

L'exploitation d'œuvres protégées pour l'entraînement des systèmes d'intelligence artificielle

L'usage des outils d'intelligence artificielle (IA) connaît une croissance rapide et s'étend aujourd'hui à tous les domaines, qu'ils soient professionnels, académiques ou personnels. De la génération automatisée de contenu à l'analyse de données, en passant par l'assistance à la prise de décision, l'IA s'impose comme un vecteur incontournable d'innovation et de transformation sociale. L'Institut national de la statistique et des études économiques (INSEE) confirme dans une étude récente qu'en 2024, 10% des entreprises françaises comptant dix salariés ou plus utilisent une technologie d'IA contre seulement 3% en 2023. Cette différence de 7% en moins d'un an reflète l'omniprésence de l'IA même au sein des entreprises et ce, peu importe leur taille.

Le développement exponentiel de l'IA pose des questions particulièrement sensibles. En effet, l'entraînement des systèmes d'IA s'appuie souvent sur l'exploitation de vastes bases de données composées de textes, images et autres créations parfois protégées par le droit d'auteur. Au titre de ce dernier, l'œuvre protégée serait *“une œuvre d'esprit, laquelle appartient à la catégorie des propriétés intellectuelles (...) elle doit être le résultat d'une activité humaine, consciente, créatrice et qu'elle s'adresse à la fois aux sens et à l'intelligence des destinataires.”*¹. Le terme œuvre de l'esprit est employé dans le Code de la propriété intellectuelle qui exige, dans son article L.112-4 un caractère original afin qu'une œuvre soit protégée. L'analyse des enjeux juridiques et éthiques que soulève l'utilisation de ces œuvres dans l'entraînement des outils d'IA requiert une clarification préalable de certains concepts fondamentaux. Par exemple, une compréhension partagée des notions d'outils d'IA, d'IA générative et de grands modèles de langage (*Large Language Models*) fournit le cadre nécessaire pour examiner la manière dont les œuvres protégées sont mobilisées dans le processus d'entraînement de ces systèmes.

1/ Système d'intelligence artificielle : Les systèmes d'intelligence artificielle sont des logiciels ou services numériques qui utilisent des techniques telles que l'apprentissage automatique, l'apprentissage profond ou encore le traitement automatique du langage naturel,

¹ C. BERNAULT, A. LEBOIS, *JurisClasseur Civil Annexes*, OBJET DU DROIT D'AUTEUR. – Œuvres protégées. Règles générales (CPI, art. L. 112-1 et L. 112-2), Fasc. 1135, 1er juin 2025.

pour automatiser des tâches normalement associées au raisonnement humain. Ils reposent notamment sur le *machine learning*, une méthode qui permet à un système d'apprendre à partir de données sans être explicitement programmé.

2/ IA générative² : l'IA générative désigne une catégorie de systèmes d'IA capables de produire de nouveaux contenus. Ces contenus peuvent prendre différentes formes : texte, code informatique, images, sons, musique ou encore vidéos. Contrairement aux systèmes d'IA conçus uniquement pour analyser ou classer des données, l'IA générative apprend à partir d'exemples afin de créer des éléments originaux qui ressemblent à ceux présents dans ses données d'entraînement.

3/ Grands modèles de langage³ : Un *Large Language Model* (LLM) est un type de modèle de langage, c'est-à-dire un système informatique conçu pour analyser et produire du texte en imitant la manière dont les humains utilisent la langue. Les LLM reposent sur le *deep learning*, une branche de l'intelligence artificielle qui utilise des réseaux de neurones artificiels très profonds pour apprendre automatiquement à partir de grandes quantités de données textuelles. Grâce à cette approche, un LLM peut effectuer différentes tâches de Traitement Automatique du Langage Naturel (TALN), ou *Natural Language Processing* (NLP), telles que : comprendre ou interpréter du texte, traduire des phrases d'une langue à une autre, prédire la suite probable d'un énoncé, générer de nouveaux contenus textuels.

