Intellectual Property for Business Series
No. 7

Learning Machines

An introduction to AI and IP for small and medium-sized enterprises

WIPO

Learning Machines

An introduction to AI and IP for small and medium-sized enterprises

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

This work is licensed under Creative Commons Attribution 4.0 International. To view a copy of this license, please visit https://creativecommons.org/licenses/by/4.0.

The user is allowed to reproduce, distribute, adapt, translate and publicly perform this publication, including for commercial purposes, without explicit permission, provided that the content is accompanied by an acknowledgement that WIPO is the source and that it is clearly indicated if changes were made to the original content.

Suggested citation: World Intellectual Property Organization (WIPO) (2026). *Learning Machines: An Introduction to AI and Intellectual Property for Small and Medium-sized Enterprises.* Geneva: WIPO. DOI: 10.34667/tind.60221.

Adaptation/translation/derivatives should not carry any official emblem or logo, unless they have been approved and validated by WIPO. Please contact us via the WIPO website to obtain permission.

For any derivative work, please include the following disclaimer: "The Secretariat of WIPO assumes no liability or responsibility with regard to the transformation or translation of the original content."

When content published by WIPO, such as images, graphics, trademarks or logos, is attributed to a third party, the user of such content is solely responsible for clearing the rights with the right holder(s).

Any dispute arising under this license that cannot be settled amicably shall be referred to arbitration in accordance with Arbitration Rules of the United Nations Commission on International Trade Law (UNCITRAL) then in force. The parties shall be bound by any arbitration award rendered as a result of such arbitration as the final adjudication of such a dispute.

The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion whatsoever on the part of WIPO concerning the legal status of any country, territory or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

This publication is not intended to reflect the views of the Member States or the WIPO Secretariat. The mention of specific companies or products of manufacturers does not imply that they are endorsed or recommended by WIPO in preference to others of a similar nature that are not mentioned.

Cover: Getty Images/MF3d; monsit
WIPO Publication No. 1090/26

© WIPO, 2026
First published 2026

World Intellectual Property Organization
34, chemin des Colombettes
P.O. Box 18 CH-1211
Geneva 20
Switzerland

ISBN: 978-92-805-3963-9 (print)
ISBN: 978-92-805-3962-2 (online)
ISSN: 2958-4620 (print)
ISSN: 2958-4639 (online)

Contents

Acknowledgments 6

Introduction 8
The intersection between IP and AI 8
Overview of this guide 9

What is AI? 10
What is machine learning? 11
What is a neural network? 12
What is generative AI? 13
What is a large language model? 13

The AI and IP landscape 14
What are the major IP issues driven by AI? 14
What are current major legislative and other developments around IP and AI? 18

Developing and protecting AI systems 23
Can I patent my AI system? 23
Can I use copyright to protect my AI model or system? 24
Can I use trade secrets to protect my AI innovation? 25
What is the role of open-source licensing in IP and AI? 26
Can I use copyrighted materials to train an AI model? 27

Creating and inventing with AI 30
Can I patent an invention created using AI? 30
Can I copyright an AI-generated work? 33
What should I look for regarding IP in AI terms of service? 37
What disclosure obligations do I have when using an AI-assisted creation? 39
How can I demonstrate originality in an AI-assisted work? 40
How can I protect my business's IP when using AI tools? 40
How can I prevent my business's IP from being used to train an AI model? 41

The future of IP and AI 42
Outstanding questions and policy challenges 42
Innovation through collaborative frameworks 43

WIPO resources 45

Acknowledgments

An initial version of this guide was prepared by Professor Ryan Abbott (Professor of Law and Health Sciences at the University of Surrey School of Law and Adjunct Associate Professor of Medicine at the David Geffen School of Medicine at University of California, Los Angeles) with the assistance of Kasiet Kachaganova (WIPO) in 2018 and further revised by him in 2022.

The text was further updated, revised and finalized by Dr. Alina Trapova (Lecturer, Faculty of Laws, University College London) in 2024.

Structural and line editing was provided by Marianne Cassidy (Content Consultant) in 2025.

Draft manuscripts received valuable comments from the following WIPO colleagues:

- Patent and Technology Law Division: Andras Jokuti (Director), Tomoko Miyamoto (former Head, Patent Law Section), and Nina Belbl (Legal Officer)
- Building Respect for IP Division: Xavier Vermandele (Senior Legal Counsellor) and Thomas Dillon (former Legal Counsellor)
- Copyright Law Division: Michele Woods (Director) and Rafael Ferraz Vazquez (Legal Officer)
- Frontier Technologies Division: Ulrike Till (Director) and Ana Andrijevic (Fellow)

An early version of the draft was reviewed by the following external experts:

- Professor Jacques de Werra (Professor of Intellectual Property Law, University of Geneva)
- Professor Stavroula Karapapa (Professor of Intellectual Property and Information Law, School of Law, University of Essex)

The guide was edited and designed by WIPO's Digital Engagement Service, led by Charlotte Beauchamp (Head of Publications and Design).

The project was managed by Tamara Nanayakkara (Counsellor, IP for Business Division) under the supervision of Guy Pessach (Director, IP for Business Division).

Introduction

Artificial intelligence (AI) is experiencing a period of unparalleled growth, driven by improvements in software, computing power, and big data. AI technologies are now poised to radically transform industries such as health care, finance, transportation, art, music, and many others with dramatic economic and social impacts already being felt. Beyond this, AI is transforming the human creative process itself and in doing so raises fundamental challenges to certain well-established tenets of intellectual property (IP) law.

Any company, big or small, developing or using AI is a key part of this transformation. The impact of AI that is developing will be felt across all industries horizontally, whether or not they are actively engaging with this technology at the moment.

The intersection between IP and AI

IP is an important source of business value for many companies. As such, it is critical for companies both creating and using AI to have a thorough understanding of how these technologies intersect with IP. Businesses should be asking themselves questions such as:

- How can IP protect AI software?
- Should IP strategies focus on software, hardware or even data?
- Is it possible to use third-party IP to train AI without permission?
- Is it an infringement for AI to collect and use third-party IP?
- Should, and if so, how can AI be licensed?
- Can AI outputs receive IP protection?
- Can companies circumvent certain IP laws via contracts?

This guide will attempt to answer these questions, with the strong caveat that this current phase of AI technology is evolving rapidly, as are the laws, regulations and litigation emerging in response to it.

AI raises other issues that go beyond IP, including personal data protection, cybersecurity, biased data, technological unemployment and disinformation are not the focus of this guide, though they may intersect with IP concerns at times.

It should also be noted that this guide is not intended to provide legal advice, much of which is jurisdiction dependent, but to help business stakeholders gain a broad understanding of the current state of AI and the potential implications for their IP strategies.

Overview of this guide

What is AI? introduces the main technologies and concepts that are captured under the umbrella term "AI", specifically: **machine learning**, **neural networks**, **generative AI**, and **large language models**. It is not necessary to dissect the technology behind AI in full but a basic understanding of these terms and systems is useful for grasping the IP implications.

The AI and IP landscape provides an overview of the major IP tensions driven by AI and the main developments in jurisdictions where we are seeing legislative reforms, consultations, and litigation. Certain jurisdictions are particularly active often driven by the fact that major AI companies are headquartered in these places.

Developing and protecting AI systems is addressed to businesses that are developing AI models or products, focusing on patents, copyright, and trade secrets. The goal is to help businesses understand their options for protecting their AI innovations under existing IP frameworks and their rights and responsibilities regarding data protection and copyright.

Creating and inventing with AI provides insights for businesses that are using AI tools to invent and create, with a focus on patents and copyrights. We will also discuss what to look for in the terms of service of an AI product regarding IP, and how to protect business IP when using AI tools.

