

Propriété intellectuelle

IA et droit d'auteur

La décision GEMA c/ OpenAI Ireland Ltd & Inc. : une décision structurante aisément transposable au droit français

Le Tribunal régional de Munich I (LG München I, 11 nov. 2025, n° 42 O 14139/24), saisi par la GEMA, pendant allemand de la Société des auteurs, compositeurs et éditeurs de musique (SACEM), s'est prononcé pour la première fois en Europe sur la qualification, au regard du droit d'auteur, à la fois de l'entraînement et de l'usage sans autorisation d'œuvres protégées par les modèles d'intelligence artificielle (IA) générative. Il qualifie la mémorisation des textes dans les paramètres d'un Large Language Models (LLM) d'acte de reproduction. En rejetant l'argument d'une simple corrélation statistique ou d'un hasard algorithmique, comme les exceptions de *text and data mining*, la juridiction affirme que les modèles de langage sont susceptibles de fonctionner, pour les contenus mémorisés, comme de véritables bases de données, rendant la qualification de reproduction juridiquement pertinente. Le tribunal juge en outre que la restitution des paroles dans les réponses générées par le *chatbot* constitue un acte de mise à disposition du public imputable à l'exploitant du LLM, lequel ne saurait s'exonérer de sa responsabilité en invoquant l'initiative des utilisateurs ou les mécanismes de décodage probabiliste. En ordonnant enfin la cessation des usages litigieux, il pose un principe clair : l'innovation algorithmique ne neutralise pas les droits des créateurs, et la responsabilité des acteurs de l'IA s'étend à l'ensemble de la chaîne, de l'*input* à l'*output*. Une décision à forte portée systémique, largement transposable au droit français.



Marie-Avril Roux Steinkühler

MARS-IP - Avocate
aux Barreaux de Paris et
Berlin - Spécialiste en droit
de la propriété intellectuelle
- Présidente de la French Tech
Berlin - Conseillère du com-
merce extérieur de la France

Les grands modèles d'IA générative, ou LLM, se nourrissent pour s'exercer et pour générer des résultats, d'immenses quantités de données. Ces données ou data sont en général des droits d'auteur : articles de presse, textes, images, vidéos, musiques, même les bases de données et tous autres contenus mis à disposition sur le web, que les systèmes d'IA « crawlent » ou explorent, puis « scrapent », c'est-à-dire extraient sans autorisation ni rémunération des ayants droit. Les industries créatives et culturelles alimentent donc, sans l'avoir choisi ni négocié, l'entraînement et les résultats générés par ces systèmes, et se révèlent être une pierre angulaire du développement de cette technologie, puisqu'elles en sont le fuel.

Faut-il pour autant renoncer, devant cette prodigieuse technologie, à rémunérer les ayants droit ? Le droit d'auteur est-il applicable, après tout, à ces usages, tant les résultats générés par les LLM diluent leurs sources ? Le droit d'auteur, frein au progrès ? Ce

phénomène n'a, en réalité, rien de nouveau : le droit d'auteur s'est historiquement affirmé comme un instrument de gouvernance des transformations technologiques. Le *Statute of Anne*, adopté en Angleterre en 1710, puis les législations révolutionnaires françaises à partir de 1791, ont précisément été conçus pour accompagner le développement de l'imprimerie et la diffusion du savoir, tout en protégeant les auteurs, qui perdaient la maîtrise des exploitations de leurs œuvres et de leurs revenus – et, ce faisant, la création elle-même.

Après l'incertitude générée par les décisions contradictoires rendues aux États-Unis, les unes faisant bénéficier aux exploitants de LLM de l'exception de *fair use*, les autres rejetant cette exception¹, après la

¹ M.-A. Roux Steinkühler, IA et droit d'auteur : l'argument du *fair use* rejeté par la justice américaine, *Légipresse* 2025. 413.

décision rendue en septembre 2024 par le Tribunal de Hambourg², qui a validé la constitution de *data sets* d'œuvres protégées sans autorisation des auteurs, motifs pris d'objectifs lointains de recherche, après enfin la décision rendue le 4 novembre 2025 par la High Court³ de Londres rejetant l'essentiel des prétentions de l'agence Getty en matière de violation du droit d'auteur par le modèle Stable Diffusion, le jugement prononcé par le tribunal de Munich⁴ apparaît comme un louable retour aux fondamentaux, tant il développe une analyse solide et claire, et dessine une ligne nettement en faveur de l'application du droit d'auteur, et des qualifications choisies.