Dans ce contexte, la présence d'œuvres protégées au sein des données mobilisées par les systèmes d'IA constitue un élément qui ne peut passer inaperçu.

Le développement des nombreux outils d'IA a conduit à se poser de nombreuses questions quant au caractère éthique de l'IA. L'article examine d'abord les limites matérielles du RGPD au regard des enjeux contemporains du droit d'auteur (I). Il se poursuit par un examen du rôle de la directive DAMUN, d'une part dans l'encadrement des pratiques de text and data mining (II), d'autre part dans la mise en place d'un double régime d'exceptions au droit d'auteur (III). La réflexion met ensuite en évidence le décalage entre le cadre juridique européen et la réalité technologique de l'IA (IV), avant d'aborder la question du maintien d'une logique d'*opt-out* dans le nouveau règlement sur l'IA (V). Enfin, l'article se conclut par une

² L'intelligence artificielle, définitions et historique », Octobre 2024, Dossier de veille – Lab de la Sécurité sociale accessible sur <https://lab.securite-sociale.fr/files/live/sites/sitelab/files/documents/Veille/2024-10/IA%20d%C3%A9finitions%20et%20historique.p> consulté le 29 novembre 2025

³ *Ibid.*

comparaison des principaux systèmes juridiques, destinée à situer l'approche européenne dans un cadre plus large (VI).

I. Les limites matérielles du Règlement général de la protection des données face aux enjeux du droit d'auteur

La loi française « Informatique et Libertés » de 1978 n'étant pas suffisante, le Règlement général de la protection des données (RGPD)⁴ est venu la renforcer et la compléter par son entrée en vigueur le 25 mai 2018⁵. Le RGPD se concentre principalement sur l'encadrement de la circulation ainsi que le traitement des données à caractère personnel au sein de l'UE. En ce sens, le champ d'application matériel du RGPD se limite au traitement des données à caractère personnel concernant les personnes physiques. L'absence de prise en compte des droits d'auteur par le RGPD a conduit à la question de l'exploitation des œuvres protégées pour l'entraînement des systèmes d'intelligence artificielle. En effet, son considérant 63 prévoit uniquement que toute personne concernée a un droit d'accès à ses données personnelles afin d'en comprendre le traitement réalisé ; ce droit ne doit également pas porter atteinte aux droits d'autrui, dont le droit d'auteur.

II. L'intervention de la directive DAMUN en matière de *text and data mining*

Afin de pallier ce silence, la directive (UE) 2019/790 du Parlement européen et du Conseil du 17 avril 2019 sur le droit d'auteur et les droits voisins dans le marché du numérique⁶ (DAMUN) a été adoptée. Dans son considérant 3, la directive évoque *“une insécurité juridique [qui] subsiste, tant pour les titulaires de droits que pour les utilisateurs, en ce qui concerne certaines utilisations, notamment transfrontières, d'œuvres et autres objets protégés dans l'environnement numérique”*. Cette dernière a pour principal objectif d'instaurer un cadre juridique sécurisé en consacrant des exceptions à la fouille de textes et

⁴ Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données).

⁵ Site du Ministère de l'Économie, des Finances et de la Souveraineté industrielle et numérique de France, accessible sur <https://www.economie.gouv.fr/entreprises/gerer-son-entreprise-au-quotidien/assurer-sa-cybersecurite-et-la-protection-de-ses/le>, consulté le 29 novembre 2025.

⁶ Directive (UE) 2019/790 du Parlement européen et du Conseil du 17 avril 2019 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les directives 96/9/CE et 2001/29/CE (Texte présentant de l'intérêt pour l'EEE.).

de données ou le *text and data mining* (TDM). Il s'agit d'un procédé automatisé de recherche et d'extraction de volumes importants de données qu'il s'agisse de textes, de sons ou d'images en vue de leur analyse et de l'acquisition de nouvelles connaissances⁷. L'article 2, 2° de la directive DAMUN apporte une définition du TDM selon laquelle *“les nouvelles technologies permettent une analyse informatique automatisée d'informations sous forme numérique, telles que du texte, des sons, des images ou des données, ce que l'on appelle généralement la fouille de textes et de données.”*. Les systèmes d'IA minent des quantités de données afin d'en dégager des corrélations ; c'est ce qu'on appelle le *machine learning*.

