



Intellectual Property newsletter



Edition 4 | 2023

Special Edition on Artificial Intelligence and Large Language Models

Contents

	1	Introduction	p. 03
	2	Gen AI and IP: Here we go again	p. 04
	3	IP rights in Gen AI artistry	p. 05
	4	AI: new regulation for Argentina	p. 06
	5	Vietnam: National Strategy on AI	p. 07
	6	TDM and Gen AI	p. 08
	7	Copyright and Large Language Models	p. 09
	8	AI Risks in Focus	p. 12
	9	The automatic generation of images by AI	p. 17

Introduction

Welcome to a special edition of the KPMG Intellectual Property newsletter focusing on Artificial Intelligence, Large Language Models and how they are reshaping the world of copyright, patents, trademarks, designs, domains and other Intellectual Property rights (“IPRs”).

We have compiled a number of interesting articles from around the world that provide insights into the current developments. KPMG firms are proud of their global network of IP lawyers, business advisors and other IP experts enabling KPMG professionals to offer an international service to clients in this area.

We start by putting the current developments in a broader context. There have always been challenges in the IP sector and we show you what we can learn from them for the future. Furthermore, we analyze how different legal systems around the world are adapting strategies to address the complex questions arising when an AI creates innovative content. We will explain this using the example of the automatic generation of images by AI.

Going into more detail on a national level, we will take a closer look at how Argentina and Vietnam are adapting their legal landscapes to navigate the challenges and opportunities presented by AI.

Gen AI thrives on a large amount of data. We shed some light on the chances and risks of Text and Data Mining and the challenges that arise with regard to copyright and data protection.

We will also focus on the question, why a comprehensive AI governance is essential for businesses and what needs to be considered to ensure the responsible and efficient use of AI-based solutions in companies.

We hope you enjoy reading.



Dr. Anna-Kristine Wipper
Partner
KPMG Law Rechtsanwalts-gesellschaft mbH
T +49 30 53019-9731
awipper@kpmg-law.com



Madlen Müllensiefen
IP Portfolio Management
KPMG Law Rechtsanwalts-gesellschaft mbH
T +49 69 95119-5641
madlenmuellensiefen@kpmg-law.com



Gen AI and IP: Here we go again

While Gen AI – a broad term to describe any AI system which generates content rather than analyzing or acting on existing data – has stunned the mainstream like no other (and seen unprecedented adoption and growth), it has particular significance for those in the business of guard railing intellectual property. Gen AI presents a fresh and emerging challenge, but there are parallels we can learn from when examining the patterns of intellectual property law over time. Bluntly, there have been IP law challenges before, and there will be so again.

As a reassuring starting point, the rules and laws around IP exist to reward human creativity. The spectrum of human creativity is vast and includes the very tools which are now enabling our current wave of Gen AI-related challenge: computers, satellites, and robots are all products of human creativity, as are medicines, music, art, and literature. Patents, one way in which we deal with protection around human inventions, have been recorded since the 15th century and the first Copyright Act came into force in 1710 in England. Since then, intellectual property protection has expanded just as the breadth of invention has expanded, now encompassing design rights, trademarks, database rights, copyright, trade secrets and know-how. In 2022, the UK Intellectual Property Office received 19,486 patent applications indicating continued appetite for innovation.

20th century challenge: the Internet

The internet's arrival changed the ways in which we communicate, consume, shop, play, watch and listen. In late 1999 – the early 2000's, various startup companies made use of peer to peer (P2P) file sharing applications that gave their users free music through their use of the P2P network, which enabled mass copying of copyrighted music. Seemed fun while it lasted, but most of these companies fell afoul of copyright laws which led to their shutdown in the early 2000's after losing several lawsuits and incurring sizeable damages as a result.

The outcome of this early challenge resulted in approaches ranging from digital rights management, piracy suits against businesses and individuals, and new music business models such as streaming and ad revenue, but tensions around streaming music remain.

21st century challenge: Gen AI

Gen AI presents immense possibilities but also immense uncertainties. As a starter, there are serious concerns around inputs into Gen AI. Does training Gen AI data sets on copyrighted material (e.g., literary works, books, news, art), without permission, infringe upon copyright laws? Under the British Copyrights, Designs and Patents Act (CDPA) 1988, it is permissible to undergo text and data mining for research purposes, but other issues are

unresolved, such as data inputs into Gen AI potentially infringing database rights.

In terms of Gen AI outputs issues, there remains many questions to resolve, such as whether AI can generate works that are subject to copyright, and exactly who owns copyright works generated by AI without human intervention.

Globally, the question of whether AI can be an inventor on a patent has seen the Americas, England, Europe, and Australia saying no, while the South African patent office appears to have given a tentative yes.

Future challenges

As we move into a global application, IP law is likely to react slowly, as is the pace of law and legal change, leading to National IPO consultations and national lawsuits. As we have seen with market reactions to the P2P applications of the early 2000's, we might expect new business models, the industry underwriting risk, exemptions, and the potential for enhanced reliance on assertions of their moral rights by creatives. While so much is currently uncertain, we do know that the IP landscape will be reformed in some fundamental ways once again.



Wahid Usman

Partner
KPMG in the UK
T +44 207694-3316
usman.wahid@kpmg.co.uk



Abeeku Mills-Robertson

Director
KPMG in the UK
T +44 754350-9764
abeeku.millsrobertson@kpmg.co.uk

3 IP rights in Gen AI artistry

In the digital realm of code and byte, Generative AI (hereinafter Gen AI) is transforming dreams into pixels, aspirations into reality, and imaginations into innovations. Gen AI fuels a creative renaissance where the fusion of human imagination and artificial intelligence paves the way for unmatched artistry. With cutting-edge algorithms, Gen AI redesigns the canvas of intellectual property with innovative strokes. The blend of artificial intelligence and intellectual property rights reforms the saga of legal evolution in a digital age.