Les faits étaient les suivants : la GEMA, organisme de gestion collective, gère, à l'image de la SACEM en France, les droits d'exploitation d'œuvres musicales qui lui ont été confiées par des compositeurs, des auteurs paroliers et des éditeurs de musique. En 2022, son assemblée générale avait décidé d'exercer son droit d'optout, qui figure depuis clairement sur son site internet⁵. Elle avait ainsi expressément réservé l'usage de son répertoire au titre de la fouille de textes et de données, conformément à l'article 4 de la directive (UE) 2019/790, transposé à l'article 44b UrhG, de la loi allemande sur le droit d'auteur⁶. Se fondant à dessein sur neuf œuvres emblématiques qui appartiennent au cœur du répertoire populaire allemand, tant par leur notoriété, celle de leurs interprètes que par leur diffusion intergénérationnelle, la GEMA avait constaté que le *chatbot* d'abord exploité par OpenAI Inc. puis par OpenAI Ireland Ltd, les deux défenderesses, était en mesure de restituer, à la suite de requêtes particulièrement simples telles que « Quel est le refrain de "Hommes" ? » (« Hommes » ou « Männer » est le titre d'une chanson), « Donne-moi aussi s'il-te-plaît la première strophe », « Qui est l'auteur du texte ? », « Donne-moi les accords de guitare de [titre de la chanson] », « Que vient-il après [premiers mots d'une strophe] ? », les paroles ou accords des chansons protégées, à l'identique ou légèrement modifiés ...

Le jugement prononcé par le tribunal de Munich apparaît comme un louable retour aux fondamentaux, tant il développe une analyse solide et claire, et dessine une ligne nettement en faveur de l'application du droit d'auteur, et des qualifications choisies.

Après un avertissement puis une mise en demeure restés infructueux, la GEMA a saisi le Tribunal de Munich.

Le tribunal déclare le droit allemand applicable, reconnaît la qualité d'œuvres protégeables aux paroles des chansons, qu'elles soient considérées dans leur intégralité ou par extraits (I.A), et juge que la mémorisation des paroles de chansons par le LLM (*input*), en particulier au cours de la phase 2, constitue un acte de reproduction (I.B), tout comme les réponses données par le *chatbot* (*output*) (I.C). Il considère la faute des défendeurs caractérisée, interdit la poursuite des usages illicites (I.E) et refuse toute application des exceptions, et en particulier la fouille de données (I.4D). Le jugement rejette

ainsi tous les arguments des défenderesses, qui contestaient toute violation, invoquant la nature probabiliste des modèles, l'absence de fixation matérielle des données et la non-applicabilité des règles classiques du droit d'auteur. Elles soutenaient également que leurs pratiques relèvent des exceptions légales de fouille de textes et de données (UrhG, § 44b) et qu'un

consentement implicite des auteurs découlerait de la mise en ligne des œuvres.

Le jugement fait ensuite très largement droit aux demandes de la GEMA : il ordonne l'interdiction de reproduire les paroles dans les modèles et de les rendre accessibles via le *chatbot*, impose la communication d'informations sur l'étendue des actes et les revenus, et reconnaît l'obligation d'indemniser les ayants droit, outre la publication du jugement devenu définitif.

Le jugement fourmille d'informations et d'analyses intéressantes, qui seront très utiles aux juristes français dans le cadre des procédures qui se profilent. Il rejette cependant les demandes formulées au titre du droit moral, pas encore unifié, et l'on peut penser que le droit français leur serait plus favorable. Il conforte la position récente prise par le Conseil supérieur de la propriété littéraire et artistique (CSPLA) sur le droit applicable (II).

En bref, il acte que l'IA générative s'inscrit pleinement dans le champ du droit d'auteur et que son développement ne peut se faire au prix d'une dilution des droits exclusifs des auteurs.

² M.-A. Roux Steinkühler, La fâcheuse et contestable affaire *Kneschke contre Laion e.V.*, *Légipresse* 2024. 688.

³ High Court of Justice Business and Property Courts of England and Wales, Intellectual Property List (ChD), 4 nov. 2025, n° IL-2023-000007, Mrs Justice Joanna Smith DBE.

⁴ À Munich siège le Deutsche Patent- und Markenamt (DPMA), pendant allemand de l'Institut national de la propriété intellectuelle (INPI), et tous les organismes de gestion collective ; c'est aussi un grand hub de médias.

⁵ Nutzungsvorbehalt für Text- und Data Mining, z. B. durch KI, GEMA News, 17 avr. 2024.

⁶ Pour une traduction en anglais de la loi allemande sur le droit d'auteur : gesetze-im-internet.de/englisch_urhg/index.html.

I - UN JUGEMENT PRÉCIS ET DÉCIDÉ, QUI QUALIFIE LARGEMENT L'INPUT ET L'OUTPUT DES MODÈLES DE LANGAGE DE REPRODUCTION DES ŒUVRES

A - La protection des paroles de chanson : un préalable indispensable et sans ambiguïté

Le tribunal commence par rappeler un point qui, en réalité, n'a pas été soulevé par les défenderesses : les paroles de chanson litigieuses constituent des œuvres protégées au sens de l'article 2, § 2, point 1, de la loi allemande sur le droit d'auteur et les droits voisins (*Urheberrechtsgesetz*). En se référant explicitement à la jurisprudence *Cofemel*⁷, qui a rappelé que, pour « être regardé comme original, il est à la fois nécessaire et suffisant que [l'œuvre] reflète la personnalité de son auteur, en manifestant les choix libres et créatifs de ce dernier », il rappelle que le seuil d'originalité exigé par le droit de l'Union est faible.