The future of IP and AI acknowledges that the continued rapid advancement of AI technologies, and the increasingly widespread availability of AI tools, leaves us with many outstanding questions. We conclude by addressing some of these, and highlighting both the challenges and opportunities they pose for businesses, IP rights holders, AI developers and providers, trade bodies, governments, universities, licensing organizations and the public.

What is AI?

The term “artificial intelligence” or “AI” has perhaps inevitably become something of a buzzword. It is most frequently used as an umbrella term to refer to a collection of related technological processes and tools, including **machine learning**, **neural networks**, **generative AI (GenAI)** and **large language models**.

This guide adopts the definition of AI from *Getting the Innovation Ecosystem Ready for AI: An IP policy toolkit* (WIPO, 2024):

“Artificial intelligence (AI) refers to the branch of computer science and engineering that focuses on creating systems capable of performing tasks that typically require human intelligence. These tasks include understanding natural language, recognizing images, making decisions and learning from data”.

It is important to distinguish between **artificial general intelligence (AGI)**, and **specific artificial intelligence** (sometimes called "narrow" or "weak" AI). AGI refers to an AI system capable of matching human level intelligence across all domains, performing *any* task that could be performed by a human, and applying reasoning and problem-solving to entirely new situations without training. Most independent researchers see AGI as theoretical for now. By contrast, specific AI can perform a particular task or narrow group of tasks to a similar standard that a human could, and in some cases consistently outperforms human experts in certain domains. Specific AI is already here and widely used for specialized tasks in many fields.

Even though AGI is not here yet, specific AI still challenges how we think about AI as a creative actor, resulting in many of the IP tensions explored in this guide. Traditionally, AI has been understood as a tool operated by a human, who remains the author or inventor behind the output. If a human uses a hammer to build a table, we do not ask if the hammer is the creator of the table. But when an AI system can generate a novel or a drug compound or a piece of music with minimal human input or direction, the question of who, or what, created it

becomes difficult to answer. This is what makes the question of AI inventorship and copyright in AI-generated works so complex.

What is machine learning?

The term "machine learning" was coined by computer scientist Arthur Samuel in 1959. He defined it as a field of study that gives computers the ability to learn new things through example and experience, as opposed to being explicitly programmed. While machine learning techniques have been around for many decades, three big developments have brought about a visible revolution in this area in recent years. These are:

- the growth and widespread availability of massive digital data sets, i.e., "big data";
- unprecedented leaps in computing power; and
- advancements in computer science.

Big data are the biggest contributor to recent advancements. Machine learning processes feed on data: the more data and the better its quality, the more accurate the systems and their outputs become. For example, a human, or even a team of humans, may not be able to make much sense of tens of millions of purchase records from a retail chain. But these data can be fed into a computer system and, through machine learning, used to extrapolate sales trends, create inventory reports, detect theft and automate labor. The more purchase records the system has access to, the more accurate its outputs will become.

In the past, computer programs relied heavily on code produced by a human programmer, who would pre-define the features of the sought outcome. This was a time consuming and imprecise process, requiring a lot of computer power. But now, big data, coupled with rapid increases in computing power, have resulted in a genuine shift in algorithms: computational systems can learn and improve as they are exposed to many examples and no longer rely on *a priori* definition by a programmer of the sought outcome.

An important caveat here is that not all big data are created "equal." The use of biased and inaccurate data in machine learning can result in biased and erroneous outputs. If a machine learning system is trained on out-of-date, unreliable, or unverifiable data (in some cases, to minimize costs), the results they produce will also be unreliable and out-of-date, and their application in particular fields (such as medicine, transportation or finance) could lead to severe and even fatal consequences.

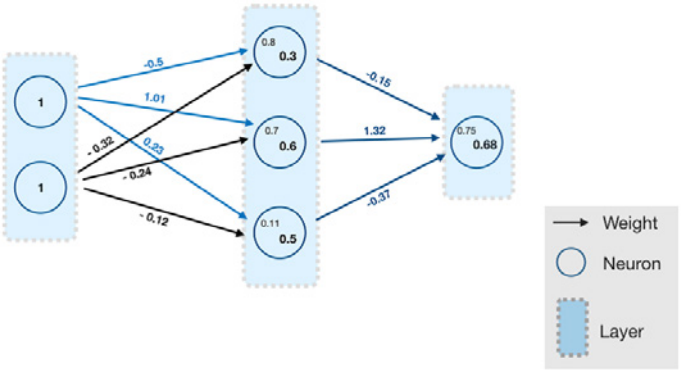
What is a neural network?

In 2019, the Max Planck Institute (MPI) for Innovation & Competition engaged in a comprehensive study of the technical aspects of AI from an IP law perspective. The study identifies the "trained machine learning model" (for short, the "model") as foundational to what we understand as "AI". These models are based on mathematical functions and generate outputs through the process of learning patterns in the data they are exposed to.

An artificial neural network is a specific type of trained model that draws inspiration from the structure and operation of the human brain to acquire patterns and representations from data. These models follow the connectionist logic of neuroscience, where neurons organized in networks establish associations according to the strength of the synapse connection. Similarly, neural network models consist of nodes, or neurons (which are mathematical functions) connected by "weights" (which represent a numeric value). "Deep" neural networks are arranged into many layers through which they can process data to produce non linear outputs.

To create an artificial neural network, the architecture must be developed by a programmer prior to the entire training process. This architecture is fixed and does not evolve during the machine learning process, and so is sometimes called a "hyperparameter." However, the weights connecting the neurons are adaptable or "trainable" parameters, meaning they can evolve autonomously as they learn from the data, rather than being preprogrammed.

Artificial neural network architecture



Source: Drexl, J. *et al.*(2019)."Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective".

What is generative AI?

GenAI, refers to AI tools capable of generating new content based on a user's prompt. A prompt is usually a short, written description of the desired output. The new content can take the form of text, computer code, images, audio, sound, and video.

The widespread availability of GenAI tools has significant IP implications across many industries, presenting both opportunities and risks. Crucially for IP purposes, GenAI models do not kick off on their own: they must be prompted by a human in a certain way in order to generate an output. Indeed, "prompt engineering" (the process of creating and refining input instructions to cause GenAI models to produce better outputs more quickly) has emerged as a distinct skill that requires an understanding of an AI model's capabilities and limitations.

What is a large language model?

Large language models (LMMs) are a type of neural network model. LLMs power many of the popular GenAI tools available on the open market today. They are so named because they are well suited for processing and correlating language. LLMs can perform various natural language processing tasks, including text generation, language translation, text summarization and more.

LLMs are not confined to processing human language. In machine learning terms, "language" refers to any system of symbols (such as words or code) that conveys meaning through structure and relationships. Just as human language derives meaning from both word choice and grammatical relationships, programming languages follow similar patterns of syntax and context. This is why some LLMs can generate computer code in common programming languages like Python or JavaScript. To effectively capture these complex relationships across any type of language, LLMs are designed to process sequences, handle large data sets and maintain sufficient contextual memory.

Launched in November 2022 by OpenAI, ChatGPT was the LLM chatbot that brought AI to the forefront of public consciousness, quickly transforming how people in many fields work and create. ChatGPT uses natural language processing and deep machine learning to generate written content in response to user prompts. While it was the first of its kind to achieve mainstream adoption, competitors such as Anthropic, Google and Meta released similar chatbots soon after.

The AI and IP landscape

What are the major IP issues driven by AI?

The rapid advancements in AI systems and widespread availability of AI tools have raised many challenges to some of the most fundamental tenets of IP. These are summarized below. Later, this guide addresses them in greater detail, with a specific lens on business concerns.