Gen AI, is a groundbreaking technology that creates innovative content like images, music, and text in response to user prompts. At present, people are weighing both the potential benefits and concerns surrounding AI wherein some fear that it will abate human creativity, while others see it as a tool that can empower creators. Gen AI armed with intricate algorithms and machine learning, has become a potent force in generating art, music, literature, and innovative designs autonomously. The proliferation of AI-generated designs raises fundamental questions about intellectual property rights when the creator is an algorithm. Accordingly, the process of determining authorship and ownership of the work becomes complex. Moreover, different legal systems around the world are adapting strategies to address these complexities, ensuring creators receive rightful recognition and protection.

Today, various countries are perceiving the concept of Gen AI as an inventor and exploring diverse approaches to define intellectual property rights. In the United States^(a), copyright regulations shield unique creations, encompassing artworks, literature, and music originating from AI processes. In the United Kingdom^(b), copyright protection is granted to a work created by a computer in the absence of a human author. The work is intended to be owned by either a human or a corporate entity, however, the computer program or AI itself cannot be the author or proprietor of the intellectual property. Instead, the copyright protection for computer-generated works is limited to the human author of the work, referred to as the person by whom the necessary arrangements for creating the work are made. Similarly, Japan's^(c) patent laws encompass inventions that are created using the Artificial Intelligence (AI) and grant rights to the individual that used AI for innovation. In addition, China, Canada, New Zealand, and other countries are also trying to find a balance between fostering innovation and establishing clear regulations for AI-generated innovations. Thus, authorship and ownership of Gen AI as an inventor varies from country to country, reflecting a variety of cultural, legal, and technological contexts. As AI technology advances, finding consensus on the concept of inventorship is essential.

The AI landscape is highly dynamic and will continue to outpace existing legislation. The complexity and unpredictability of AI makes it difficult to create an all-encompassing law. For that reason, current legislation should be amended to create guardrails around the various AI risks. To start, AI models should be able to moderate content and there should be a mechanism to suspend unlawful content. The legal approach should be towards the trusted and responsible use of Gen AI that does not impede innovation. Also, the legal system should elucidate the limits of what constitutes a "derivative work" under intellectual property law, based on various federal circuit court interpretations.

The collaboration between IP experts and tech innovators is essential in crafting a future where Gen AI coexists seamlessly with intellectual property laws. As we navigate this digital frontier, striking a balance between fostering innovation and safeguarding creators' rights is imperative. Moreover, the fundamental step of deploying Gen AI in creating patents, copyrights, and other IPs will be to review the terms and conditions of the AI platform in question and ensure that the rights to utilize the final output are based on your desire.

Note: (a) <https://crsreports.congress.gov/product/pdf/LSB/LSB10922>
 (b) https://www.wipo.int/edocs/mdocs/scp/en/scp_35/scp_35_7_summ.pdf
 (c) <https://books.openedition.org/putc/15392?lang=en>



Sumantra Mukherjee
 Partner
 KPMG Assurance and Consulting Services LLP
 T +91 1243369711
sumantram@kpmg.com



Alka Chauhan
 Associate Director
 KPMG Assurance and Consulting Services LLP
 T +91 1243369001
alkachauhan@kpmg.com



Deepna Mehta
 Assistant Manager
 KPMG Assurance and Consulting Services LLP
 T +91 1243369001
deepnamehta@kpmg.com

AI: new regulation for Argentina



The Agency for Access to Public Information (the “AAPI”) issued Resolution No. 161/2023 which created a Transparency and Personal Data Protection Program in the use of Artificial Intelligence (AI).

On September 4, 2023, the AAPI, as enforcement authority of the Personal Data Protection Law, created the program with the aim of promoting analysis processes, regulating and strengthening institutional capacities regarding the development and use of AI.

It is based on recommendations from the Organization for Economic Co-operation and Development (OECD) that aim to guide governments, organizations and individuals in relation to the design of systems in order to responsibly address the effects of AI technologies, prioritizing the interests of people and, in turn, assigning responsibilities in relation to their operation.

The program focuses on guaranteeing the effective exercise of citizens' rights in terms of transparency and protection of personal data, both in public and private sectors. In other words, it seeks to provide greater security and legitimacy to the use of the AI systems to protect such sensitive data and the right to privacy.

Furthermore, the program proposes the creation of an observatory on AI to follow up regional and global progresses on technological developments based on AI; and a multidisciplinary advisory board with experts in the field to agree on sectorial policies and elaborate regulatory strategies to prevent negative impacts from the use of AI. It will further provide best practice guidelines, training, campaigns and technical assistance on transparency and personal data protection.

Lastly, it will be in charge of the National Directorate for the Evaluation of Transparency Policies and the National Directorate for Personal Data Protection. However, the AAPI will provide support in the development processes and the integration of AI to technological solutions in multiple fields.



Juan Martín Jovanovich
Partner
KPMG in Argentina
T +54 11 4316-5805
mjovanovich@kpmg.com.ar



María Ximena Pérez Dirrocco
Director
KPMG in Argentina
T +54 11 4316-5915
mperezdirrocco@kpmg.com.ar



Andrea Gisela Martínez
Senior Consultant
KPMG in Argentina
T +54 11 4316-5700
andreagiselamartinez@kpmg.com.ar

Vietnam: National Strategy on AI

What can be expected for the IP system?

The application of AI in innovation, content creation and business has become significant across industries of Vietnam over the last years. Sectors where AI is most used include banking and finance (e.g., for eKYC), healthcare (diagnosis, telehealth), automobile (in-car virtual assistance), FMCG (supply chain), education (content communication), e-commerce (personalization of shopping experience), etc. AI applications have also been adopted widely in the automation of public services^(a). In 2022, Vietnam was ranked 55th place globally in the Government AI Readiness Index. Remarkably, most of the AI solutions employed are reported as developed by Vietnamese technology companies. The Government has strongly upheld the production and application of AI products, with heavy public investment in machine learning, big data analytics, and IoT^(b). Vietnam has aimed to be a key AI player with the Prime Minister's National Digital Transformation Program until 2025^(c), and National Strategy on Research, Development and Application of Artificial Intelligence (AI) until 2030 ("**National Strategy on AI**")^(d). It is notable that the National Strategy on AI sets out a number of goals tasks in relation to IP, suggesting that IP is among the Government's key considerations when it comes to the growth of AI.