L'intérêt de la décision réside moins dans ce rappel de principe que dans l'attention portée par le jugement aux parties d'une œuvre – ce qui sera utile pour la suite de l'analyse. En s'appuyant sur la décision *Infopaq*⁸, le tribunal anticipe toute critique possible tirée de la reprise de quelques mots et affirme ainsi la protection d'extraits de refrains ou de strophes relativement courts : la protection ne dépend ni de la longueur du texte ni de son caractère complet, mais de l'expression d'un libre choix créatif.

Cette approche est particulièrement adaptée dans le contexte des systèmes génératifs, dont les sorties reposent précisément sur la reproduction de fragments et non nécessairement d'œuvres intégrales.

B - La mémorisation dans le modèle : une reproduction au sens du droit d'auteur

1 - Une lecture extensive de l'article 2 de la directive InfoSoc, invoquant le principe de neutralité technologique

Le cœur de la décision réside dans la qualification juridique de la mémorisation des paroles de chansons dans les modèles de langage. Le tribunal retient sans hésitation que cette mémorisation constitue une reproduction au sens de l'article 16 de l'UrhG⁹, qui transpose fidèlement l'article 2 de la directive 2001/29/CE du 22 mai 2001 sur l'harmonisation de certains aspects du droit d'auteur et des droits

voisins dans la société de l'information (InfoSoc). Il rappelle que la notion de reproduction doit être interprétée de manière autonome et uniforme en droit de l'Union, et qu'elle couvre toute fixation directe ou indirecte, temporaire ou permanente, quelle que soit la technologie utilisée pour la matérialisation (copie, numérisation, stockage dans un format de fichier, compression avec perte...).

En se référant notamment aux arrêts *Infopaq*, *Copydan*¹⁰, à la jurisprudence allemande et aux termes larges de la directive 2001/29/CE¹¹ repris par l'article 16, § 1^{er}, de la loi allemande, le tribunal rappelle que la reproduction ne suppose pas nécessairement une copie identique mais « par quelque moyen et sous quelque forme que ce soit ». Peu importe que l'œuvre ne soit pas stockée sous une forme perceptible directement par l'homme et que les paroles soient, en l'espèce, décomposées en paramètres : il suffit qu'elles soient intégrées dans un système de manière durable et qu'elles puissent être rendues perceptibles, fût-ce indirectement, au moyen de dispositifs techniques. Ici, il s'agit de séquences de *tokens* statistiquement probables qui reproduisent de manière reconnaissable les paroles des chansons.

Dans le même sens, peu importe la durée de la reproduction¹², qui peut être temporaire dès lors qu'elle satisfait à ce critère de rendre l'œuvre perceptible par les sens humains de manière directe ou indirecte.

2 - La démonstration de la mémorisation et le rejet de l'argument du hasard

Fort de ce préalable, le tribunal procède à une analyse détaillée du fonctionnement des modèles de langage, en distinguant les trois différentes phases du traitement des données :

- l'extraction et la préparation des corpus ;
- puis la deuxième phase d'analyse et d'entraînement ;
- et, troisièmement, la génération des sorties.

En se fondant sur les travaux scientifiques de référence de Carlini¹³ et sur leur analyse juridique faite

¹⁰ CJUE 5 mars 2015, aff. C-463/12, JAC 2015, n° 29, p. 38, chron. P. Mouron ; RTD com. 2016. 114, obs. F. Pollaud-Dulian) et *TU Darmstadt* (CJUE 11 sept. 2014, aff. C-117/13, *Légipresse* 2014. 520 et les obs. ; *ibid.* 611, comm. J.-M. Bruguière ; AJDA 2014. 1743 ; *ibid.* 2295, chron. E. Broussy, H. Cassagnabère et C. Gänser ; D. 2014. 1872 ; JAC 2014, n° 17, p. 6, obs. E. Scaramozzino ; RTD com. 2014. 810, obs. F. Pollaud-Dulian ; RTD eur. 2014. 965, obs. E. Treppoz.

¹¹ Consid. 21 et art. 2.

¹² En ce sens, BGH 23.10.2024, I ZR 112/23, pt 66, *Manhattan Bridge*, et implicitement dans CJUE 29 juill. 2019, aff. C-476/17, *Pelham*, pt 52, *Légipresse* 2019. 452 et les obs. ; *ibid.* 541, obs. V. Varet ; *ibid.* 2020. 69, étude C. Alleaume ; D. 2019. 1742, note G. Querzola ; Dalloz IP/IT 2019. 465, obs. N. Maximin ; *ibid.* 2020. 317, obs. A. Latil ; RTD com. 2020. 74, obs. F. Pollaud-Dulian ; RTD eur. 2019. 927, obs. E. Treppoz ; *ibid.* 2020. 324, obs. F. Benoît-Rohmer.

¹³ N. Carlini et al., *Extracting Training Data from Large Language Models*, 30th USENIX Security Symposium, USENIX Security, 2021, p. 2633-2650.

⁷ CJUE 12 sept. 2019, aff. C-683/17, pt 30, D. 2019. 1759 ; *ibid.* 2020. 1588, obs. J.-C. Galloux et P. Kamina ; RTD com. 2020. 54, obs. F. Pollaud-Dulian ; RTD eur. 2019. 930, obs. E. Treppoz.