La directive DAMUN affirme en son considérant 6 que *“Les exceptions et limitations prévues dans la présente directive tendent vers un juste équilibre entre les droits et les intérêts des auteurs et autres titulaires de droits, d'une part, et des utilisateurs, d'autre part. Elles ne peuvent s'appliquer que dans certains cas particuliers qui ne portent pas atteinte à l'exploitation normale de l'œuvre ou autre objet protégé ni ne causent un préjudice injustifié aux intérêts légitimes des titulaires de droits.”*. C'est cet objectif d'équilibre des droits, qui justifie, *in fine*, l'extraction massive via l'exception de fouille de textes prévue aux articles 3 et 4 de la directive.

III. La mise en oeuvre par la directive DAMUN d'un double régime d'exceptions au droit d'auteur

A. Un double régime d'exceptions

La directive DAMUN prévoit un double régime d'exceptions à ces mêmes articles 3 et 4 :

1. Une exception à des fins de recherche scientifique (article 3) : il est prévu une autorisation des organismes de recherches et les institutions de patrimoines culturels à avoir recours au TDM sur les différentes œuvres. L'exception est ici impérative et les titulaires de droits ne peuvent s'y opposer ; en effet, le régime de recherche scientifique bénéficie d'un statut d'ordre public.
2. Une exception générale (article 4) : le TDM est permis pour toute personne (y compris des développeurs d'IA par exemple) à condition d'avoir eu un accès licite au

⁷ Jean-François Debarnot JurisClasseur Communication, Fasc. 244 : Institut national de l'audiovisuel (INA), 25 avril 2024.

contenu. En revanche, cette exception n'est pas absolue puisque les titulaires de droits conservent une faculté d'opposition dite *opt-out*.

En ce sens, la consécration d'exceptions pour la fouille de données transforme un aspect du droit d'auteur en remplaçant le consentement préalable de l'auteur ou *opt-in* par une logique d'*opt-out*. En d'autres termes, l'*opt-in* est le mécanisme selon lequel l'utilisation d'une œuvre est interdite par défaut ; en ce sens, il faut un consentement préalable et explicite du titulaire des droits d'auteur avant toute utilisation. *A contrario*, l'*opt-out* vise le mécanisme par lequel l'utilisation de l'œuvre est autorisée par défaut ; le titulaire de droits doit ainsi faire une démarche afin de s'y opposer et rétablir une interdiction d'utilisation de l'œuvre protégée. Les chercheurs soulignent également plusieurs limites structurelles du mécanisme d'*opt-out* prévu par l'article 4 de la directive DAMUN.

B. Le mécanisme d'*opt-out* à l'épreuve de la technique

Ce mécanisme permet la fouille de données tant que le titulaire des droits ne s'y oppose pas "*de manière appropriée*"⁸, c'est-à-dire au moyen d'un signal technique que les systèmes peuvent reconnaître automatiquement. En pratique, cela renvoie souvent à l'utilisation d'outils simples comme le fichier robots.txt⁹, un petit fichier placé sur un site web pour indiquer ce qui peut être collecté ou non. Pourtant, l'interprétation de ce qui constitue un refus "*lisible par machine*"¹⁰ reste incertaine : une décision récente en Allemagne (*Robert Kneschke v. LAION e.V.*¹¹) a considéré qu'un refus rédigé en langage naturel dans des conditions d'utilisation pouvait également être valable. Cette ambiguïté est d'autant plus problématique que de nombreux titulaires de droits n'ont pas la maîtrise technique de la page web où leur œuvre est publiée. C'est le cas, par exemple, lorsqu'un livre, un article ou une photographie est mis en ligne par un éditeur, un journal ou une plateforme tierce. Ils peuvent donc matériellement se trouver dans l'impossibilité d'exprimer

⁸ Directive (UE) 2019/790 du Parlement européen et du Conseil du 17 avril 2019 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les directives 96/9/CE et 2001/29/CE, art. 4

⁹ T. Marcelin with F. Cassetti, AI and Copyright: The Training of General-Purpose AI (EPRS At a Glance European Parliamentary Research Service).