The potential implications of AI on the IP system have long been discussed in Vietnam, with the most concerned issues involve the protectability and authorship in AI-generated creations. This started from the increasing use of smart machine tools in creative activities and innovative research to replace human power. Lots of creative and innovative products have been generated by, or with substantial assistance of AI. Under the Vietnamese IP Law, works or technical solutions completely or significantly generated by machines are unlikely to be eligible for protection under the copyright or patent regime. The law only recognizes humans who create works or inventions as "authors" for the purposes of IP protection and registration. AI-generated creations are left unregulated under the current law and may therefore be vulnerable to infringements. This situation is expected to be changed as the National Strategy on AI sets its very first task as "to develop and complete legislative documents regarding IP relating to AI" (Section IV – Tasks and Solutions). This task is assigned to the Ministry of Science and Technology ("**MOST**"), who shortly after that launched their own plan to implement the National AI Strategy^(e). According to the MOST's plan, this task is handled by the national IP office and scheduled to be complete in 2025.

Another aspect of IP promoted under the National IP Strategy is the IP commercialization practice. In this regard, the Strategy requests the MOST's support for projects of cooperation on AI technology transfer, and utilization of inventions and industrial property rights of AI between Vietnamese and foreign enterprises. This suggests stimulation made to the yet-to-be dynamic market to commercialize IP in general and AI products in particular. One may expect initiatives from the MOST to invite more inbound transfers of AI technologies as well as to accelerate domestic technology transactions in this area, possibly in form of funds, workshops or advice. The agency to take charge of this mission is the State Agency for Technology Innovation ("**SATI**") under the MOST.

Last but not least, one of the key goals set by the National AI Strategy is to "rapidly increase the number of scientific works and patent applications on AI in Vietnam". Researchers in this area can therefore expect strengthened support from the Government in courses of patent generation. It remains to be seen in which form the support would be made. It is notable that in recent years, the IP office has been active in raising public awareness in this area, with a series of training programs on patent organized for SMEs and startups, with the topics ranging from state-of-the-art search, patent filing to technology licensing.

- Note:
- (a) <https://english.mic.gov.vn/Pages/TinTuc/tinchitiet.aspx?tintucid=157621>
 - (b) <https://en.baochinhphu.vn/viet-nams-ai-leadership-status-is-blossoming-forbes-111230225131949533.htm>
 - (c) Prime Minister's Decision No. 411/QĐ-TTg dated 31 March 2022 on Approval For The National Strategy For Development Of Digital Economy And Digital Society By 2025, Orientation Towards 2030
 - (d) Prime Minister's Decision No. No. 127/QĐ-TTg dated 26 January 2021 on National Strategy for Research, Development, And Application of Artificial Intelligence Until 2030
 - (e) Decision No. 1678/QĐ-BKHCN issued on 23 June 2021 by the MOST to promulgate the implementation plan of the "National Strategy on Artificial Intelligence research, development and application to 2030" ("**Decision 1678**")



Nguyen Thi Nhat Nguyet
Director, Attorney-at-law, Certified IP Agent
KPMG Law in Vietnam and Cambodia
T +84 28 3821-9266
nguyetnnguyen@kpmg.com.vn



Tran Bao Trung
Associate Director
KPMG Law in Vietnam and Cambodia
T +84 28 3821-9266
trungbtran@kpmg.com.ar

TDM and Gen AI

Copyright Boundaries in the Age of AI

Gen AI has made a significant progress and received a great amount of attention recently. AI models in this category have demonstrated the ability to generate human-like text and content quickly and with a minimum effort. This ability greatly depends on the vast amount of data on which it is trained. The more data the AI model has for training, the better results it may deliver. However, what if the data includes works protected under copyright or any other law?

On one hand, it is paramount for a successful Gen AI model to have a steady flow of diverse large-scale data. On the other hand, the larger and more diverse the data sets are, the bigger the probability that the data is protected by law. The level and kind of protection usually depends on what type of data is used for training and how it is acquired.

Increasing demand for data due to AI training gives attention to a popular method of acquiring large volumes of data – web scraping. This method serves to access and extract available internet data. If such data contains photographs, images, music, or other data which may be subject to copyright protection, acquiring such data and using it to train Gen AI models legally may pose a significant challenge for Gen AI service providers.

In general, if the copyright protection applies to data, the person who intends to use such data (e.g. reproduce the data) should acquire a permission from the right holder (author) or rely on one of the legal exceptions under which no further permission from the rightsholder is required. In this regard, a recent exception for text and data mining (TDM) adopted in the European Union plays an important role in terms of acquiring and using copyright protected data for training Gen AI models.

This exception was introduced by the EU Copyright in the Digital Single Market Directive 2019/790 (CDSM Directive). The CDSM Directive obliged Member States to transpose two exceptions to copyright protection for TDM in their national copyright legislation. The first exception is set out in Article 3 which allows research organizations and cultural heritage institutions to use TDM for scientific purposes without greater limitations.

The second exception under Article 4 further provides that the TDM exception may be used by all other entities (including commercial sector) provided that the right holders of the relevant works have not reserved the use of their work for TDM. The exclusion is based on the so-called “opt-out principle” and must be made in an appropriate manner, such as machine-readable means in

the case of content made publicly available online (for instance via metadata or terms and conditions).

At first glance, the TDM copyright exception could provide a fair amount of certainty for Gen AI developers when conducting their data scraping activity and AI model training. The existence of this exception ensures that the AI development would not be slowed down by a lack of data necessary for AI model training. At the same time, it gives the right holders an opportunity to opt-out and protect their works.