⁸ CJUE 16 juill. 2009, aff. C-5/08, D. 2011. 2164, obs. P. Sirinelli ; RTD com. 2009. 715, obs. F. Pollaud-Dulian ; RTD eur. 2010. 939, chron. E. Treppoz.

⁹ « (1) Le droit de reproduction est le droit de faire des copies de l'œuvre, que ce soit temporairement ou définitivement, par quelque moyen que ce soit et en quelque quantité. »

par Cooper et Grimmelmann¹⁴, la juridiction identifie clairement la deuxième phase, celle d'entraînement, comme juridiquement déterminante. En effet, c'est à l'occasion de l'entraînement que les données d'apprentissage sont extraites et mémorisées, constituant une reproduction au sens du droit d'auteur.

Et, s'appuyant sur d'autres travaux de recherche¹⁵, le tribunal relève que la mémorisation des données d'entraînement peut être établie par comparaison directe des passages issus du système ou par l'analyse de paramètres de perplexité et d'entropie¹⁶, lesquels permettent d'évaluer le degré de certitude avec lequel un modèle reproduit une sortie, particulièrement élevé en présence de contenus mémorisés. Les prompts simples effectués ne constituent pas une condition de génération des données d'entraînement, mais un instrument probatoire destiné à démontrer l'existence d'une mémorisation.

De nombreux tests sont cités, s'appuyant sur des prompts très simples, qui tous concordent vers la reproduction exacte des paroles des chansons.

Compte tenu de la longueur et de la complexité des paroles reproduites, l'identité des résultats aux chansons ne peut être le fruit du hasard ou de lois statistiques. Selon la décision de Munich, peu importe que les reproductions soient le fruit de statistiques. OpenAI fonctionne *de facto* comme une base de données des contenus mémorisés, malgré les mécanismes de randomisation introduits au stade du décodage.

C - Les sorties du *chatbot* : reproduction et mise à disposition du public

Le tribunal retient également que la restitution des paroles dans les réponses générées par le *chatbot* constitue un acte autonome de reproduction et de mise à disposition du public au sens de l'article 19a de l'UrhG¹⁷. Il insiste sur le fait que les choix de conception relatifs au décodage et à la génération des réponses relèvent de la sphère de contrôle des exploitants du système, lesquels ne sauraient s'exo-

nérer de leur responsabilité en invoquant le caractère prétendument imprévisible des sorties, dès lors que ces mécanismes relèvent de choix de conception contrôlés par eux.

La juridiction souligne en outre que, lorsque les probabilités associées aux *tokens* mémorisés sont extrêmement élevées, les probabilités associées aux *tokens* correspondants atteignent souvent des valeurs proches de 100 %. La reproduction devient pratiquement inévitable. La distinction opérée par les défendeurs entre le modèle, l'algorithme de décodage et l'interface utilisateur est ainsi rejetée comme artificielle au regard des finalités du droit d'auteur.

Dans cette perspective, le *chatbot* est assimilé, pour les contenus mémorisés, à une base de données fonctionnelle.

D - Les exceptions au droit d'auteur ne sont pas applicables

1 - L'inapplicabilité des exceptions de fouille de données

Certes, les modèles de langage comme ceux d'OpenAI sont visés par les exceptions de fouille de données prévues à l'article 44b de l'UrhG et à l'article 4 de la directive (UE) 2019/790, mais pas pour l'ensemble de leur activité. Seule la première phase de développement peut bénéficier de l'exception, celle de la préparation des données. L'idée sous-jacente est que ces reproductions autorisées pour la fouille soient uniquement réalisées à des fins d'analyse ultérieure et ne portent donc pas atteinte aux intérêts d'exploitation de l'auteur sur l'œuvre, ce qui explique qu'il n'y ait pas d'obligation de rémunération. La fouille de données vise à analyser le niveau d'expression, par exemple le choix des mots, les répétitions. Elle a pu permettre ainsi de découvrir dans ses dernières œuvres, les indices d'une maladie d'Alzheimer avancée chez Agatha Christie¹⁸. En « fouillant » les œuvres de la romancière pour y étudier l'évolution de son style, l'évaluation automatisée de son style n'affecte pas l'œuvre, ne s'y substitue pas. Les résultats des fouilles de données en question sont des analyses de ces données.

En revanche, lors de la phase 2 d'entraînement des modèles d'OpenAI, la reproduction des œuvres s'est faite à grande échelle non pas pour leur analyse, mais pour être durablement mémorisées, et pour fournir des résultats, les *outputs*, qui sont les œuvres elles-mêmes ou des résultats s'en inspirant. Les

¹⁴ A. F. Cooper et J. Grimmelmann, *The Files are in the Computer: On Copyright, Memorization, and Generative AI*, 100 *Chi.-Kent L. Rev.* 141 (2025).

¹⁵ Z. Yang et al., *Unveiling Memorization in Code Models*, in *Proceedings of the 2024 ACM/IEEE 44th International Conference on Software Engineering (ICSE '24)*, IEEE Computer Society Press, 2024, p. 867-879.

