convention shall be the life of the author and fifty years after his death".* In the Community framework, Art. 1 Directive 2009/24/EC89, Art.6 Directive 2006/11690, and Art.3 Directive 2006/1169/EC state that for a creation to be protectable by copyright, it must be original in the sense of constituting *"the author's own intellectual creations".*

11   WIPO (n.9)

12   Bernt Hugenholtz et al. 'Trends and Developments in Artificial Intelligence, Challenges to the Intellectual Property Rights Framework, Final Report' (2020) <https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=71915>

13   For instance, Ana Ramalho, 'Will robots rule the (artistic) world? A proposed model for the legal status of creations by artificial intelligence systems' (2017) *21 Journal of Internet Law,* 12,25: proposes "a disseminator's right, bearing a similar regime to the publisher's right in the publication of previously unpublished works as prescribed by the EU Term of Protection Directive.

The latter is supported by part of the academic community, which rejects the idea of any kind of AI-generated works falling into the public domain. Some argue, however, that it may not even be necessary considering the available tools, such as trade secrets, factual control, and unfair competition, to protect the results of creative AI systems.[14]

c. Liability

Another question that is keeping stakeholders busy is the one of liability, in particular considering scenarios where the AI engineer will have limited to no influence on the behaviour of the AI system. In particular it might be problematic if an AI engineer generates a method and different companies apply it for different use cases. The AI engineer cannot know where it will be applied and what the use case may imply. Moreover, with global interconnectivity, it will become increasingly difficult to identify who among the many actors had caused certain damage and, if several of them were responsible, to which degree each of them.

d. IP protection of AI features: Copyright, patents, and trade secrets

Companies wishing to benefit from their own investments in AI are wise to adopt an efficient IP strategy to protect the different elements of AI systems. A starting point of such strategy should consist in identifying the AI features that

> **The definitions adopted by WIPO do not reflect the state of the current debate.**

are protectable by IP rights.

AI systems are formed by algorithms, which by themselves are not copyrightable. Nevertheless, they are encoded in a programming language and incorporated into software that is protectable by copyright. However, copyright does not protect the program's underlying ideas and principles but rather the way in which they are expressed. Consequently, the functional aspects of the software are not covered by copyright.s[15] Against this background, there is an active debate on whether Machine Learning models can qualify as learning algorithms, AI software or super-software. Some argue that they do not meet the originality requirement[16]; others that complex, dynamic Machine Learning models might be protected by the *sui generis* database right.[17] Further research is needed on this issue.

The algorithms, weights, models, and evaluation mechanisms that compose an AI system are of an abstract mathematical nature. Therefore the European Patent Office typically excludes them from patentability when claimed as such. Nevertheless, these features applied in an invention with a technical character can be protected as elements of the invention.