Compared to the US or other countries where the AI model providers widely rely on fair use doctrine, the providers in the European Union will probably not experience the same number of lawsuits as they currently do in the USA or other countries without specific TDM exception. However, scraping data from the internet for the purpose of AI model training should always be carefully considered in terms of whether the TDM exception applies in full and whether other kinds of data do not trigger protection under different laws (for instance personal data protection or contractual law protection).

To avoid lawsuits from the right holders, AI model providers are introducing various features which allow the right holders to choose whether they want their works to be subject of AI models training. Enabling these features is also motivated by an endeavor to increase confidence in the services based on Gen AI models. Ultimately, it may be the users’ trust or the lack of it that force the Gen AI service providers to acquire and use the training data in a fair and proportional manner that will show respect for copyrighted works and their authors. From this point of view, relying solely on the TDM exception may not be enough for a successful operation of the services based on Gen AI.



Martin Čapek

Associate
KPMG Legal s.r.o., advokátní kancelář
T +420 222 12-3967
mcapek@kpmg.cz



Eliška Houdková

Junior Lawyer
KPMG Legal s.r.o., advokátní kancelář
T +420 222 12-3808
ehoudkova@kpmg.cz

Copyright and Large Language Models

Large Language Models (LLMs) as a subset of Artificial Intelligence (AI) have arrived in the legal world since the start of ChatGPT and have triggered discussions in copyright law. LLMs are models that are trained using large amounts of data including texts with the goal of being able to predict the next elements of the text itself. Hereby, the LLM is able to calculate the probability of word sequences and to evolve them independently into texts in the next step. The answers (output) of the LLM are based on the most probable word sequence calculated from the words of the input (prompt). The central questions are (1.) whether the output of LLMs infringes copyrights of third parties, (2.) whether own copyrights arise from the output and (3.) whether copyright law allows the reproduction and storage of data for training LLMs.

1. Copyrights of third parties

First and foremost, it is important for users who are concerned about copyright infringement when using LLMs to understand that LLMs do not create plagiarized content. Unlike typical search engines, they do not locate existing texts and display them; instead, they generate entirely new texts. However, depending on the instructions provided by the user to the LLM, this could still potentially lead to a copyright violation. For instance, if the output is a derivative work based on copyrighted material, it may constitute a copyright infringement. The answer to the question whether the output of LLMs infringes copyrights of third parties therefore depends, among other things, on the copyright classification of the generated text. Different constellations have to be distinguished. In some constellations rights to the input may persist in the output.



Copyright and Large Language Models

a) Reproductions and translations

If a LLM is asked to reproduce a specific text like a song text and you receive a text that is not yet in the public domain, the copyrights of the original author are still valid.

The situation is similar if the LLM is asked to translate a copyrighted text into another language. In such cases the text can therefore not be exploited freely.

b) Rewritings and technical texts

If a LLM is asked to rewrite an existing text, the rights of the original author may also persist. Parts of a text or individual characters in a text can also enjoy copyright protection. Whether a copyright infringement exists or not depends on how much of the original can still be found in the output of the LLM.

If the output of the LLM consists of an informative technical text, a copyright infringement is also possible but less likely. Protection requirements for technical texts are relatively strict according to jurisprudence in Europe. Although a few consecutive words can express a separate intellectual creation, technical texts are essentially determined by the information they contain, and it is often not possible for the author to express creativity when drafting them. Furthermore, the output of LLMs does not consist of text fragments but is formulated anew by the AI.

2. Own copyrights

The answer to the question whether copyright law grants own copyrights regarding the output autonomously produced by LLMs depends on whether the output is a protected work or a related right according to copyright law.

a) Protected work

The question of whether content can be considered a protected work under copyright law must likely be answered in the negative in most cases, as content created solely by AI does not qualify for copyright protection as a work. This is due to the concept of a copyrighted work established in most legal jurisdictions. According to Section 2 subsection 2 of the German Copyright Act (UrhG) for example, works under copyright law are only personal intellectual creations. The creator has to be a human being, and the creation must result from a thought process that is also "personal." The creation has to be the result of a purposeful intellectual creative process. Random results, such as those arising from unintentional paint splatters or a photograph taken by a monkey playing with a camera, are not protected by copyright due to the absence of an intellectual creative process. According to German law, only a human with

human intelligence (HI) and not AI can be considered an author, and only they can create copyright-protected works. The crucial factor is that the author is genuinely free to make creative decisions.

However, when using LLMs the way we are already used to using them, no sufficiently creative decision by the user occurs at any stage (conception, execution, editing). The writing of the input may be a copyrighted performance but does not lead to protection on the output. The user usually has no significant influence on the actual machine execution, the actual production of the text. The further prompting and editing by the user will also generally not lead to protection.

However, there may be cases where a different assessment is justified, where the users employ and operate the LLM as a tool that merely executes their personal creative intent. This would be akin to using a paintbrush – if the brush merely rolls across the paper due to, for instance, being dropped, no copyright-protected work is created. However, if the painter deliberately wields the brush, a protected painting emerges. If AI is used in a manner analogous to how a painter uses a brush, a copyright-protected work can indeed be produced.

This immediately raises the question of who the author of this work is. Various solutions are possible. It could be solely the user, or it could be a joint effort between the user and the AI programmer. This question is sure to occupy copyright experts in different legal jurisdictions for some time to come.



Copyright and Large Language Models

b) Related rights

Related rights, such as the database manufacturer's right, can also be considered for protection of the output. However, the requirements for protection will often lack. Databases are collections of works, data or other independent elements arranged systematically or methodically and individually accessible by electronic means or otherwise. The prerequisite for this *sui generis* property right is a substantial investment, in terms of type and scope, in the acquisition, verification or presentation of the data. The production of the data does not yet lead to this investment protection. Therefore, the output does probably not meet the requirements of the database producer right. The same applies to press publisher law.

3. Text and Data Mining

A crucial question currently sparking controversial discussions in copyright law: Does copyright law, for example Section 44b of the German Copyright Act (UrhG), allow the reproduction and storage of data for training AI systems, especially LLMs, and how long can these data be retained?