¹⁶ La perplexité est surtout utilisée pendant l'entraînement et l'évaluation du modèle. Elle mesure à quel point un modèle est « surpris » par un texte qui l'entraîne : plus elle est basse, mieux le modèle pouvait prédire les mots. Par ex., la phrase « Je bois du café » aura une faible perplexité, alors que « Je bois le soleil » aura une perplexité élevée. L'entropie intervient au moment de la prédiction, quand le modèle choisit le prochain mot, car elle mesure le niveau d'incertitude dans un choix. Si le modèle hésite fortement entre plusieurs mots possibles (« Je bois du café, de l'eau, du vin... »), l'entropie est élevée au moment du choix.

¹⁷ Art. 19a, droit de mise à disposition du public : « Le droit de mise à disposition du public désigne le droit de rendre l'œuvre accessible au public par fil ou sans fil, de manière que les membres du public puissent y avoir accès à partir des lieux et aux moments de leur choix. »

¹⁸ Expérience citée par L. Kuschel et D. Rostam (ZUM 2025. 71-73), qui se réfèrent à I. Lancashire et al., *Longitudinal detection of dementia through lexical and syntactic changes in writing: a case study of three British novelists*, *Literary and Linguistic Computing*, vol. 26, n° 4, déc. 2011, p. 435-461.

résultats des usages génératifs sont concurrents des œuvres. Et là, les intérêts légitimes des auteurs sont atteints, « les reproductions dans le modèle nuisent durablement à l'exploitation de l'œuvre », souligne le tribunal.

Le présupposé fondant le régime du *text and data mining* et ses limitations – l'idée que le traitement automatisé de simples données factuelles ne nuit pas aux droits d'exploitation des auteurs – fait ici défaut. Au contraire, les reproductions par le modèle et issues par lui portent atteinte au droit d'exploitation des titulaires de droits. Et le tribunal de rappeler les considérants de la directive (UE) 2019/790 selon lesquels l'introduction de restrictions en matière d'exploration de textes et de données vise non seulement à promouvoir l'innovation et les nouvelles technologies, mais aussi à protéger les auteurs¹⁹.

Selon ces critères, les conditions de la restriction prévue à l'article 4, § 1^{er}, de la directive sur le droit d'auteur dans le marché unique européen (DSM) et à l'article 44b, § 2, de la loi allemande ne sont pas remplies en l'espèce.

OpenAI a bien essayé d'invoquer également l'exception de fouille de données aux fins de recherche scientifique prévue à l'article 3 de la directive et transposée à l'article 60d de la loi allemande, mettant en avant sa qualité d'entreprise dite « d'utilité publique » et d'institut de recherche, mais sans succès. Le tribunal estime à juste titre qu'elle ne l'est pas.

2 - Le rejet des autres limitations

Le droit de citation, pourtant moins exigeant en Allemagne²⁰ qu'en France, est également écarté. En s'appuyant sur l'arrêt *Spiegel Online*²¹, le tribunal rappelle que la citation suppose une finalité d'illustration ou de débat intellectuel, incompatible avec le fonctionnement d'un système dépourvu d'intention et de capacité critique. L'existence d'un prétendu consentement, l'argument tiré du caractère accessoire de l'œuvre²² par rapport à la reproduction, l'atteinte à l'intérêt général des utilisateurs des

modèles, les exceptions de pastiche et de copie privée sont rejetées pour des motifs analogues, tenant tant à l'absence de finalité artistique qu'au caractère commercial et public des usages en cause.

E - Une responsabilité affirmée et une décision à forte portée systémique...

La première réparation que la loi allemande prévoit en cas de violation des droits de l'auteur, est la cessation de l'usage litigieux. Mais si le contrevenant parvient à prouver qu'il n'agit ni intentionnellement ni par négligence, il peut indemniser financièrement la partie lésée si le fait de satisfaire à la cessation lui cause un préjudice disproportionné. Le tribunal est catégorique : certes, la portée du droit de reproduction et les limites du *text and data mining* n'ont pas encore été clarifiées par les Cours suprêmes et la doctrine est encore partagée, mais il est certain que les sociétés OpenAI ont agi avec négligence. Et de souligner que les défenderesses avaient été informées par le groupe de chercheurs de Carlini de leurs résultats et des reprises *verbatim* même avant la publication de l'étude en 2021²³, donc de la possibilité de violations du droit d'auteur par leurs modèles.

Toute invocation d'une erreur de droit excusable est ainsi écartée.

La décision ordonne en conséquence la cessation des reproductions litigieuses, assortie de sanctions dissuasives, et autorise la publication du dispositif du jugement, en raison de l'intérêt général attaché à la clarification du cadre juridique applicable à l'IA générative.

Loin de constituer une réaction de principe hostile à l'innovation, ce jugement rappelle que l'IA générative s'inscrit dans un ordre juridique existant et qu'elle ne saurait prospérer sur une dilution des droits exclusifs des auteurs. Il pose, en creux, une exigence claire : celle de la mise en place de mécanismes de licence ou de corpus autorisés, faute de quoi l'entraînement et l'exploitation des modèles s'exposent à la qualification de contrefaçon.