Copyright owners vs. GenAI developers

GenAI models require vast amounts of training data. For example, OpenAI's original GPT-3 model was trained on 570 gigabytes of text. In some cases, large portions of these training data sets were scraped from the internet and contain copyright-protected text and images. This has raised issues of potential copyright infringement. As such, there are two main conflicting viewpoints regarding the balance between the interests of copyright owners and the need for GenAI developers to use copyrighted works to train their models.

The developers of GenAI models sometimes argue that the models do not keep a permanent copy of any copyright-protected work and only access the information contained in the materials temporarily: in this way, they argue that the use of copyrighted training data sets is like a human learning by reading books or viewing artworks. When a copy is made, either temporarily or permanently, developers argue that those copies fall under existing copyright exceptions and limitations, and therefore no infringement occurs.

Many copyright owners, on the other hand, assert that unauthorized use of their works as part of these training data sets is copyright infringement. Several lawsuits have been initiated by copyright owners against AI developers, but clear answers to these complex legal questions are unlikely to emerge in the immediate future. A particular point of contention is that there is no easy way to assess how much a single work contributed to training an AI model and how to

potentially compensate the copyright owner. Also, copyright registration is not compulsory in many jurisdictions, making the identification of copyright works difficult at times.

Copyrighting GenAI outputs

GenAI can produce an extensive array of creative outputs, including text, audio, video, and images. There is an ongoing international discussion on whether these AI-generated outputs should benefit from copyright protection.

For example, a high-profile case in the United States of America involved an application for copyright protection for a work entitled "Théâtre D'opéra Spatial", which was created almost entirely using the AI image generator Midjourney. This application was rejected on the basis that the human creative input was *de minimis*. However, in a more recent case centered around a work called "A Single Piece of American Cheese, the author was able to demonstrate sufficient human intervention combined with AI tools. The US Copyright Office was satisfied that the work met the requirements for registration.

The "Théâtre D'opéra Spatial" was generated in 2022 by Jason M. Allen using the AI image generator Midjourney.

A landmark case that came before the Beijing Internet Court in 2023 concluded that an image generated by Stable Diffusion (another AI image generator) counted as an original work because the human prompts required to create it amounted to sufficient human creativity. This decision contends that GenAI can be used as a tool for humans to fuel creativity, and that if a human creator is using GenAI as a tool while also making a significant creative contribution, the resulting work could potentially be awarded copyright protection. How to measure such contributions and establish a threshold remains an open question.

AI and inventorship

Can an AI system hold a patent? Can AI be credited as the inventor of a new technology or tool?

DABUS ("Device for the Autonomous Bootstrapping of Unified Sentience") is an AI system created by Dr. Stephen Thaler that reportedly conceived of two inventions without human input. Thaler filed patent applications for these inventions, listing DABUS as the inventor, around the world. Most patent offices have rejected the DABUS applications on the grounds that patent laws require the naming of a human inventor. South Africa's patent office became a notable exception by granting the patent, albeit without examination. Beyond these legal rejections, many computer scientists remain skeptical that current AI systems are truly capable of inventing autonomously.

What counts as a human contribution?

While the concept of a fully autonomous AI invention remains contentious, policymakers still must engage with the reality that AI is playing an increasingly large role in the inventive process. This raises critical questions, such as:

- When AI is used in the inventive process, what should be the threshold of human contribution for a patent to be granted?
- What kind of human contributions count as part of the inventive process, as opposed to the mere use of a tool?
- Should the AI model's developers and data providers have entitlement to patents arising from inventions made using their systems?

Each question carries widespread implications for innovation incentives, patent validity and the fundamental purpose of the patent system.

Disclosure and transparency concerns

Patent law requires that applications name an inventor and disclose the invention in enough detail for others to reproduce it. However, they do not require disclosure of *how* the invention was made. This means that the role of an AI system in the inventive process can potentially remain hidden both from patent examiners and the public.

If jurisdictions decide that the extent of AI involvement is material to patentability, and if human contribution thresholds matter for granting patents, then the current disclosure framework is insufficient. Patent examiners would need mechanisms to ascertain the respective contributions of humans and AI systems when evaluating applications. Without such transparency, whatever standards that emerge for AI-assisted invention cannot be applied consistently, creating uncertainty for both patent applicants and examiners.

If there are no changes to the current IP laws, and the patent law thresholds for human contribution are high, will the lack of patent protection for AI-generated inventions lead to an increased reliance on trade secrets protection? This, in turn, could drive major transparency concerns and compound the "black box" problem of AI. Would investment in AI be disincentivized? Would there be grounds for patent revocation if a human were wrongly named as an inventor of an AI-generated invention?

Potential policy solutions

Potential policy solutions include:

- Exploring joint inventorship between AIs and humans
- Removing the requirement to name an inventor at all
- Naming the person with the closest connection to the AI as the inventor
- Naming a human inventor but requiring disclosure of the involvement of AI
- A brand new IP framework for AI-generated inventions

Whatever the outcome, policymakers will need to go back to basics, with analyses of the economic and social purposes of the patent system, i.e., to incentivize human creativity and the disclosure of new inventions by granting inventors the exclusive right to exploit their invention for a set period.

What are current major legislative and other developments around IP and AI?

China

China has taken a pragmatic and increasingly permissive approach to AI and IP, and has been implementing comprehensive AI governance through sector-specific regulations.

In November 2023, in a landmark case, the Beijing Internet Court ruled in an infringement lawsuit that an AI-generated image is copyrightable and that a person who prompted the AI-generated image is entitled to the right of authorship under Chinese copyright law. A similar decision was made by the Changshu People's Court in March 2025. The court considered that an image created through AI can be an original intellectual achievement as it involves making design choices, writing prompts, setting parameters and selecting the final output.

"Spring Breeze Brings Tenderness" was created in 2023. The disputed image in *Li* v. *Liu* was generated by Li Yunkai using the AI image generator Stable Diffusion.

On December 31, 2024, China's National Intellectual Property Administration issued "Guidelines for Patent Applications for AI-Related Inventions". These guidelines emphasize that the inventor(s) listed in a patent document must be a natural person(s). In September 2025, "Measures for the Labeling of Artificial Intelligence-Generated and Synthetic Content" came into effect, mandating

explicit and embedded labeling of all AI-generated text, images, audio, video and immersive virtual scenes distributed online within China.

European Union

The European Union's AI Act is a comprehensive regulatory framework that became law on August 1, 2024, with a phased implementation approach. It placed a ban on AI systems posing unacceptable risks, starting in February 2025, with codes of practice applying nine months after entry into force and rules for general-purpose AI models applying 12 months later.

The AI Act has two provisions that address copyright: Article 53(1)(c) and (d). The first requires "general purpose AI" developers to comply with copyright law and the conditions set out in the Directive on Copyright and Related Rights in the Digital Single Market to allow text and data mining. In essence, this means that copyrighted materials can be used to train AI models so long as right holders do not express their refusal. The second provision requires that AI developers report on the data sets used to train general-purpose AI models. This creates transparency obligations that could expose potential copyright infringement in training data sets.

In February 2026, the European Parliament's Committee on Legal Affairs adopted a report addressing the intersection of generative AI and copyright law, signaling its intent to take a stronger stand in favour of copyright holders.

Japan

Japan was the first country in the world to explicitly exempt text and data mining from copyright liability. Notably, unlike similar exemptions in other countries, this applies to both commercial and non-commercial uses. This was done to encourage the growth of Japan's digital economy. Since then, Japan has expanded exemptions for AI-based activities, with new laws exempting the following activities from copyright infringement:

1. certain uses of copyrighted works for machine learning;
2. the making of incidental electronic copies of works; and
3. the use of copyrighted works for data verification.