Text and Data Mining is the automated analysis of individual or multiple digital or digitized works to extract information, especially patterns, trends, and correlations. To enable this, data must be stored, which constitutes reproduction, a right generally reserved for the copyright holder. According to Section 44b subsection 2 UrhG, reproductions of lawfully accessible works for text and data mining are permitted. However, these reproductions must be deleted when no longer needed for text and data mining.

The legal situation regarding this question is contentious in the literature, and there is no jurisprudence on the matter. Section 44b UrhG was created in its current form as part of the 2021 German copyright reform. Some argue that it does not apply to AI training at all, so the copyright exception does not apply here. This view should be rejected, not least because the legislative rationale for Section 44b UrhG explicitly cites AI training as a justification. It states that the use of copyrighted works should be expanded to promote innovation, particularly for machine learning as a foundational technology for AI. Therefore, we can assume that Section 44b UrhG allows the reproduction and storage of lawfully accessible data for AI training in general.

However, the question arises whether this can unequivocally apply to LLMs as well. The German copyright reform is based on the European Directive on Copyright in the Digital Single Market (DSM Directive) from 2019, a time when legislators had not yet contemplated LLMs. Text and data mining is essentially about extracting

information, so there is a question of whether the training of LLMs falls within its scope. In any case, one may argue that an analogous application is permissible until there is specific legislation addressing LLM training. The interests of copyright holders are adequately addressed as they can reserve and prohibit such usage.

Nevertheless, another question arises: How long can such training data be stored? Is there a time limit after which the data must be deleted, or does the justifying purpose continue as long as the AI is operational? This question has not yet received a definitive answer.

Some argue that the data may need to be deleted after the initial training to ensure the protection of copyright and prevent disproportionate infringement. Others contend that continuous storage and use of this data for AI development may be justified. It remains to be seen how legislation and jurisprudence in these areas will develop in the coming years.

4. Conclusion

The output of LLMs raises legal questions regarding the possible infringement of third party copyrights, possible own copyrights and the possibilities of reproduction and storage of data for training LLMs.

The answer to the first question is "maybe", because in some constellations rights to the input persist in the output.

The answer to the second question will probably be "no" in most cases, because the output created by AI alone cannot be considered as work from a human creator and related rights lack the necessary requirements.

The answer to the third question is that reproduction and storage of lawfully accessible data for AI training should be allowed because the interests of copyright holders are adequately addressed as they can reserve and prohibit such usage.



Dr. Anna-Kristine Wipper

Partner
KPMG Law Rechtsanwalts-gesellschaft mbH
T +49 30 53019-9731
awipper@kpmg-law.com



Dr. Thomas Beyer

Senior Manager
KPMG Law Rechtsanwalts-gesellschaft mbH
T +49 30 53019-9822
thomasbeyer@kpmg-law.com

AI Risks in Focus

Why comprehensive AI governance is essential for businesses

The development of AI-based applications is progressing at a rapid pace. Indeed, the capabilities of programs like the text generator ChatGPT are impressive. Though, it is not surprising ChatGPT has attracted media attention. However, what the global public is currently witnessing is just the beginning of a comprehensive technological revolution. Much of it is still in its early stages. The relevance and potential of Gen AIs are not yet foreseeable at present.

Even though the current hype remains unabated, it is essential to consider that this development is not without risks. Simplifying and automating routine decisions inevitably entails risks. How should these risks be managed in practice? Especially at the board and executive levels, it is crucial to inquire about the specific functions within companies that AI solutions are suitable for and the necessary conditions to be fulfilled. AI governance must become a central concern on the agenda of decision-makers to avoid compliance risks.

It is anticipated that the complexity of applications will rise in this field more than in conventional data and analytics

scenarios. Simultaneously, transparency and traceability of results are expected to decrease. Furthermore, the impact that autonomous systems have on processes and the shortening of decision-making paths must be closely monitored. Legal questions regarding data foundations and the use of outcomes remain inconclusive.

AI governance – risks, regulation, and evaluation of AI systems

To ensure the responsible and efficient use of AI-based solutions in companies, the board and executive management must establish comprehensive AI governance as an overarching framework for managing the risks of modern AI systems. The supervisory board ought to engage with this issue in a critical manner. Ultimately, it is necessary to design and implement a target operating model with responsibilities and processes for the required AI governance.

Understanding how AI-based systems work

Incidents related to the use of AI systems reveal the associated risks clearly. Who is responsible for fatal accidents caused by autonomous vehicles? Who is accountable if the performance of new technology is overestimated?



And what image and reputation damage can result from faulty systems? For instance, a business that had implemented an AI recruitment system scanning CVs to find the most qualified candidates terminated its use following a short trial period due to the discrimination that the AI showed towards women.

The reason for such malfunctions lies in the functioning of the underlying algorithms. These algorithms need to be trained for their specific use cases. This training often relies on historical data, which could already contain biases, be flawed, or subject to copyright protection. In the case of the recruiting AI, it was trained based on applications from the past ten years, most of which came from men, causing the algorithm to learn that the "male" gender trait was a good hiring criterion^(a).

Another example is a chatbot that had to be shut down after just 16 hours because it adapted its language use based on user feedback, leading to suddenly using racist and discriminatory language. The issue at hand was that the chatbot's "adaptation" took place without any controlling component^(b).

Identification and classification of AI risks

This practical insight illustrates why AI governance fundamentally differs from conventional IT governance. When it comes to monitoring traditional IT systems, the binary question is merely, "Is the server/system running or not?" With AI systems, it is crucial to determine not only, "Is the AI system running?" but also to clarify, "What is its content?", "What direction is it taking?", "Has it developed biases?" Therefore, specific AI risks must be identified and classified accordingly. These include:

- **Outage Risk:** Is there a comprehensive AI disaster recovery plan in place? Does monitoring cover the entire machine learning or data science lifecycle? What are the consequences if the AI falls victim to a cyberattack and needs to be shut down?
- **Information Risk:** Do the outputs of AI systems align with reality? Does the model need to be adapted to real-world conditions? Does the model behave in the real world as it did during training in the "lab"? How does the AI's learning path evolve in reality? How quickly does the AI develop biases when it reinforces its own decisions?
- **Financial Risk:** Are the costs associated with the development and deployment of AI systems justified, considering that they can be quite high? How can these costs be capitalized? What is the permissible transfer pricing when implementing a centralized AI that operates across a corporation's global network? What functions cause the central AI to provide

feedback and utilize elements itself within the global application?