Finally, there are other elements that, when not protectable by copyright and patents, or holders are not willing to disclosed them, are protectable by trade secrets. But are trade secrets the best option considering our need for
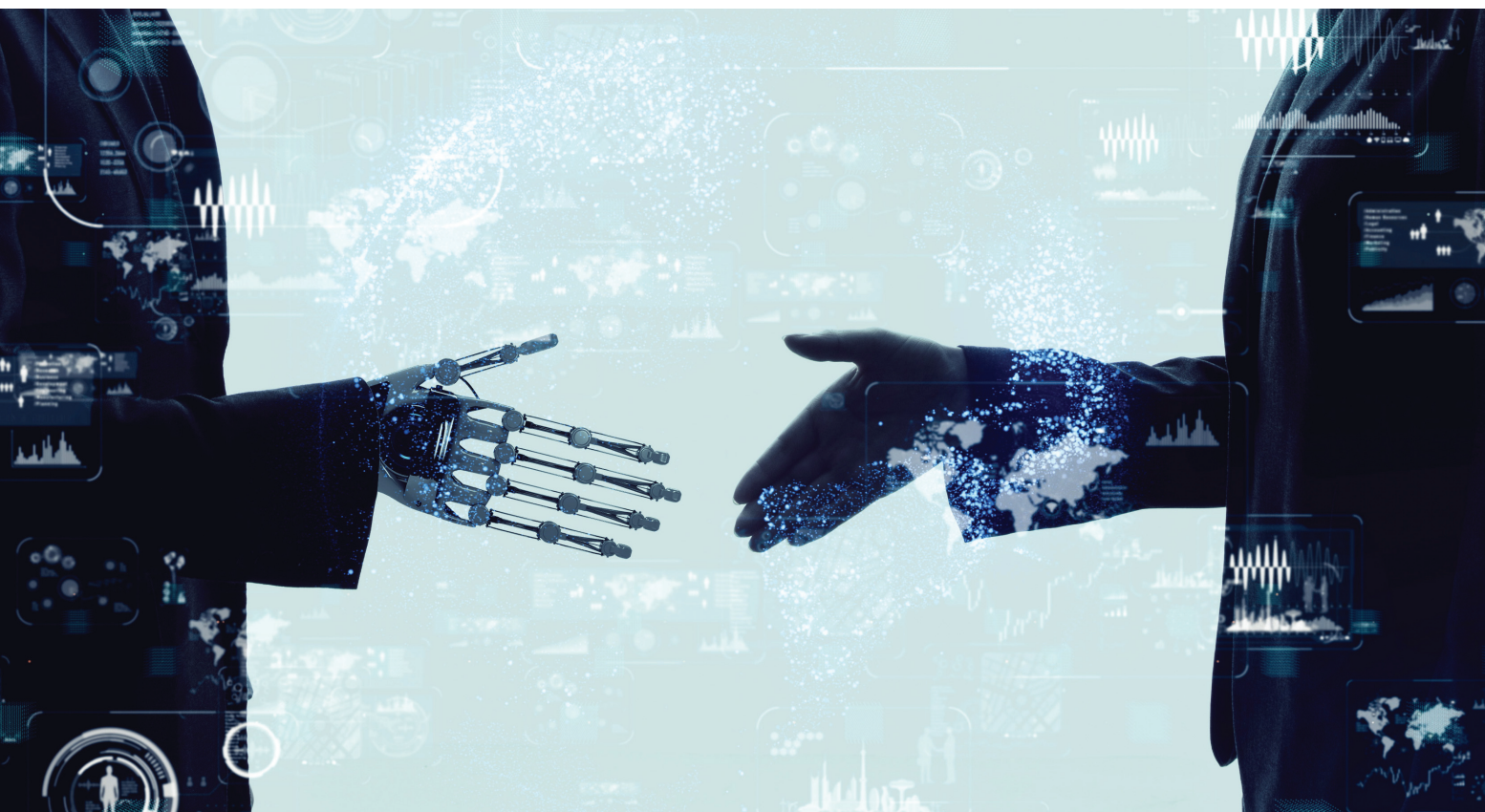
maximal diffusion and further collaboration to develop AI?

e. IP rights embedded in the training datasets

One important legal barrier for data sharing is the uncertainty about which IP rights are embedded in the training datasets. Training datasets often include data that is publicly accessible and freely available on websites. While raw data is not protected by IP rights, other data, such as images or sounds, can be protected by copyright or related rights. Consequently, if the latter data are not covered under the Text and Data Mining exceptions of the Directive on copyright and related rights in the Digital Single Market,[18] a license will be needed for their use. It is also unclear whether the training datasets can be protected by copyright and the sui generis database right. Thus, there is a risk that companies choose to restrict access to raw data and datasets by means of factual control.[19]

In practice, triggering business-to-business (B2B) data sharing is resulting in a very challenging process, among other reported reasons, because of the lack of confidence among economic operators that the data will be used in accordance with the contractual agreements, or the fear of losing a competitive advantage.[20]

f. Competition law

In this situation, it must be evaluated if competition law could be used to correct imbalances. Yet, companies cannot be forced to license their datasets merely because they have a competitive advantage and have refused to license them. Access to data under competition law can only be granted in the circumstances set out in the essential facilities doctrine. The application of such doctrine to this case is problematic because in most cases datasets are not 'essential' since it would be feasible to