Japan has maintained this relatively permissive approach to AI training data, though it may be moving toward a more nuanced stance as litigation increases globally.

United Kingdom

The United Kingdom has protections for "computer-generated works," a somewhat unique form of copyright available in the United Kingdom, Ireland and a few other jurisdictions. Simply put, these rules give copyright protection to works created entirely by computers with no human involvement. The rule solves the "who owns it?" problem by naming the human who created the computer system that generated the work as the copyright beneficiary.

On the surface, these laws may seem like they solve the conundrum of copyright protection for AI-generated works. However, this approach remains fundamentally at odds with copyright laws worldwide, notably the Berne Convention. These are designed around the core assumption that creators are humans who benefit from protection. The United Kingdom's approach goes against the basic "author-centric" nature of copyright law, and so functions more as a problematic workaround rather than a proper solution to the AI authorship challenge

In December 2024, the UK Government published a consultation on "Copyright and Artificial Intelligence" that proposes, among other things, to reform protections for computer-generated works, which could have significant impacts for AI. Following these consultations, in March 2026, the Government published its *Report on Copyright and Artificial Intelligence*. This report made no concrete proposals, abandoned some its previous preferences (such as text and data mining exceptions and right holders opt-out), and overall adopted more of a wait-and-see approach to policymaking in this area.

The UK Data (Use and Access) Act 2025 does not impose immediate legal obligations on AI developers to secure consent from right holders, but strongly signals the potential emergence of a formal framework for AI and IP soon.

United States

In February 2025, in *Thomson Reuters* v.. *ROSS*, a US federal court issued the first major AI copyright decision of 2025, upholding the plaintiff's copyright infringement claim and rejecting the defendant's fair use defense. It is worth noting that the AI tool in this case was not a GenAI model writing new content, but rather a search engine style model that functioned as a directly competing product to the copyrighted material it was trained on. Since this case, other decisions in the US (notably *Kadrey* v. *Meta* and *Bartz* v. *Anthropic*) have upheld the fair use defense.

Executive Order 14179, issued in January 2025, reoriented US AI policy by revoking previous directives emphasizing data protection and transparency and aiming to eliminate federal policies perceived as impediments to innovation. A subsequent order in December 2025 was aimed at blocking states from enforcing their own AI regulations.

The US Copyright Office published a three-part report on *Copyright and Artificial Intelligence*. The third part, on GenAI training, was released as a pre-publication version in May 2025, and essentially concluded that legislation was premature at this time. Instead, it recommended that fair use doctrine continue to be applied on a case-by-case basis, with licensing as a fallback where fair use does not apply.

Cross-jurisdictional concerns

When an AI system potentially infringes on copyright and the case reaches a court, it raises some challenging practical issues rooted in private international law. The core question here is: does the court have the competence to hear the case?

This is a complex issue that involves looking into not only the intricate corporate structures of the companies creating AI products (often spanning several jurisdictions) but also where complex GenAI models are trained, refined and deployed (again, often in several jurisdictions).

The EU AI Act has tried to solve this in Recital 106. It states that:

"providers of general-purpose AI models should put in place a policy to comply with Union law on copyright and related rights."

It goes on to say that:

"any provider placing a general-purpose AI model on the Union market should comply with this obligation, regardless of the jurisdiction in which the copyright-relevant acts underpinning the training of those general-purpose AI models take place."

In plain language: *"If you want to sell your AI model in the European Union, you must follow EU copyright law, regardless of where the model was trained."* The underlying logic here is that no company should gain an unfair advantage in the European market by training their AI in countries with weaker copyright protections.

However, Recital 106, while sweeping, is not legally binding and ignores the fact that copyright law is governed by the principle of territoriality. In practice, if the text and data mining activities take place outside of the European Union and the model is placed in the European Union market only afterwards, then the provider will be compliant with the recital. On the contrary, should any of the training stages happen within European Union territory, such as web-scraping, then the recital kicks in and the provider will have to respect European Union law.

Despite these provisions being narrowly focused on copyright law and the European Union, it would not be surprising to see this approach exported in the future in other areas of IP law and beyond the territory of the European Union. The goal is to avoid a "race to the bottom" mentality, where AI companies shop for the most permissive jurisdictions to develop their models while still accessing stricter markets for sales.

Developing and protecting AI systems

Can I patent my AI system?

Most AI can be considered software. In many jurisdictions, software cannot be protected by patents. To understand the nuance here, we need to break down the key terms in this area.

Software is a **computer program** that instructs hardware (i.e., the computer's physical equipment) to perform specific tasks. A **computer program** is an **algorithm**, which is a series of instructions to perform a computation.

Algorithms are usually written in a programming language, like Java or C++ or Python, which is similar to natural language. This is called the **source code**. The source code is then usually translated to machine language (**object code**), which consists of binary values (0s and 1s) that provide operations a computer can execute directly.

The European Patent Office (EPO) does not consider a computer program to be patent eligible subject matter (and considers the term "software" to be ambiguous). Specifically, in its 2018 annual update to its "Guidelines for Examination," the EPO notes that AI and machine learning are based on algorithms that are of an "abstract" (and so, unpatentable) nature and, irrespective of training, cannot be patented. However, the EPO may allow patents on **computer-implemented inventions** (CII), which functionally allows patents on some software, including AI applications.

A CII "involves the use of a computer, computer network or other programmable apparatus, where one or more features are realized wholly or partly by means of a computer program." Eligible patent claims must have a technical character, which may involve claiming a computer-implemented method rather than an algorithm. Claims must also have a technical contribution to an inventive step.

For example, a method for having a computer generate random numbers by itself would not be patentable, but the EPO has found the same method patentable when used to simulate a circuit in a particular manner. The EPO guidelines provide other examples of patentable AI technical applications: use of a neural network in a heart-monitoring apparatus to identify irregular heartbeats, or the classification of data based on low-level features (e.g., edges or pixel attributes in images). AI may also be protectable where it is specifically adapted to run on a computer, where "the design of the mathematical method is motivated by technical considerations of the internal functioning of the computer."

As is the case for all patents, a CII needs to meet general patentability criteria: the invention must be novel, provide an inventive step, a "technical solution to a technical problem," and have industrial applicability.

The US also permits certain software patents. Before the AI era, the United States Supreme Court found specific software patents to be patent-ineligible "abstract ideas" implemented on a computer. Subsequently, lower-level courts have upheld a limited number of software patents, but there is no clear rule distinguishing the patentable from the unpatentable. The US Patent and Trademark Office guidance states that in order to refuse a CII as an abstract idea, patent examiners must factually prove that claims are "well-understood, routine, and conventional".

When successfully obtained, or sometimes even as pending applications, patents can offer significant value to businesses. That said, due to the relatively high cost of obtaining patents, protection is usually sought selectively in key jurisdictions. This means that patent holders can only prevent third parties from using their inventions in certain jurisdictions. Even for large enterprises, "worldwide" patent protection is usually both cost prohibitive and unnecessary. At the moment, considering the countries that are racing to become the global leader in AI, the most important jurisdictions are the US, France, Germany, the United Kingdom, Japan, Republic of Korea and China.

Can I use copyright to protect my AI model or system?

Copyright is a critical source of protection for software, which is protected as a "literary work" under Article 4 of the WIPO Copyright Treaty (WCT) and Article 10 of the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS). When it comes to machine learning and neural networks, copyright only

protects the code. The model will be protected by a combination of copyright and patents law, and possibly by trade secrets.