II - LA DÉCISION GEMA C/ OPENAI AU REGARD DU DROIT FRANÇAIS : UNE SOLUTION LARGEMENT TRANSPOSABLE, VOIRE RENFORCÉE

A - Les points communs certains

Sur la loi applicable, la décision rendue par le Tribunal régional de Munich I dans l'affaire *GEMA*

¹⁹ Consid. 3, phrase 4 ; Consid. 8, 17 et 18.

²⁰ Art. 51, citations : « La reproduction, la diffusion et la communication publique d'une œuvre publiée sont autorisées à des fins de citation, dans la mesure où l'utilisation est justifiée par l'objectif particulier. Cela est notamment autorisé lorsque 1. des œuvres individuelles sont intégrées après leur publication dans une œuvre scientifique indépendante afin d'en expliquer le contenu, 2. des passages d'une œuvre sont cités après leur publication dans une œuvre linguistique indépendante, 3. des passages individuels d'une œuvre musicale publiée sont cités dans une œuvre musicale indépendante. Le droit de citation conformément aux phrases 1 et 2 comprend l'utilisation d'une illustration ou d'une autre reproduction de l'œuvre citée, même si celle-ci est elle-même protégée par un droit d'auteur ou un droit voisin. »

²¹ CJUE 29 juill. 2019, aff. C-516/17, *Légipresse* 2019, 451 et les obs. ; *ibid.* 541, obs. V. Varet ; *ibid.* 2020, 69, étude C. Alleaume ; *ibid.* 322, étude N. Mallet-Poujol ; D. 2019, 1605 ; Dalloz IP/IT 2020, 317, obs. A. Latil ; RTD com. 2020, 83, obs. F. Pollaud-Dulian ; RTD eur. 2019, 927, obs. E. Treppoz ; *ibid.* 2020, 311, obs. F. Benoit-Rohmer ; *ibid.* 324, obs. F. Benoit-Rohmer.

²² Dir. InfoSoc, art. 5, § 3, i), repris in *UrhG*, art. 57 ; v. *CPI*, art. L. 122-5, 11°.

²³ N. Carlini et al., *Extracting Training Data from Large Language Models*, préc., ann. K 23.1, p. 2634 : « [Nous] constatons que plus de 600 des [échantillons examinés] sont des échantillons *verbatim* provenant des données d'entraînement du [modèle]-2 (confirmé en collaboration avec les créateurs du [modèle]-2). » En outre, une scientifique qui était employée par la société mère de la défenderesse au moment de la publication a participé à l'étude en tant qu'auteure.

c/*OpenAI* apparaît, à bien des égards, comme une mise en œuvre jurisprudentielle anticipée des réflexions développées par le CSPLA dans son rapport de décembre 2025 sur la loi applicable aux modèles d'IA générative²⁴.

Le rapport du CSPLA part du constat que les atteintes au droit d'auteur liées à l'IA générative ne se limitent pas à l'*output*, mais trouvent leur source en amont, dans la phase d'*input*, caractérisée par des reproductions transnationales d'œuvres protégées aux fins d'entraînement des modèles. Il insiste sur la difficulté, voire l'artificialité, qu'il y aurait à rattacher ces reproductions à des critères purement techniques tels que la localisation des serveurs, et privilégie une lecture de la *lex loci protectionis* orientée vers le pays dans lequel la protection est revendiquée, c'est-à-dire, en pratique, le territoire de commercialisation et d'utilisation du modèle. C'est précisément cette approche que consacre le Tribunal de Munich lorsqu'il applique globalement la loi allemande, sans distinguer selon que les actes de reproduction liés à l'entraînement auraient pu matériellement intervenir à l'étranger.

La décision *GEMA* illustre également la logique d'éviction de la loi étrangère analysée par le rapport, fondée sur l'ordre public international et la reconnaissance du droit d'auteur comme droit fondamental au sens de l'article 17, § 2, de la Charte des droits fondamentaux de l'Union européenne. En choisissant d'appliquer la loi allemande à l'ensemble des actes litigieux, le juge anticipe la possibilité, envisagée par le CSPLA, d'écartier toute législation étrangère qui priverait les titulaires de droits de la faculté de s'opposer à l'exploitation de leurs œuvres par des modèles d'IA commercialisés sur le territoire de l'Union.

Pour le reste, on remarque l'harmonisation européenne à l'œuvre, car les réponses françaises ne seraient pas sensiblement différentes. **La protection des chansons** en tout ou partie ne soulèverait ainsi pas davantage de difficulté en droit français, en application de l'article L. 112-2 du code de la propriété intellectuelle²⁵.

La question centrale de la **qualification de la mémorisation** des œuvres dans un modèle d'IA comme acte de reproduction appelle, elle aussi, une réponse

largement convergente en droit français, puisqu'elle résulte aussi du droit européen. L'article L. 122-3 du code de la propriété intellectuelle définit la reproduction comme la fixation matérielle de l'œuvre « par tous procédés qui permettent de la communiquer au public de manière indirecte ». Cette définition, volontairement large et technologiquement neutre, a permis de qualifier de reproduction des actes de fixation transitoire ou immatérielle, dès lors qu'ils permettent l'accès ultérieur à l'œuvre.