Google Developers, Présentation du fichier robots.txt accessible sur

<https://developers.google.com/search/docs/crawling-indexing/robots/intro?channel=31&hl=fr> (consulté le 29 novembre 2025).

¹⁰ T. Marcelin with F. Cassetti, AI and Copyright: The Training of General-Purpose AI (EPRS At a Glance European Parliamentary Research Service).

¹¹ WIPO, '2024 WIPO IP Judges Forum Informal Case Summary – Hamburg Regional Court, Germany: Robert Kneschke v LAION e.V., Case No 310 O 227/23 (DE048-j)' [2024].

un *opt-out* effectif. Ces limites nourrissent les interrogations quant à la portée réelle de la protection assurée par le cadre juridique actuel.

IV. Un cadre juridique en décalage : une lente adaptation à la réalité de l'IA

Finalement, il aura fallu attendre 2019 pour que le droit européen commence à se préoccuper de l'utilisation d'œuvres protégées par le droit d'auteur par les systèmes d'IA. En effet, de nombreuses critiques soulignent que cette absence de réflexion anticipée révèle un certain décalage entre l'évolution rapide des technologies d'apprentissage automatique et la lenteur de l'adaptation du cadre juridique. En effet, les grands modèles de système d'IA existent depuis 2010 et s'appuyaient déjà sur des données existantes probablement protégées pour certaines d'entre elles. En 2016, la Commission européenne avait déjà initié cette proposition de directive qui a suscité un bon nombre de débats au sein du Parlement européen retardant fortement l'adoption de cette directive DAMUN n'ayant vu le jour qu'en 2019. Ce retard peut s'expliquer par l'opposition féroce entre les géants du numérique et les industries culturelles durant les négociations. Les débats se sont focalisés sur la rémunération de la presse et la responsabilité des plateformes bloquant l'ensemble du texte et repoussant d'autant plus l'adoption des règles sur l'IA¹².

Néanmoins, L'UE avait eu l'occasion de se pencher sur les œuvres et objets protégés bien avant 2019 avec l'adoption de la directive du Parlement européen et du Conseil du 22 mai 2001 sur l'harmonisation de certains aspects du droit d'auteur et des droits voisins dans la société de l'information¹³. Cette dernière harmonise le droit d'auteur au début de l'ère numérique en renforçant le contrôle exclusif des auteurs sur la diffusion de leurs œuvres en ligne. Or, les systèmes d'IA n'étant pas encore créés ni popularisés, la directive n'évoque le droit d'auteur que sous un point de vue limité aux problématiques initiales liées au début de l'ère numérique.

¹² Toute l'Europe, Qu'est-ce que la directive européenne sur le droit d'auteur ? de Camille Richir, disponible sur <https://www.toutteleurope.eu/economie-et-social/qu-est-ce-que-la-directive-europeenne-sur-le-droit-d-auteur/>, (consulté le 29 novembre 2025).

¹³ Directive 2001/29/CE du Parlement européen et du Conseil du 22 mai 2001 sur l'harmonisation de certains aspects du droit d'auteur et des droits voisins dans la société de l'information.