Exactly what qualifies for copyright protection varies by jurisdiction. In the UK, preparatory design work qualifies for protection if a computer program can result from it at a later stage.

In the EU, the protection of software is regulated under the Software Directive, which establishes that protection subsists whenever the work is the "author's own intellectual creation."

Copyright protects expression and does not extend to ideas, procedures, methods of operations or mathematical concepts. This means copyright does not prevent competitors from creating their own code that can behave like the protected code. If a competitor can access protected code, they may be able to rewrite the code so that it no longer infringes but has the same functionality. Even without direct access to the code, it may be possible for a competitor to generate a similar program from observing how software behaves. If they write their own code, and it does not include the protected source or object code, there will generally be no infringement. In fact, a competitor's code will then be protected by its own copyright.

Can I use trade secrets to protect my AI innovation?

Software may benefit from trade secret or confidential information protection. If a competitor cannot access the source code or otherwise determine how to recreate the software's functionality, trade secret protection may be effective. For AI, this could be particularly effective if third parties do not have direct access to the AI system and only the AI's output is shared outside an organization. For example, weights and parameters of an AI model could be protected as confidential information.

Trade secret protection requires constant vigilance since it relies on secrecy. To constitute a "trade secret" under IP (with the caveat that different jurisdictions have different rules):

- the information may need to have economic value by virtue of the fact it is not generally known; and
- the owner must take reasonable steps to keep the information secret.

This is one reason why non disclosure agreements (NDAs) are common and essential, particularly as part of contracts with employees and independent contractors. If a company fails to take reasonable steps to protect information (including technical measures) and discloses it without any confidentiality agreement in place, then that information generally does not obtain trade secret protection from the outset or loses that protection.

If a company takes reasonable steps to restrict disclosure of information that is lawfully in its control, and the information has value due to its confidentiality, then it may be possible to prevent that information from being disclosed, acquired or used without their consent in a manner contrary to honest commercial practices. In practice, this means that if the information is leaked through, for example, a hacker attack or a breach of contract, there may be legal remedies available to restrict use of the information or seek damages.

What is the role of open-source licensing in IP and AI?

Special considerations apply to the use of **open-source software** (OSS). OSS usually refers to source code that is publicly available, and that may be used by third parties without cost. However, while OSS is usually free of monetary cost, it generally comes with restrictions on how the code may be used.

Using OSS may require a license that prevents users from making modified versions of the source code or from using it for commercial purposes. OSS licenses might also restrict anyone incorporating the source code into other software unless they agree to make the entire program open source, effectively eliminating its proprietary status.

Thousands of different types of OSS licenses exist, many of which limit the IP protections available to any product incorporating OSS. This applies both to the use of code libraries and to snippets that may be incorporated into a broader codebase.

The discussion around open-source AI is still in its infancy. According to the open-source Initiative, an Open Source AI is a system made available under terms that grant users the freedoms to use, study, modify and share the system.

As the landscape currently stands, open-source AI exists on a spectrum from fully open to fully closed, depending on which components are publicly available under open-source licenses. Such components can include documentation, software code, training data, weighting factors and model architecture. The

EU AI Act defines "free and open-source AI" as AI components made accessible under open-source licenses, though this definition has limitations and doesn't specify how many components must be open. This is important, because the AI Act provides lighter requirements for stakeholders releasing “open-source AI” (though this excludes high-risk systems and monetized services). Some critics have argued this creates incentives for "open-source washing" to avoid transparency obligations.

To summarize, open-source AI requires different concepts than traditional open-source software due to its complexity and multicomponent nature. Clearer definitions and legal frameworks for what constitutes “open-source AI” are still emerging. In the future, reliance on open-source frameworks could provide a key source of transparency around AI, provided that the entire open-source system is considered and open standards are adopted horizontally. Users should be aware that different AI OSS licenses may come with their own conditions and restrictions.

Can I use copyrighted materials to train an AI model?

As outlined in Chapter 1, recent rapid advancements in machine learning have been made possible thanks to the widespread availability of large quantities of digital data, often referred to as “input material”. This can be in various forms: from statistical information such as numbers, where no copyright concerns would arise, to text, image, video and voice recordings, where copyright issues have become a key point of contention and debate.

In some cases, input materials are collected as training material for an AI system through text and data mining (TDM), which is the highly automated processing of large amounts of data to reach new knowledge. (Again, think of the example from Chapter 1: tens of millions of purchase records from a retail chain are unintelligible to humans due to the sheer scale of the data, but machine learning can be used to accurately predict sales trends.)

Generally, a TDM operation involves the following stages:

1. Extracting information from accessible online sources using bots or other automated systems. This is often called “crawling” or “scraping.”
2. Formatting the collected data.
3. Scanning the input materials to find correlations, trends and patterns.
4. Processing and analyzing data.
5. Storing data for validation purposes.

The key point of contention here is that each of these stages may involve a different kind of **reproduction**, which is one of the exclusive economic rights reserved for copyright holders in all jurisdictions, and the principal right at stake in these processes.

That said, many jurisdictions see significant societal and economic value in TDM activities, and so legislators are seeking to balance copyright holders' rights and users' interests by allowing exceptions for TDM. As noted in Chapter 2, Japan was the first country in the world to explicitly exempt TDM from copyright liability in both commercial and non-commercial cases.

In the US, there are a few decided cases (and several pending) on whether "fair use" is a defense against copyright infringement where copyright-protected material has been used for training AI systems.

When deciding if "fair use" is at play, the four factors of section 107 of the US Copyright Act have to be considered:

1. the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
2. the nature of the copyrighted work;
3. the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
4. the effect of the use upon the potential market for or value of the copyrighted work.

In two recent decisions, *Bartz* v. *Anthropic* and *Kadrey* v. *Meta*, the courts held that there was fair use in the use of the copyright-protected material in training AI models. The decisions were based largely on the first factor, "purpose and character of use", where the key issue is analyzing whether or not the use of the copyrighted material was "transformative." In both these cases, the courts decided that AI training is indeed transformative, and so counts as "fair use." In *Kadrey* v. *Meta*, the court also looked at the fourth factor: the impact on the market and the value of the copyrighted works. The conclusion was that because the AI system could not fully reproduce the copyrighted works or replace them in the market, the training was fair use.

However, it is worth noting that these decisions are specific to the facts of the cases and for now the US Copyright Office continues to recommend that fair use is applied case-by-case. Without further clarity, it is best to avoid using copyrighted content to train AI models, or obtain permission from the right holder(s) where necessary.

Other jurisdictions, such as Israel, have a similar, open-ended, fair use doctrine. Israel's Ministry of Justice has been vocal in the AI/copyright intersection, stating openly that TDM training falls within either the fair use exception or the temporary copying exception.

At the EU level, TDM exceptions have been introduced through two mandatory exceptions in Articles 3 and 4 of the Directive on Copyright and Related Rights in the Digital Single Market Directive (the CDSM Directive): one addressing scientific purposes, and a second for the purposes of TDM, including AI training. The EU legal framework states that for the second exception, rights holders are allowed to place contractual restrictions limiting TDM through their Terms and Conditions or Terms of Service. All member states had to implement both exceptions by June 7, 2021 and would be held liable for lack of compliance if they failed to introduce them into their national copyright law.

In conclusion, there is no simple answer to the general question of whether copyrighted materials can be used to train AI models. Representatives of creative industries have engaged in extensive lobbying to defend their position and ensure that authors, artists, musicians, photographers, and other rights holders are remunerated for the use of their copyrighted works. They maintain that when TDM is used on copyrighted works to train AI models this could amount to copyright infringement, and that rights holders should be approached for licenses and able opt out altogether, if they wish.

Creating and inventing with AI

Can I patent an invention created using AI?