> **A starting point of such strategy should consist in identifying the AI features that are protectable by IP rights.**

find a substitute of the same.[21] Notwithstanding this, the EC has stated that compulsory access to data on fair, reasonable and non-discriminatory terms (FRAND) *"will be needed where specific circumstances so require".*[22] It will therefore be necessary to bring some clarity and identify those "specific circumstances" and how FRAND is to be interpreted in this context. Also, it may be recalled that competition law only operates ex-post, and that ex-ante regulation seems more appropriate to address B2B data-sharing concerns.

g. Soft law

In 2018, the EC issued non-binding guidance setting some principles to which the parties might adhere for promoting the voluntary B2B data sharing based on fair contractual arrangements.[23] However, they have proven to be insufficient. As a result, the EC has recently stated that it *"will continue to assess whether amended principles and possible codes of conduct are sufficient to maintain fair and open markets, will address the situation"*, and if needed, [it] *"take appropriate actions".*[24]

h. Data protection

AI development is dependent on the availability of large quality datasets for its training, at least for most AI systems. In general, for many AI systems to work properly, specific data must be collected, organised and prepared in a very particular way with the know-how of the AI engineer. In other words, as the algorithm will not work if the AI system is provided with a random selection of data, one needs to filter or 'clean' the 'lake' of data. This means removing any inconsistencies, duplicates or incorrect entries, and verifying that the data is accurate, complete, reliable, and up to date. Companies can easily spend around 80% of the resources on collecting and preparing the data. To pre-process the data to be used in the AI system

14 Bernt Hugenholtz *et al.* (n.13)

15 Peter R Slowinski ´Rethinking Software Protection, ´ Draft Chapter, in J.-A. Lee, K.-C. Liu, R. M. Hilty *(eds.), Artificial Intelligence & Intellectual Property, Oxford, Oxford University Press,* 2020 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3708110>

16 Begoña Gonzalez Otero, 'Machine Learning Models under the copyright microscope: is EU Copyright fit for purpose?' (2021) GRUR International 70(11), 1043,1055.

17 Josef Drexl, Reto M. Hilty, Luc Desaunettes-Barbero, Jure Globocnik, Begoña Gonzalez Otero, Jörg Hoffmann, Daria Kim, Shraddha Kulhari, Heiko Richter, Stefan Scheuerer, Peter R. Slowinski, Klaus Wiedemann, ´Artificial Intelligence and Intellectual

Property Law Position Statement of the Max Planck Institute for Innovation and Competition of 9 April 2021 on the Current Debate´ (2021) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3822924>

18 Arts. 3 and 4 Directive (EU) 2019/790.

19 Josef Drexl, 'Designing Competitive Markets for Industrial Data – Between Propertisation and Access' (2017) *8 JIPITEC*, para 6,12.

20 Communication from The Commission to The European Parliament, The Council, The European Economic and Social Committee and The Committee of The Regions 'Building A European Data Economy', COM/2017/09 final [2017].

21 Giuseppe Colangelo and Mariateresa Maggiolino, 'Big Data as Misleading

Facilities'(2017), *13 European Competition Journal 13(2),* 272,274.

22 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 [2020] 13

23 Commission Staff Working Document Guidance on Sharing Private Sector Data In The European Data Economy, SWD(2018) 125 final [2018].

24 Communication from The Commission to The European Parliament, The Council, The European Economic and Social Committee and The Committee of The Regions, 'Towards A Common European Data Space', COM (2018) 232 final [2018].

there are two options. Either humans process the data, or they use automation tools or even human-created AI system to do so. Having taken into consideration the significant work and research behind the filtering of data, companies generally wish to protect 'data cleaning' systems with patents. However, some patent offices are reticent to recognise the technical purpose of that invention because they perceive the system as 'only' manipulating and reorganising data. Unfortunately, to date there is no harmonisation amongst patent offices on this topic.