V. Un maintien controversé de la logique d'opt-out par le nouveau Règlement sur l'intelligence artificielle

Alors que la directive de 2019 se concentrait davantage sur la réglementation de la fouille de données ou *mining*, l'adoption du Règlement sur l'intelligence artificielle (RIA) marque un réel changement de paradigme. Il devrait constituer une révolution normative dans le paysage numérique européen en ayant une approche bien plus moderne et adaptée avec de nouvelles règles. Entré en vigueur le 1er août 2024, il devrait être pleinement applicable en France le 2 août 2026 selon le calendrier d'application du RIA publié sur le site du Gouvernement¹⁴. Si le RIA apporte des nouveautés quant à ses règles de classification des systèmes d'IA à haut risque, il n'opère aucune modification substantielle quant aux œuvres protégées ; loin l'idée de faire un régime *ex nihilo*, le règlement opère un renvoi textuel explicite à la directive DAMUN. L'unique changement réside dans une obligation de transparence, prévue à l'article 19¹⁵ de la même directive et reprise dans son considérant 75, pour certains systèmes d'IA permettant d'une certaine façon à ce que les titulaires de droits puissent avoir un moyen de vérification de respect de leurs droits. Néanmoins, ce mécanisme d'opt-out survit même au sein du RIA illustrant un refus de l'UE de passer à l'opt-in malgré le nombre de critiques à l'égard du premier système. Nombreux pensent encore que l'opt-out est un danger pour la protection des œuvres protégées et donc, en ce sens, pour les titulaires des droits d'auteur ; ce n'est pas l'avis de la Commission européenne qui estime que l'opt-out est suffisant pour garantir leurs droits¹⁶.

VI. La comparaison du droit de l'Union européenne avec différents systèmes juridiques

Le paysage juridique international se complexifie encore davantage en raison des divergences entre les cadres juridiques¹⁷. Aux États-Unis, par exemple, la doctrine du *fair use*¹⁸ permet, dans certaines circonstances, l'utilisation non autorisée d'œuvres protégées afin

¹⁴ Service public - entreprendre, AI Act : quels changements pour les entreprises ? publié le 6 octobre 2025.

¹⁵ Article 19 de la directive DAMUN : «Les États membres veillent à ce que les auteurs et les artistes interprètes ou exécutants reçoivent, régulièrement et au minimum une fois par an, et en prenant en compte les spécificités de chaque secteur, des informations actualisées, pertinentes et complètes, sur l'exploitation de leurs œuvres et les exécutions de la part des parties auxquelles ils ont octroyé sous licence ou transféré leurs droits, ou des ayants droits de celles-ci, notamment en ce qui concerne les modes d'exploitation, l'ensemble des revenus générés et la rémunération due.».

¹⁶ Marie Sellier et Colette Bouckaertle, Quelles protections pour la propriété intellectuelle ? Les droits d'auteur et les droits de la presse face à l'IA générative le 17 octobre 2025.

¹⁷ Dentons, 'Can copyrighted material be used to train AI systems?', 11 June 2025.

¹⁸ US Copyright Office, 'U.S. Copyright Office Fair Use Index'.

de favoriser la liberté d'expression¹⁹. Cependant, il convient de noter qu'aux États-Unis, le *Copyright Office* a publié un rapport plus tôt cette année²⁰, concluant que l'utilisation de matériels protégés par le droit d'auteur pour entraîner des modèles d'IA pourrait constituer, *prima facie*, une violation du droit de reproduction de ces œuvres. Selon la décision *Bartz v. Anthropic PBC*, l'entraînement d'un modèle d'IA à partir d'exemplaires licitement acquis constitue un usage hautement transformateur relevant du *fair use*. À l'inverse, l'utilisation d'œuvres obtenues illicitement, notamment via des sites de piratage, demeure intrinsèquement contrefaisante, de sorte que le *fair use* ne peut s'appliquer lorsque l'accès initial aux œuvres est illégal²¹. De l'autre côté, l'Union européenne prévoit des exceptions limitatives et exprès, notamment en matière de *text and data mining*. Cette hétérogénéité crée une incertitude importante, tant pour les développeurs d'IA que pour les créateurs. Par exemple, au Royaume-Uni, l'affaire *Getty Images c. Stability AI*²² est une bonne illustration des enjeux liés à l'utilisation d'œuvres protégées dans l'entraînement des systèmes d'IA²³. Getty reprochait à Stability AI d'avoir utilisé ses contenus protégés pour former le modèle Stable Diffusion, alléguant une "*secondary copyright infringement*" c'est-à-dire une violation du droit d'auteur par la mise à disposition, la diffusion ou la facilitation de l'accès à une œuvre protégée, même sans avoir créé directement l'œuvre contrefaisante. La Cour a jugé que les paramètres internes du modèle, c'est-à-dire les données numériques reflétant les motifs et caractéristiques appris lors de l'entraînement, ne contiennent ni ne reproduisent les œuvres de Getty et ne constituent donc pas des « copies contrefaisantes » au regard du droit d'auteur britannique²⁴. La Cour a jugé qu'un « article » au sens du *Copyright, Designs and Patents Act 1988* (CDPA) peut être intangible²⁵, par exemple un fichier numérique ou les paramètres internes d'un modèle d'IA. Selon le CDPA²⁶, un article désigne tout objet dans lequel une œuvre protégée est incorporée et qui peut être importée, exportée, vendue ou autrement distribuée, incluant aussi les copies stockées sur tout support, physique ou