There are several provisions of patent law that do not easily mesh with AI-generated inventions. These are:

1. the notion of an inventor;
2. the inventive step requirement (referred to as the “nonobviousness” test in certain jurisdictions); and,
3. disclosure requirements.

The notion of an inventor

Different jurisdictions have different criteria for inventorship. In the US, the “threshold question in determining inventorship is who conceived the invention. Unless a person contributes to the conception of the invention, he is not an inventor.” Conception is the “the formation in the mind of the inventor of a definite and permanent idea of the complete and operative invention as it is thereafter to be applied in practice". After conception, a person having ordinary skill in the subject matter of the invention should be able to use the invention without undue experimentation.

In the UK, an inventor is the “actual deviser” of an invention. There is a two-step approach to determining inventorship: first, the inventive concept of an application must be identified, after which it must be determined who devised that concept. The “inventive concept” is “concerned with the identification of the core (or kernel, or essence) of the invention – the idea or principle, of more or less general application which entitles the inventor’s achievement to be called inventive”. The inventive concept varies by application, but may reside in an idea, a means of realizing an idea or the combination of individually known elements. Ultimately, the inventive concept must be determined on a case-by-case basis, but superficial contributions, including to patent claims, do not qualify for inventorship.

The inventive step requirement

Generally, for patents to be granted, an invention must be new (i.e., the **novelty requirement**), involve an **inventive step** and have **industrial application**. It is the **inventive step** requirement that makes it difficult for AI-generated inventions to meet the traditional threshold.

An invention can be said to include an inventive step if the invention is not **obvious** to a person skilled in the relevant field, as measured against the current state of the field or the "art". This is different from the novelty requirement in patent law, which is a relatively objective assessment that asks the question: "did someone else get here first?"

By contrast, for the inventive step requirement, the core comparison is done with the prior art. The core question is: "does the prior art get us close enough to the invention that it is an obvious invention to come up with?" To put it another way: "would this invention seem obvious to a person skilled in this field?"

This test exists so that only those inventions that represent significant advances qualify for patent protection. That is because patents can have significant social costs, and too low a bar would allow for excessive patenting, which would ultimately be counterproductive.

For AI-generated inventions, the difficulty of the inventive step requirement lies in the definition of a person "skilled in the art". According to the EPO guidelines, a person skilled in the art is "presumed to be a skilled practitioner in the relevant field of technology, who is possessed of average knowledge and ability and is aware of what was common general knowledge in the art at the relevant date".

The challenge is that an AI system could have access to far more than "average" or "common general" knowledge in the "relevant field of technology". This means the scope of "obviousness" for AI is much larger than it is for a natural person. This challenges definitions of both "skilled in the art" and what counts as "obvious".

The use of AI in research and development (R&D) is gradually becoming the norm, though it is not yet an essential and integrated element of all fields. In any case, once "inventive" AI is a standard part of R&D in a particular field, it may be challenging for a patent examiner or court to determine what an AI would find "obvious." Going forward, will our definition of a person skilled in the art need to account for their access to AI systems? This is a tricky question, and it may force

patent offices to refocus on different factors, such as: objective; economic facts about how an invention was received in the marketplace; whether others tried and failed to generate the same invention; and whether there was a long-felt need for the invention.

Disclosure requirements

One of the central requirements of patent law is that the patent filing must disclose the invention in sufficient detail so that a person skilled in the art can comprehend it. Essentially, the description of the invention should be clear and complete enough for the person skilled in the art to reproduce it without too much difficulty (according to the EPO Boards of Appeal, a "reasonable amount of trial and error is permissible").

For AI-based inventions, this raises important practical question: what type of information and what level of detail need to be provided in the patent application about the AI system itself? To what extent should the application explain the AI model's architecture, training data and methods? How can applications satisfy the disclosure requirement while navigating the inherent "black box" nature of many AI systems, where the developers themselves may not fully understand how the model arrived at specific outputs?

"Disclosure" in the AI context carries a second, equally important meaning related to transparency about the inventive process itself. Current patent law in most jurisdictions does not require applicants to explain *how* an invention was made: only what the invention is and how to reproduce it. This means the use of AI in the inventive process can remain hidden.

This creates a significant policy challenge. If jurisdictions decide that a certain threshold of human contribution is required for patentability (as suggested by the rejection of the DABUS applications in most countries), then patent examiners would need to know not just *if* but *how* AI was used in the inventive process. Introducing disclosure requirements to support this would represent a fundamental shift in patent practice, requiring applicants to disclose their methods of inventing. Such a change could have far-reaching ripple effects on how the patent system operates and how innovation is documented and protected.

Can I copyright an AI-generated work?

When AI started appearing prominently in the IP discourse, it was primarily in the context of AI-generated outputs. The central question was whether there is an author of the outputs of these complex systems and, if so, could they trigger a legitimate copyright claim? Academic literature had already been grappling with this issue before the AI explosion of the last decade. The landscape was already complex when machine learning systems began to challenge the allocation of authorship. It has become even more complex considering the widespread availability of GenAI tools.

The two crucial IP concepts in this discussion are “authorship” and “originality".

But before discussing these concepts, we need to address an important underlying point. When trying to determine who should be considered the "author" of AI-generated works and whether these works deserve copyright protection, we must examine the rationale for why copyright law exists in the first place.

There are two main theories that justify copyright protection. The first is called "natural rights justification" and includes ideas like personality rights (works express the creator's personality) and Lockean labor theory (people deserve to own what they create through their work). These theories require a human creator, who is largely missing when an AI system generates content.

The second theory is the "utilitarian justification", which argues that copyright's main purpose is to benefit society by encouraging people to create new works and share intellectual content with the public. The idea is that copyright protections motivate creators to keep creating by granting them certain rights, such as the right to exclusively profit from their work.

So, if we cannot identify a specific person or entity who would be motivated by these copyright incentives, then it can be argued that there is no good reason to randomly assign copyright protection for the sake of it, without any solid theoretical basis.

Authorship

When determining whether AI-generated content can be protected by copyright, the first key question is: who or what counts as an "author?"

The Berne Convention for the Protection of Literary and Artistic Works (the main international instrument governing copyright) does not provide a clear definition of an author. Similarly, European law does not offer a definition of an author, except in two specific technical areas: software and databases.

However, despite the lack of an explicit definition, much of copyright law is human-centric and assumes human authorship. Because copyright law assumes human authorship, an AI-generated work can only receive copyright protection if we can identify a human being who can be called the "author". That person would then own both the economic rights (such as the right to sell copies of the work) and moral rights (such as the right to be credited) associated with the work, where that right is available.

Originality

The second key requirement for copyright protection is "originality". Courts have different understandings of the criterion of originality, depending on the jurisdiction and the type of work.

In the US, the landmark Supreme Court case *Feist Publications, Inc.* v. *Rural Telephone Service Co.* established that originality requires two things: independent creation (the work was not copied) and a "modicum of creativity" (at least a tiny spark of creative expression).

In the EU, the situation was chaotic for years because each country had its own definition of originality. While economic rights were harmonized as part of the InfoSec Directive in 2001, the basic question of what works deserve protection varied from country to country. This changed in 2009 with the *Infopaq* case, where the European Court of Justice created a unified standard: a work is only protected if it represents "the author's own intellectual creation". The Court explained this means the work must reflect the author's "free and creative choices".

Regardless of the legal test used, it is broadly understood that to meet the originality standard, a work must be the product of human intellectual creation involving genuine creative choices.