- **Liability Risk:** Who is responsible for decisions in which AI systems are directly or indirectly involved? What penalties can be expected for insufficient AI governance?
- **Reputation Risk:** Do all use cases of AI systems adhere to ethical standards? Are certain groups discriminated against by the AI?
- **Data Risk:** Is the data (from input to output) processed in compliance with regulations? What is the source of the data (ownership vs. possession)? Are all data protection regulations being followed?
- **IP Infringement Risk:** What copyrights or other intellectual property rights must be considered when using data for the training of the AI? Are there intellectual property rights, such as copyright, associated with the results, and if so, who holds these rights?

Given the broad spectrum of risks, it is essential to take a comprehensive and unified approach to the entire construct of an "AI system", considering all aspects involved, ranging from the use case to the data and models up to the systems and interfaces used.

Developments on the regulatory front

Legislators have also recognized the risks associated with AI systems and have already responded with corresponding regulations. At present, there are four significant developments in the field of legislation, with laws and initiatives being passed or underway at both national and EU levels, as well as through international agreements.

• The "EU AI Act"

From a corporate perspective, we think that the "EU AI Act" is of high relevance. This is a comprehensive EU-wide approach to regulating AI applications. The greater the degree of risk posed by an AI system, the more significant the accompanying responsibilities and obligations become. Failure to comply with these requirements can lead to financial penalties of up to €30 million or up to six percent of the total global annual turnover for the previous financial year, whichever is greater. The "EU AI Act" is expected to come into effect in 2024 and can be compared in

- Note:
- (a) Source: Wilken, Felicitas: Bewerbungsroboter: Künstliche Intelligenz diskriminiert (noch), Zeit online, 18.10.2018
 - (b) Source: Graff, Bernd: Rassistischer Chat-Roboter: Mit falschen Werten bombardiert, Süddeutsche Zeitung digital, 03.04.2016

AI Risks in Focus

terms of scope and penalty to the General Data Protection Regulation (GDPR). In addition, current liability gaps in the use of AI systems are expected to be addressed through a new AI liability directive and an update to the product liability directive. The liability risk for the distribution and use of AI systems is expected to increase due to causality presumptions and evidentiary facilitations in favor of victims. Compliance with the "EU AI Act" will be even more critical. Although alterations to the current draft regulation are expected, the principal framework of the regulation will remain unchanged.

- **The Directive on Copyright in the Digital Single Market and the German copyright law reform**

The Directive on Copyright in the Digital Single Market (DSM Directive) ensures that text and data mining is

enabled in all European Union member states. It establishes a legal framework for commercial text and data mining to stimulate innovation, including AI training. In Germany, the Directive was implemented through the copyright reform that came into effect in 2021. According to Section 44b subsection 2 of the German Copyright Act, reproducing legally accessible works for text and data mining is now allowed also for commercial purposes. However, copies are to be deleted once the objective is achieved and they are no longer needed for conducting text and data mining.

- **The IDW EPS 861**

The IDW EPS 861 provides the relevant AI framework from Germany. Although it is currently in draft form, it has been declared a voluntary standard for examining artificial intelligence. It serves as a good preparation



for the "EU AI Act" as it helps companies implement AI systems in compliance with current law. The basis for EPS 861 is the ISAE 3000 "Assurance Engagements Other than Audits or Reviews of Historical Financial Information," an acknowledged and tested standard with room for adjustments.

- **The Global Partnership on Artificial Intelligence (GPAI)**

On an international level, the "Global Partnership on Artificial Intelligence" (GPAI) emerged in June 2020. The central goal of GPAI is that AI must be developed in harmony with human rights and democratic values to ensure public trust in the technology, as outlined in the OECD Principles on Artificial Intelligence (2019). The specific integration and interpretation of these principles are implemented through the legal systems of each respective country.

Evaluation of an AI system under the "EU AI Act"

The obligations of the "EU AI Act" apply to manufacturers, providers, and sellers of AI systems, product manufacturers incorporating AI systems into their products, as well as users of AI systems. This will potentially encompass nearly every organization in the future. The initial stage is to evaluate, within the framework of the existing regulations of the "EU AI Act," whether an AI system is operational within the company, its risk classification, and guaranteeing compliance with all responsibilities.

- Example criteria for risk assessment include^(c):
- Evaluation of the purpose of use
- Scope of use (local to international)
- Scope of affected individuals (few to many)
- Type of data (non-personal to sensitive personal data)
- Degree of automation (manual to automatic)
- Complexity of algorithms (low to high)
- Overall complexity of the AI system (low to high)
- Individuality (custom development to off-the-shelf software)
- Level of autonomy in AI decision-making (low to high)
- Use of results (informative to strategic)

Based on the risk assessment, companies must classify their AI systems into one of three risk classes: "unacceptable," "high," or "low/minimal."

- **Risk Class "Unacceptable":**

The marketing, deployment, or use of AI systems that

pose an unacceptable risk is prohibited. This includes especially AI systems that are designed specifically to subliminally influence human behavior detrimentally or to take advantage of vulnerable individuals, including children or those who are physically or mentally disabled. Additionally, the use of AI systems for assessing or classifying the trustworthiness ("social scoring") of natural persons by or on behalf of authorities is prohibited. The use of AI systems for biometric real-time remote identification of natural persons in publicly accessible spaces for law enforcement purposes is largely restricted and only permissible under limited exceptions.