La restitution des paroles dans les réponses générées par un *chatbot* constituerait, en droit français, un acte distinct de représentation au sens de l'article L. 122-2 du code de la propriété intellectuelle, et un acte de mise à disposition du public, conformément à la directive 2001/29/CE. Le fait que cette restitution soit déclenchée par une requête de l'utilisateur ne devrait pas faire obstacle à la responsabilité de l'exploitant du système, dès lors que celui-ci organise, contrôle et monétise le service. La jurisprudence française, notamment en matière de plateforme

numérique, a déjà admis une responsabilité fondée sur la maîtrise de l'architecture technique et sur le rôle actif de l'opérateur, indépendamment de l'initiative immédiate de l'utilisateur²⁶.

L'examen des exceptions et limitations au droit d'auteur conduirait également, en droit français, à un résultat très proche de celui retenu par le Tribunal de Munich. L'exception de fouille de textes et de données, transposée à l'article L. 122-5-3 du code de la propriété intellectuelle, est interprétée de manière restrictive, conformément à la tradition française et au principe de spécialité des exceptions²⁷. Elle est expressément subordonnée au respect du triple test et ne saurait, en tout état de cause, couvrir la restitution au public des œuvres reproduites. Surtout, la doctrine française est largement majoritaire pour considérer que l'entraînement des modèles d'IA générative excède la simple « analyse » de données au sens de la directive (UE) 2019/790 sur le droit d'auteur, en ce qu'il vise la production de contenus nouveaux susceptibles de concurrencer les œuvres originales²⁸. À cet égard, le droit français offre aux titulaires de droits le levier supplémentaire qu'est le principe d'interprétation stricte des exceptions.

largement convergente en droit français, puisqu'elle résulte aussi du droit européen. L'article L. 122-3 du code de la propriété intellectuelle définit la reproduction comme la fixation matérielle de l'œuvre « par tous procédés qui permettent de la communiquer au public de manière indirecte ». Cette définition, volontairement large et technologiquement neutre, a permis de qualifier de reproduction des actes de fixation transitoire ou immatérielle, dès lors qu'ils permettent l'accès ultérieur à l'œuvre.

²⁴ T. Azzi et Y El Hage, Rapport de mission sur la loi applicable, en vertu des règles de droit international privé, aux modèles d'IA générative commercialisés dans l'UE, CSPLA, 15 déc. 2025.

²⁵ Par ex., Paris, pôle 5 - 2^e ch., 8 mars 2024, n° 22/03274, RTD com. 2024. 275, obs. F. Pollaud-Dulian.

²⁶ Encore très récemment, Com. 7 janv. 2026, n° 24-13.163, D. 2026. 52 ; JT 2026, n° 293, p. 10, obs. X. Delpech.

²⁷ Par ex., Paris, pôle 5 - 2^e ch., 26 mars 2010, n° 08/13185.

²⁸ V. Benabou, IA - Comment associer les créateurs à la valeur produite par l'utilisation de leurs œuvres par les systèmes d'IA ?, CCE 2005, n° 4.

Enfin, la mise en balance des intérêts aurait sans doute été encore plus explicitement marquée en droit français, à la lumière du statut constitutionnel du droit de propriété, dont relève le droit d'auteur. La reconnaissance du droit d'auteur comme droit fondamental, consacrée tant par la Déclaration des droits de l'homme et du citoyen que par l'article 17, § 2, de la Charte des droits fondamentaux de l'Union européenne, pourrait renforcer l'argument selon lequel l'entraînement non autorisé de modèles d'IA générative porte atteinte à l'exploitation normale des œuvres et cause un préjudice injustifié aux intérêts légitimes des auteurs.

B - La grande différence du droit moral d'auteur

Les demandes de condamnation au titre de la violation du droit moral des auteurs et compositeurs, fondée sur l'absence de mention du nom des créateurs, sont rejetées par le juge munichois. Il est fort à parier que le juge français répondrait autrement. Le droit moral reste une compétence nationale, et n'est pas harmonisé ; la vision allemande est très différente du droit français.

Dans ses termes déjà, le droit allemand ne parle pas de « droit moral » mais de « droit de la personnalité d'auteur » (*Urheberpersönlichkeitsrecht*), attaché à la personne et qui s'éteint soixante-dix ans après sa mort, comme le droit patrimonial. Ce droit se réfère ainsi au droit de la personnalité comme droit fondamental, en référence à la Convention européenne des droits de l'homme et à l'article 1^{er} (dignité humaine), et de l'article 2, § 1^{er} (libre développement de la personnalité) de la Loi fondamentale (*Grundgesetz*) ou Constitution allemande. En tant que droit fondamental, elle doit être mise en balance avec les autres droits fondamentaux. La violation du droit au respect du nom des auteurs non mentionnés est ainsi rejetée, car elle ne conduit pas en l'espèce à une atteinte à la personnalité de l'auteur. Les juges estiment que l'absence de nom n'aurait pu être sanctionnée que si elle s'était accompagnée d'une erreur sur l'attribution de l'œuvre (BGH 08.06.1989, I ZR 135/87, *Emil Nolde*), ou d'une atteinte à l'image de l'auteur, ce qui ne serait pas le cas en l'espèce. Le tribunal précise que la mention des auteurs lors de la modification et de l'adaptation de leurs œuvres relève exclusivement de leur contexte professionnel, c'est-à-dire de leur vie sociale, et non de leur vie privée. Or une atteinte à la personnalité ne pourrait être caractérisée que dans des hypothèses exceptionnelles d'atteintes graves à la sphère sociale, telles que la stigmatisation ou l'exclusion sociale, ce qui n'est pas le cas en l'espèce. Le juge n'estime pas nécessaire de répondre au reproche de déformation de l'œuvre (UrhG, art. 14), équivalent de l'atteinte au respect de l'œuvre de l'article L. 121-1 du code de la

propriété intellectuelle, puisqu'il a donné droit à la demande de cessation.