This creates a major challenge for AI-generated content. When we break down how machine learning and neural networks work (see Chapter 1), we see that a trained AI model is essentially a huge collection of mathematical functions linked together by numeric weights. In such a highly technical, algorithmic

process, it is very difficult to identify the "free human creative choices" and "modicum of creativity" that copyright law requires.

However, as the MPI study points out, these AI systems do not operate in a vacuum. They all require some degree of human intervention to get started and guide the process, even if the actual creation is driven by algorithms.

Several high-profile AI-generated works have tested the boundaries of our traditional concepts of originality and authorship, forcing us to reconsider these fundamental principles. Two specific instances of works that were rejected for registration by the US Copyright Office are worth mentioning: *Zarya of the Dawn* and *Théâtre D'opéra Spatial*. Both works were generated with Midjourney, a text-to-image GenAI tool.

Zarya of the Dawn is an 18-page comic book. The work was registered at the United States Copyright Office, but this was later revoked as the applicant had not disclosed the use of AI in the application. The applicant, Ms. Kristina Kashtanova, argued that she had provided “hundreds of thousands of descriptive prompts” to Midjourney until the “hundreds of iterations [created] as perfect a rendition of [her] vision as possible". She sustained that she was the author of that work. The US Copyright Office was not convinced, concluding that Midjourney generates images in an unpredictable way and the person who provides the prompts does not “actually form” the generated images and as such is not the “mastermind” behind them. That said, the text and arrangement of the images in the comic are protected by copyright, as they are considered the creative work of Kashtanova and not AI.

The cover page of *Zarya of the Dawn* generated in 2022 by Kristina Kashtanova using the generative AI

model Midjourney.

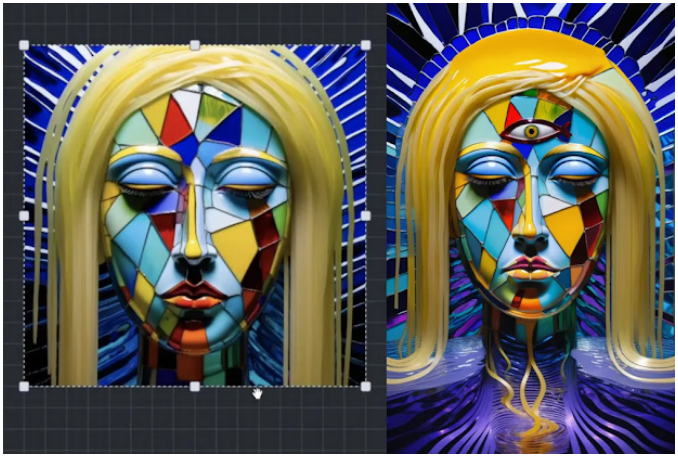
A similar application was registered by Jason M. Allen for the work *Théâtre D'opéra Spatial*. Allen sustained that he put 624 text prompts into Midjourney to arrive at the initial version of the work, which he then modified further with Adobe Photoshop and another AI tool called Gigapixel AI. The Office found that the work was not eligible for copyright protection as the human creative input was *de minimis*, with AI-generated elements dominating.

In 2023, the Beijing Internet Court came to the exact opposite conclusion in a landmark copyright infringement case, *Li* v.. *Liu*. Plaintiff Mr. Li sued Defendant Ms. Liu for copyright infringement after she allegedly used an image he generated using Stable Diffusion without permission. Mr. Li claimed this violated his right of authorship and dissemination. The court determined that Li's intellectual investment in creating the image through prompts and parameter adjustments met the requirements for a copyrighted work, and ruled that Ms. Liu had indeed infringed copyright. This sets a bold precedent for future cases involving AI-generated content in China.

Also in 2023, the US Copyright Office maintained its restrictive approach by granting only partial copyright protection to another AI-assisted artwork entitled "Rose Enigma", also created by Kristina Kashtanova. The Office registered copyright solely for the human-created sketch that served as the basis for the AI generation, while explicitly excluding all AI-generated or AI-modified portions. As with "Zarya of the Dawn", this case seems to indicate that while humans can retain copyright in their original contributions to AI-assisted works, the AI-generated elements do not qualify as copyrightable expression.

Kristina Kashtanova's original sketch for the work "Rose Engima", 2023 , alongside the final work generated by Midjourney.

In January 2025, a work entitled "A Single Piece of American Cheese" received copyright protection in the US. It was created by Mr. Kent Keirsey, CEO of Invoke AI. Using Invoke AI's tool, Keirsey iteratively modified an initial AI-generated image 35 times, making specific creative decisions about colour, composition and placement of elements. After an initial rejection, Invoke AI successfully appealed by providing documentation demonstrating Keirsey's active creative role, including a time-lapse video. The US Copyright Office determined that the final work contained "sufficient human original authorship," comparing it to a copyrightable collage where an artist compiles elements created by others into an original composition. The original AI-generated image was excluded from protection.



"A Single Piece of American Cheese" (2025) by Kent Keirsey (right) alongside the AI-generated image that provided the starting point for the work (left).

What should I look for regarding IP in AI terms of service?

As AI companies face lawsuits and new regulations, they are using their terms of service (ToS) agreements to shift legal risks away from themselves and onto their users. This was demonstrated by a 2024 study looking at the terms of service of 13 major AI companies and covering text, image, audio and video generators. Three aspects of copyright law were studied in-depth:

- ownership of inputs and outputs;
- liability if outputs infringe copyright; and,
- content moderation with the goal of reducing copyright infringement.

Ownership of input and outputs

While the study found that there is barely any reference in the ToS to ownership of user inputs into GenAI tools, more recent anecdotal evidence shows that AI companies are quite clear about the fact that they use inputs provided by users to train their AI models.

The study also identified a trend where providers assign ownership of outputs to the user, but also grant themselves a broad license to reuse the outputs. Some models (Stable Diffusion in particular) operate on an open-source license in this regard. If you are creating using AI, this means you technically "own" your AI-generated output, but the company can also use it however it wants, essentially nullifying any exclusive rights you might hold.

Liability for copyright infringement

If an AI-generated output potentially infringes on copyright, the AI provider's ToS typically place liability for this squarely on the shoulders of the user. The providers position their products as neutral intermediaries rather than active participants in content creation.

Content moderation and prompt filtering

Almost all companies use content moderation systems, including:

- filtering prompts to prevent users from requesting copyrighted content; and
- "notice and takedown" systems (similar to YouTube's copyright strikes).

This is consistent with broader trends in internet law: the more control a platform exercises over content, the more legal responsibility it is considered to have. The Digital Services Act in the EU has been pioneer legislation in this area, imposing serious transparency and reporting obligations and fines for noncompliance. By actively monitoring and removing content, companies are trying to demonstrate responsibility in this regard while still maintaining some protection.

In conclusion, users should be aware that most ToS for widely-used AI products are one-sided contracts that:

- put the bulk of the risk on the user;
- give the company maximum control over both inputs and outputs; and
- frame companies as neutral platforms rather than service providers.

The "output ownership" question will be particularly important to watch in the future, especially as AI systems become more capable and help users create more sophisticated and commercially valuable content.

What disclosure obligations do I have when using an AI-assisted creation?

In commercial contexts, disclosure obligations vary significantly by industry and jurisdiction, but transparency is increasingly becoming both legally prudent and ethically expected. Many professional bodies and sectors (such as academia, law and creative agencies) now require explicit disclosure of AI assistance in any work produced. Even where not legally required, failing to disclose AI assistance in commercial work can create liability risks around misrepresentation, breach of client expectations or violation of professional standards.

Additionally, as mentioned above, if an AI-assisted creation is found to violate copyright, the ToS typically place liability entirely on the user, which could have severe consequences in a commercial context.