- **Risk Class "High":**

AI systems that are deemed to pose significant risks to the health and safety or fundamental rights of individuals, particularly when used as product components, are commonly known as "high-risk AI systems." The extent of the negative impact of the AI system on the rights protected by the Charter of Fundamental Rights of the European Union is of particular importance when classifying an AI system. These rights include human dignity, respect for private and family life, personal data protection, freedom of expression and information, freedom of assembly and association, non-discrimination, consumer protection, workers' rights, the rights of people with disabilities, the right to an effective remedy and a fair trial, the presumption of innocence, the right to defense, and the right to good administration. It should be emphasized that children have distinct rights beyond these.

"High-risk AI systems" include, in particular, AI systems that are:

- used as safety components of a product or themselves constitute a product subject to the regulations listed in Annex II of the "EU AI Act," such as the Machinery Directive or the Toy Directive,
- themselves or as safety component subject to third-party conformity assessment as specified in the regulations listed in Annex II of the "EU AI Act," or
- listed in Annex III of the "EU AI Act." This encompasses specifically AI systems for biometric identification of individuals and AI systems utilised in an employment setting, particularly in relation to selection decisions, monitoring of employees, or personnel management.

Note: (c) Source: KPMG Germany, 2023

AI Risks in Focus

The "EU AI Act" imposes stringent requirements on the design and use of high-risk AI systems. These requirements include:

- Quality requirements for data governance
- High standards for accuracy and security, including cybersecurity
- Transparent operation and interpretability of results
- Technical documentation and automated operational logging
- Operation of quality and risk management systems and compliance with market surveillance obligations
- Extensive recording, documentation, and logging obligations
- Ensuring human oversight capabilities

High-risk AI systems should generally display the CE marking as proof of compliance with the "EU AI Act." This also includes the intended registration of high-risk AI systems in a European Union database. Serious incidents and malfunctions of high-risk AI systems are expected to be subject to mandatory reporting in the future.

• Risk Class "Low/Minimal":

If AI systems are not classified as unacceptable or high-risk AI systems, they fall into the third category. These AI systems are subject to less stringent requirements. Providers of such systems are encouraged to establish codes of conduct and voluntarily comply with regulations for high-risk AI systems.

Additionally, the "EU AI Act" requires that even these AI systems must adhere to safety standards if they are placed on the market or put into operation. From a company's perspective, it is possible to demonstrate the necessary level of safety if the regulations for a high-risk AI system are willingly adopted. Therefore, it is advisable to actively evaluate and incorporate every AI system into a governance structure based on the risk pattern of the AI system.

AI governance – what you need to consider

As AI becomes more prominent in applications, it must be integrated into corporate governance. For you, this means that you must consider all AI-based solutions in use comprehensively as "AI systems" and understand their use cases and associated risks. It is crucial to ensure that, as an end-product manufacturer, you meet the provider obligations outlined in the "EU AI Act".

Specifically, you must ensure that the embedded AI system fulfils the stipulated requirements. Risks also include liability risks arising from regulations, particularly in terms of deadlines and penalties. To meet all requirements, establishing AI governance is essential.

The key issue here is determining who bears the responsibility for risk classification. Is it a self-assessment, which could be subjectively biased, or a more objective third-party assessment? Another aspect frequently disregarded concerns the multifaceted composition of the team responsible for the evaluation.

Moving from risk management to unlocking the full potential of AI

When it comes to practical risk management solutions, the need to establish corresponding guidelines, processes, and monitoring solutions for risk mitigation is essential. Various institutions and organizations, such as BSI, IDW, or DIN, are already developing standards.



Dr. Justus H. Marquardt
Partner, Consulting
KPMG AG Wirtschaftsprüfungsgesellschaft
T +49 211 475-7263
jmarquardt@kpmg.com



Francois Heynike, LL.M. (Stellenbosch)
Partner
KPMG Law Rechtsanwalts-gesellschaft mbH
T +49 69 951195-770
fheynike@kpmg-law.com



Oleg Brodski
Partner
KPMG AG Wirtschaftsprüfungsgesellschaft
T +49 211 475-7621
obrodski@kpmg.com



Dr. Anna-Kristine Wipper
Partner
KPMG Law Rechtsanwalts-gesellschaft mbH
T +49 30 53019-9731
awipper@kpmg-law.com

The automatic generation of images by AI

Legal issues with regard to IP and AI

Nowadays anyone can quickly generate unlimited images, reproducing or imitating artistic styles or art works, due to a new range of different platforms and apps using a combination of machine learning, user-uploaded images, and/or written instructions. Moreover, Gen AI themselves can create digital images or art models by extracting existing text-to-image based on natural language text inputs. By way of example, a recent cover of a famous magazine has been created by one of the leading image generation engine AI's who is able to paint and create new images in a photorealistic style and, such illustrations or images, generated by text-to-image, are starting its commercial exploitation.

A number of legal and ethical issues have arisen due to the unclear current approach over the kind of protection to be granted to AI generated images and digital artworks complicated by the different approaches taken internationally and an apparent lack of reciprocity between countries, mainly due to the non-human factor. The legal and commercial challenges facing the creative

and other industries include new issues to be addressed as, how much is AI's "borrowing" from existing copyrighted works, going beyond the boundaries of inspiration, as well as, is the AI output exploitable and if so, by whom and under which terms and conditions.

In one of the first cases involving non human creations, the well known "monkey selfie copyright", the US Court of Appeal addressed the issue of the copyright status of "selfies" taken by macaques using equipment set up by a human photographer, finding that under US law only natural or legal persons are entitled to be listed as authors of copyright works and that animals are not entitled to own copyright or sue for infringement. Whether a work may qualify for copyright protection and in which territory may also depend on a certain status of the author, as its nationality, residency or citizenship, and in case of a company, on the place of incorporation. The Authorship status will be granted to the person or persons that have creatively contributed to the output. When AI systems are used to create content, co-authorship claims by AI developers will also be likely to arise.