Moreover, some questions arise regarding ownership and transfer of data. What can data holders do with the data and how do they maintain the ownership? If data is generated, for example, by the operators running through a network, who has the ownership rights? If the algorithm uses data from another data holder and changes it, who is the owner of the transformed data? How can companies protect personal data in compliance with GDPR and data protection regulations when transferring data between different countries? International companies or institutions may require that their employees exchange data in order to create and make AI systems work. Data holders may

> **In this situation, it must be evaluated if competition law could be used to correct imbalances.**

wish to allow others to use the data in some cases, but does this mean we need an open database?

i.   Natural language processing
Some patent offices employ a system for the prior art search that uses so-called 'natural language processing'. This system enables the use of algorithms to 'match best' the text of patent applications to a big database of text. As patent examiners still need to filter, verify and sometimes correct the results, this processing is only an 'assisting tool' for patent examinations. Patent offices also use the system for their classifications and statistics, which they later on publish in several reports. Let's say "Top 10 companies with patents in AI Machine Learning". This has led to companies including some particular keywords in their patent claims (in this example 'AI Machine Learning'). Should they choose not to do so, they may risk being perceived as not innovative in that area.

In any case, there is a great deal of potential in those AI tools. Google and IBM have developed a very elaborated natural language processing system using algorithms that enable, for example, to generate human speech or human text that is so realistic that it is very hard for the

individual to distinguish it from that produced by humans. On the downside, let us imagine AI generates thousands of articles on a certain field (impacting the state of art). It would then be very challenging for inventors to obtain a patent because of lack of novelty. Another concern would be if someone with sufficient financial resources decides to pollute the patent environment. That person/association/government could overload patent offices by generating, with AI, thousands of automated patent applications.

j. Enforcement

Let us now imagine one obtains a patent for an AI invention and needs to enforce it. It would not work as it does nowadays with, for example, patents essential to a standard. For essential patents companies can show a claim chart, mapping the patent with the standard. But with some AI systems, it may be difficult to show infringement. Infringement may be in the internal workings of the algorithm, or in the filtering of data. The way some other AI systems work is often not well understood by their own designers. How can you enforce it if the patent office or the court requests to reproduce it? Can we give the data of the operator incorporated in the network for which the AI patent is granted?

k. Functional and geographical distribution

Another challenge regards the functional and geographical distributions of the AI system. Let's imagine that, in the architecture, part of the algorithm is performed at the edge, e.g., in collecting or filtering the data in the mobile phone, and the other part is performed in a node, in the network. Also, how can you enforce your right to an algorithm that is in a device in Germany but the node is in Spain and the execution is in the US? In these scenarios it is usually decisive the territory where the technical result occurs.

III. Proposed way forward

As we are preparing for making the unimaginable possible,[25] it is indispensable that we start addressing the above-mentioned challenges.

Definitions of AI-generated results, such as the one given by WIPO, do not appear to reflect the current state of the debate and may lead to confusion. Therefore, a first step in advancing this debate is for (legal) practitioners to better understand how AI technology works and its actual capabilities, and to make decisions in line with reality.

While the volume of data production is increasing, its potential is still underused, so greater data availability and interoperability should be fostered, especially in the B2B

> **" A first step in advancing this debate is for (legal) practitioners to better understand how AI technology works and its actual capabilities, and to make decisions in line with reality. "**

context. The EC should create a framework for B2B data sharing, taking into consideration the interests of the various players involved and the rapid changes in the data and AI sector, which requires flexibility. However, it remains open what the appropriate legal instruments for this purpose are.

We also need a B2B data-sharing framework, given the importance of data in AI development. But there are still some open questions. For example, would the recommendation of standard licensing rules help to create a common framework? What is the role of the OSS community? Would it be desirable to apply a method of controlling unfair terms?

Regarding liability aspects, one could consider whether to create an insurance tailored to AI driven products.

A final remark is dedicated to courts and patent offices. We need harmonisation by courts and patent offices in getting protection for AI systems and in the enforcement. Equally important is for them to keep up to date with the fast development of AI. Finally, creative solutions may be required to prove infringement. For example, one could present to the patent office the feeding of specific data to the AI system and observe the outcome. If one gets the output expected, then the patent office (or the court) could conclude there is a high likelihood that there is an infringement. Another possibility would be to use 'comparable tests' (with data X one is able to get the claimed result, but with other data that result is not achieved) or to exploit techniques currently developed in the growing field of Explainable Artificial Intelligence (XAI).

[25] Machine Learning and other AI technologies will lead to innovation we cannot even imagine today. See more at https://www.ericsson.com/en/careers/better-brighter-tomorrow