¹⁹ 17 USC § 107.

²⁰ US Copyright Office, *Copyright and Artificial Intelligence: Part 3 – Generative AI Training* (May 2025) - III. *PRIMA FACIE INFRINGEMENT* accessible sur - <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf> (consulté le 29 novembre 2025).

²¹ Ropes & Gray, Regina S. Penti, M. J Rizzolo and Y. Schaal, 'Anthropic's Landmark Copyright Settlement: Implications for AI Developers and Enterprise Users' (8 September 2025) accessible sur : <https://www.ropesgray.com/en/insights/alerts/2025/09/anthropics-landmark-copyright-settlement-implications-for-ai-developers-and-enterprise-users> (consulté le 5 décembre 2025).

²² *Getty Images (US) Inc and Others v Stability AI Ltd* [2025] EWHC 2863 (Ch).

²³ Latham & Watkins, 'Getty Images v Stability AI: English High Court Rejects Secondary Copyright Claim', 13 November 2025.

²⁴ *Ibid.*

²⁵ *Getty Images (US) Inc and Others v Stability AI Ltd* [2025] EWHC 2863 (Ch), paragraph 586.

²⁶ Copyright, Designs and Patents Act 1988, ss 22, 23, 27.

électronique. Traditionnellement, un article désignait des supports physiques incorporant l'œuvre (livres, CD, photographies), mais la Cour a élargi cette notion pour inclure des objets numériques, comme des modèles ou fichiers d'IA hébergés. Elle a fondé cette position sur le principe dit “*always speaking*”²⁷, selon lequel les lois doivent toujours être interprétées de manière à pouvoir s'appliquer aux nouvelles technologies et circonstances au fil du temps, et sur une lecture intégrée du CDPA, en soulignant que la Section 17 reconnaît explicitement le copiage “*dans tout support, par tout moyen électronique*”. Cette interprétation diverge du droit européen²⁸, où la notion d'articles pour la distribution ne concerne que des copies matérialisées susceptibles d'être mises en circulation sous forme d'objets tangibles, dans le sens de la directive 2001/29²⁹ et du traité de l'OMPI sur le droit d'auteur³⁰.

L'étude montre que, malgré les avancées introduites par la directive DAMUN et confirmées par le Règlement sur l'intelligence artificielle, le régime applicable à l'utilisation d'œuvres protégées pour l'entraînement de l'IA demeure marqué par un déséquilibre entre les besoins technologiques et la protection des titulaires de droits, et une complexité juridique accrue en raison des divergences entre différents cadres juridiques dans le monde. Par exemple, le maintien du mécanisme d'*opt-out*, associé à des exigences limitées de transparence, laisse subsister une incertitude quant à l'effectivité du contrôle des auteurs sur l'exploitation de leurs œuvres. Face à des technologies dont le développement dépasse le rythme de la réponse juridique, la question reste ouverte quant à la capacité du cadre actuel à assurer une protection réellement satisfaisante.

²⁷ *Getty Images (US) Inc and Others v Stability AI Ltd* [2025] EWHC 2863 (Ch), paragraph 580.

²⁸ Latham & Watkins, ‘Getty Images v Stability AI: English High Court Rejects Secondary Copyright Claim’, 13 November 2025.