Some major AI providers, such as OpenAI and Microsoft, have made commitments to cover some of their users' legal fees arising from copyright issues with AI-generated outputs. These commitments only apply to commercial and enterprise-level users and are hedged with significant conditions. In any case, they have yet to be tested in practice, as litigation in this space has so far been directed at the AI providers themselves, and direct claims against end users have not materialized.

Beyond formal requirements, commercial disclosure serves important business purposes: it manages client expectations, demonstrates technological competency, and can even be a competitive advantage when positioned as innovation and efficiency. However, the level of detail required varies: you may need to disclose that AI was used and for what purpose, but not the exact prompts or technical details (though it may be prudent to document these in any case). When in doubt, err on the side of transparency, as undisclosed AI use discovered later can damage relationships and professional credibility more than upfront disclosure ever could.

How can I demonstrate originality in an AI-assisted work?

Two landmark decisions – *Li* v. *Liu* in China in 2023 and "A Single Piece of American Cheese" in the US in 2025 – provide some guidance for creators who use AI heavily in their creative process. These cases indicate that demonstrating originality in commercial AI-assisted work requires showing substantial human creative input and decision-making throughout the process. For creators, this could involve documenting in detail:

- your conceptualization of the original idea;
- strategic prompting decisions;
- selection criteria for AI outputs and editorial choices; and,
- how you combined, refined or transformed AI-generated elements.

The key is proving that the final work represents human intellectual judgment and creative vision, with AI serving as a sophisticated tool rather than the primary creator.

The focus should be on areas where human creativity is most evident: strategic planning, conceptual framework, editorial curation and the integration of AI elements into a cohesive whole. To support this, it would be useful to maintain work files that show iterations, rejected alternatives, reasoning for creative decisions, or even video documentation of the creative act, as was provided for in "A Single Piece of American Cheese".

How can I protect my business's IP when using AI tools?

Given the uncertain legal landscape, the burden of protection remains squarely with users for the time being.

If you are using AI tools in a commercial context, regardless of jurisdiction, protecting your business IP requires a multilayered approach starting with careful vendor selection and contract negotiation. Choose AI services that offer robust data protection guarantees, including explicit commitments not to use your data for model training, clear data retention and deletion policies, and compliance with relevant data protection regulations. Negotiate specific contractual protections around data handling, including data residency requirements, encryption standards and the right to audit data handling practices. Many leading AI providers now offer professional and enterprise-grade tiers with enhanced privacy protections designed for commercial use.

Internally, it is wise to implement robust data governance policies that classify information by sensitivity level and restrict what can be input into AI systems. Never input highly sensitive information like trade secrets or confidential business strategies into AI tools, unless you have explicit contractual protections. Consider using on-premises or private cloud AI solutions for the most sensitive work, and establish clear employee guidelines about what data can and cannot be shared with AI systems. Regular security audits and monitoring of AI tool usage can help ensure compliance with your data protection policies and identify potential risks before they become problems.

How can I prevent my business's IP from being used to train an AI model?

Preventing your business IP from being used for AI training requires both technical and legal measures.

Some jurisdictions maintain opt-out provisions for copyright holders who do not wish to have their works used in TDM activities. So, from a legal perspective, clear copyright notices and licensing terms that explicitly prohibit use for AI training may provide some protection. Some companies are exploring digital watermarking and other technical solutions to track their content and prove unauthorized use. Technical barriers, including robots.txt files, can explicitly disallow AI crawlers, as can access controls on public-facing content and clear terms of use that prohibit scraping or automated data collection.

Be prepared to issue takedown notices when you discover unauthorized use, and consider working with legal counsel to send cease-and-desist letters to AI companies that may be using your content. While the legal landscape is still very much in flux globally, taking proactive steps to signal your intent not to participate in AI training strengthens your position in any future disputes and may deter casual scraping of your content.

The future of IP and AI

Outstanding questions and policy challenges

The intersection of AI and IP presents both unprecedented opportunities and complex challenges that will shape the innovation landscape for decades to come. Several fundamental questions remain unresolved, with major implications for how we balance innovation incentives with traditional IP protections.

After a decade in which judicial instances and jurisdictions have addressed the question of AI inventorship, the current direction seems to be as follows: while pure AI inventions are not patentable, AI inventions that involve human contribution could be patented. The key question requiring clarification is the extent of human contribution that would suffice in this context and the evidence that should be provided.

Similarly, the question of originality in AI-generated creative works challenges copyright law's human-centric foundations. The Beijing Internet Court's recognition that AI-assisted works can demonstrate sufficient human creativity through strategic prompting and editorial choices suggests a path forward, but the threshold for human contribution remains undefined. As AI tools become more sophisticated, determining where tool use ends and creative authorship begins will require nuanced legal frameworks that can adapt to rapidly evolving technology.

For policymakers, these challenges remain careful balancing acts. Overly restrictive approaches risk stifling innovation and putting jurisdictions at competitive disadvantages. Conversely, overly permissive frameworks may undermine creator rights and reduce incentives for human creativity. The EU's risk-based approach through the AI Act demonstrates one model for managing this balance, while Japan's permissive stance on training data reflects a different strategic choice prioritizing AI development.

Innovation through collaborative frameworks

Looking ahead, the intersection of AI and IP offers remarkable potential for driving innovation across industries. Rather than the zero-sum competition often portrayed in current debates, thoughtful frameworks can lay the groundwork for a future where AI and IP protections work together to enhance human creativity and accelerate technological progress.

The concept of "AI-assisted innovation" offers a framework for understanding how human creativity can be amplified by artificial intelligence. In patent contexts, this might involve AI systems helping inventors explore vast design spaces or identify novel combinations of existing technologies, while maintaining human judgment in defining problems, setting objectives and evaluating solutions. Such collaboration can accelerate innovation cycles and enable breakthrough discoveries that neither humans nor AI could achieve independently.

In creative industries, AI tools are already demonstrating their potential to democratize content creation, enabling small businesses and individual creators to produce high-quality work previously requiring substantial resources. When combined with clear attribution and fair compensation mechanisms, these tools can expand creative opportunities rather than displacing human creators.

This future relies on several key developments. First, adaptive legal frameworks that can evolve with technological advancement while maintaining core principles of fairness and innovation incentives.

Second, industry self-regulation and ethical standards should go beyond minimum legal compliance. The emergence of responsible AI principles and voluntary licensing frameworks suggests that many stakeholders recognize the long-term benefits of fair and transparent practices. Major media companies like the *Financial Times*, Associated Press and *Le Monde* have pioneered content licensing agreements with AI developers, creating precedents for fair compensation while enabling continued AI innovation.

Third, technological solutions should enable fine-grained rights management and compensation without introducing prohibitive practical challenges for AI developers who need access to high-quality data. Early innovations are emerging, such as automated licensing and royalty distribution and “ethical AI marketplaces.” These could dramatically reduce AI development and transaction costs while ensuring creator attribution and compensation.

Finally, education and capacity-building to help all stakeholders navigate this evolving landscape. As AI tools become ubiquitous, understanding their IP implications becomes essential for businesses, creators and policymakers alike. This guide is one contribution toward this end.

The challenges are significant, but so are the opportunities. By embracing collaborative approaches that respect both innovation and creator rights, AI's transformative potential can be harnessed while strengthening the intellectual property systems that have long served as engines of human progress.
The future of IP and AI lies not in choosing between human and artificial intelligence, but in creating frameworks that enable them to work together in the service of continued innovation and human flourishing.

WIPO resources

WIPO (2024). Generative AI: Navigating Intellectual Property.

WIPO (2024). Getting the Innovation Ecosystem Ready for AI: An IP policy toolkit.

WIPO (2020). Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence.

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

wipo.int