The automatic generation of images by AI

Under Italian Copyright Law (law no. 633 of 22 April 1941, along with its various amendments), a work of authorship also needs to meet the originality requirement, and this is met when there is an expression of the authors' own creation, implying a degree of human creativity in the work reflecting the personality of the author and its creative input and choices. The uncertainty on whether AI created images meet the "originality" criteria derives from the AI systems' use of mechanic and random methods based on algorithms rather than on creative choices on the shape, the color or other.

The leading scholars have argued in favor of granting copyright protection to AI independent creations taking into consideration: (i) the degree of investment and economic effort vested in the development of the underlying algorithm, and (ii) the fixation on a tangible form of expression requirement. AI text-to-image generation seems gone beyond cases where the system is being used as the mere tool of a human author and, to be able to generate non existing images, AI must receive guidance and training to process existing images and to

identify patterns which predefine how much importance is given to certain parameters in the model, with a huge amount of data required. Among these high quality data a number of images or texts are likely to be protected by copyright requiring license for use for commercial purposes, and this new situation is leading to an increase in the number of legal complaints being raised. Among recent cases, a photography platform and image licensing company announced it was commencing UK legal action for copyright infringement against an AI system, alleging stock images owned or represented by the licensing company had been unlawfully accessed and processed without authorization. In the US, several artists have launched a claim against two primary AI systems, alleging the companies have used their copyright artwork to train AI models without consent.

The outcome of such cases will be interesting to define the limits of using copyright images and data in AI training without permission and without payment to the copyright owner.



Legal issues on the automatic generation of images by AI Systems

The UK government has anticipated its approach in permitting parties to mine datasets and use them to train AI models, thereby creating exemptions to infringement of copyright. However, this could not prevent copyright owners from claiming their rights.

At this point, the issue to be addressed is the one concerning the question if AI generated contents shall consequently fall into public domain allowing any kind of reproduction. A recent decision involving the use of an image generation AI model, pre-programmed to create images, in an authorial book issued by the US Copyright Office, denied registration to AI-generated images within a comic work, ruling that as they were created by a Gen AI tool through a mechanical process, and not by a human, they were not entitled to copyright protection. As a consequence, the US Copyright Office, canceled the original certificate of registration of the novel and issued a different one covering only the materials created by the author. The policy of the US author does not imply that technological tools may not take part of the creative process, however the extent of human input, guidance and selection within the creative process and expression used to create the work, may determine registrability (21 February, 2023- ID 1-5GB561K).

On the other side, Italy appears to have reached a contrasting approach in a recent decision by the Italian Supreme Court, in favor of copyright protection. The Court found that the use of software shall not automatically exclude the processing of an intellectual work protectable under copyright law, however in such case a more rigorous factual assessment creativity test would need to be conducted in order to identify any creative human contribution to the AI creative process (Case n. 1107/2023, 16 January 2023). The subject matter of this case was copyright protection on a scenography named "The scent of the night" used during an Italian popular music festival broadcasted internationally, where such scenography had been created through a software program able to process its colors, details and shapes by using set algorithms reproducing the images on different scales, together with a process guided by the author that was limited to the programming of the algorithm and approval of the software generated results. This case reformed the previous Italian leading case law in finding that the legal concept of creativity should not be mistaken with different concepts as originality and novelty, this as the creativity threshold may be met also in the event of a work including the expression of the authors' individuality, even if to a minimum degree. In the same case, the previous Court of Appeal decision had found that the main image in the scenography did not consist in mere reproduction of a flower, but it incorporated a re-elaboration of the work

in which the personality of the artist was expressed.

In conclusion, the EU copyright framework appears sufficiently flexible to deal with the current challenges arising from AI-assisted creation and producers of AI-assisted output will in many cases be entitled to copyright protection, which is likely to be invoked in the event the user of the AI provides evidence proving that the AI system represents a tool, or a moment, within an overall more complex creative process, while the non-authorial output might still qualify for protection under unfair competition, trade secrets, sui generis database protection, and other, in accordance to the facts.

The core issue being whether the AI-assisted output is the result of human creative choices "expressed" in the output. Three different phases of the creative process can be distinguished: phase one being the "conception", involving design and specification and where the role of human authors often is essential; phase two being the "execution" or the producing of draft versions in which it is the AI systems playing a dominant role; and finally the "redaction" phase involving selection, editing, refinement and finalization, often lead by human beings. Assuming creative choices are expressed in the final output, even if it has been AI-assisted, such output will then qualify as a copyright-protected work depending on the facts. While, if an AI system is programmed to automatically execute content without human contribution, this will not be entitled to become a "work of authorship".

In the light of the above, before AI images are put into use, or are commercially exploited, the user of AI text-to-image generators should consider how to ensure ownership of copyright, develop strategies on how to react and prevent infringement, and verify rights to use the images internationally and in accordance with local laws. Another aspect to be kept in consideration is ensuring that the license, permitting use of the AI text-to-image tool, addresses copyright correctly, since this may impact on copyright first ownership in generated images and also on the nature of rights granted to users. In general, where AI images are being used as part of projects involving or commissioned by third parties, the contract with third parties should consider the legal status of AI images and the issues outlined above.



Marina Lanfranconi
Principal
Studio Associato
Consulenza legale e tributaria
T +39 02 6764-4308
mlanfranconi@kpmg.it



kpmg.com

Throughout this document, "we", "KPMG", "us" and "our" refers to the global organization or to one or more of the member firms of KPMG International Limited ("KPMG International"), each of which is a separate legal entity.

KPMG refers to the global organization or to one or more of the member firms of KPMG International Limited ("KPMG International"), each of which is a separate legal entity. KPMG International Limited is a private English company limited by guarantee and does not provide services to clients. For more detail about our structure please visit kpmg.com/governance.

The information contained herein is of a general nature and is not intended to address the circumstances of any particular individual or entity. Although we endeavor to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act on such information without appropriate professional advice after a thorough examination of the particular situation.

© 2023 Copyright owned by one or more of the KPMG International entities. KPMG International entities provide no services to clients. All rights reserved. The KPMG name and logo are trademarks used under license by the independent member firms of the KPMG global organization.