²⁹ Directive 2001/29/CE du Parlement européen et du Conseil du 22 mai 2001 sur l'harmonisation de certains aspects du droit d'auteur et des droits voisins dans la société de l'information.

³⁰ World Intellectual Property Organization (WIPO)

TRT/WCT/001 https://www.wipo.int/wipolex/en/text/295166#P65_7507

Bibliographie :

Soures internationales :

- World Intellectual Property Organization (WIPO) TRT/WCT/001

Sources législatives européennes :

- Directive 2001/29/CE du Parlement européen et du Conseil du 22 mai 2001 sur l'harmonisation de certains aspects du droit d'auteur et des droits voisins dans la société de l'information.
- Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données).
- Directive (UE) 2019/790 du Parlement européen et du Conseil du 17 avril 2019 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les directives 96/9/CE et 2001/29/CE, art. 4.

Les jurisprudences :

- Getty Images (US) Inc and Others v Stability AI Ltd [2025] EWHC 2863 (Ch)
- Robert Kneschke v LAION e.V., Case No 310 O 227/23 (Regional Court of Hamburg, Germany).

Les articles et encyclopédies :

- Carine BERNAULT, Audrey LEBOIS, OBJET DU DROIT D'AUTEUR. – OEuvres protégées. Règles générales (CPI, art. L. 112-1 et L. 112-2), JurisClasseur Civil Annexes, Fasc. 1135 le 1er juin 2025.
- Dentons, 'Can copyrighted material be used to train AI systems?' (11 June 2025).
- European Parliamentary Research Service, Tristan Marcelin with Filippo Cassetti, 'AI and Copyright: The Training of General-Purpose AI' (EPRS At a Glance).
- Jean-François Debarnot JurisClasseur Communication, Fasc. 244 : Institut national de l'audiovisuel (INA), 25 avril 2024.
- Lab de la Sécurité sociale, 'L'intelligence artificielle, définitions et historique' (Dossier de veille, octobre 2024).
- Latham & Watkins, 'Getty Images v Stability AI: English High Court Rejects Secondary Copyright Claim' (13 November 2025).

- Marie Sellier et Colette Bouckaertle, *Quelles protections pour la propriété intellectuelle ?* Publié le 17 octobre 2025.
- Les droits d’auteur et les droits de la presse face à l’IA générative le 17 octobre 2025.
- Ropes & Gray, Regina S. Penti, M. J Rizzolo and Y. Schaal, ‘Anthropic’s Landmark Copyright Settlement: Implications for AI Developers and Enterprise Users’ (8 September 2025).

Les sites internet :

- CNIL, Entrée en vigueur du règlement européen sur l’IA : les premières questions-réponses de la CNIL, disponible sur <https://www.cnil.fr/fr/entree-en-vigueur-du-reglement-europeen-sur-lia-les-premieres-questions-reponses-de-la-cnil>, consulté le 30 novembre 2025.
- Google Developers, Présentation du fichier robots.txt accessible sur <https://developers.google.com/search/docs/crawling-indexing/robots/intro?channel=31&hl=fr>
- Toute l’Europe, article *Qu’est-ce que la directive européenne sur le droit d’auteur ?* de Camille Richir, <https://www.touteleurope.eu/economie-et-social/qu-est-ce-que-la-directive-europeenne-sur-le-droit-d-auteur/>, consulté le 30 novembre 2025.
- Service public - entreprendre, AI Act : quels changements pour les entreprises ? publié le 6 octobre 2025, disponible sur <https://entreprendre.service-public.gouv.fr/actualites/A18475>, consulté le 30 novembre 2025.
- US Copyright Office, ‘U.S. Copyright Office Fair Use Index’ accessible sur <https://www.copyright.gov/fair-use/>
- US Copyright Office, Copyright and Artificial Intelligence: Part 3 – Generative AI Training (May 2025) - III. PRIMA FACIE INFRINGEMENT accessible sur <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf>, consulté le 30 novembre 2025.