Au contraire, on sait qu'en droit français, le droit moral est un droit autonome, distinct du droit de la personnalité. Il est imprescriptible, inaliénable et perpétuel (CPI, art. 121-1). L'atteinte au droit de paternité ou à l'intégrité de l'œuvre ne suppose aucun risque de confusion d'identité, ni aucune atteinte à la vie privée ou à la sphère personnelle de l'auteur. Il suffit que l'œuvre soit modifiée sans autorisation ou que le nom de l'auteur soit omis ou altéré pour que la violation soit constituée. Si l'affaire avait été jugée en droit français, la solution aurait sûrement été différente sur ce point ; la modification ou l'adaptation de l'œuvre aurait été analysée comme une atteinte au droit au respect de l'intégrité de l'œuvre et l'absence ou l'altération de la mention du nom de l'auteur aurait pu constituer une atteinte au droit de paternité, sans qu'il soit besoin d'invoquer une atteinte à la vie privée ou à la sphère sociale.

III - BONUS : UN PEU D'ORDRE DANS LA JURISPRUDENCE ACTUELLE – VERS UNE RECOMPOSITION DES LIGNES DE FRACTURE DU DROIT D'AUTEUR FACE À L'IA

Pris isolément, le jugement rendu par le Tribunal régional de Munich I dans l'affaire *GEMA c/ OpenAI* pourrait être analysé comme une application rigoureuse, mais somme toute classique, des principes fondamentaux du droit d'auteur européen. Pourtant, replacée dans le contexte jurisprudentiel plus large dessiné par les décisions de Hambourg et américaines, cette décision révèle des lignes de tension profondes et, surtout, une fragmentation croissante des approches judiciaires face à l'entraînement des IA.

Dans le jugement rendu par le Tribunal régional de Hambourg le 27 septembre 2024 dans l'affaire *LAION e.V. contre Kneschke*, confirmée en grande partie en appel²⁹, la juridiction hambourgeoise a admis que l'intégration d'une photographie protégée dans un vaste ensemble de données destiné à l'entraînement d'un modèle d'IA relevait de l'exception de fouille de textes et de données prévue par l'article 4 de la directive (UE) 2019/790. Le tribunal a estimé que la constitution d'un *dataset*, même à grande échelle, ne portait pas atteinte à l'exploitation normale de l'œuvre ni aux intérêts légitimes de l'auteur, dès lors qu'aucune reproduction directe de l'image n'était démontrée dans les sorties du modèle. Cette décision, contestée en appel et aujourd'hui devant la Cour suprême allemande, reposait sur une lecture extensive de l'exception de *data mining* et reléguait au second

²⁹ M.-A. Roux Steinkühler, La fâcheuse et contestable affaire *Kneschke contre Laion e.V.*, préc.

plan le risque de concurrence potentielle entre l'œuvre humaine et les contenus générés par l'IA.

Le jugement *GEMA* marque, sur ce point, une inflexion nette, car le Tribunal de Munich refuse de considérer l'entraînement comme une zone juridiquement aseptisée, indépendante des usages ultérieurs du modèle. Cette divergence révèle une hésitation structurelle du droit européen entre, d'une part, une logique fonctionnelle, favorable à l'innovation et fondée sur la fouille de données, et, d'autre part, une logique patrimoniale, attentive à la maîtrise par l'auteur des nouveaux modes d'exploitation de son œuvre.

La mise en perspective avec la décision américaine du Tribunal de district du Delaware du 11 février 2025, *Thomson Reuters c/Ross Intelligence*, permet d'éclairer encore différemment les enjeux. Dans cette affaire, le juge Bibas a rejeté l'argument du *fair use*, pourtant traditionnellement central en droit américain, en mettant au cœur de son raisonnement la question de la concurrence et de la substitution de marché. En estimant que l'IA du demandeur Ross poursuivait le même objectif économique que les défendeurs Westlaw et qu'elle pouvait se substituer au produit original, le tribunal a opéré un recentrage du *fair use* autour du critère de l'impact sur le marché, dans la droite ligne de la décision *Andy Warhol Foundation v. Goldsmith* rendue par la Cour suprême.

À première vue, les raisonnements du Delaware et de Munich reposent sur des fondements juridiques distincts. Le premier mobilise une exception ouverte et souple, le *fair use*, tandis que le second applique un régime européen d'exceptions limitatives, interprétées strictement. Pourtant, une convergence conceptuelle apparaît. Dans les deux cas, l'IA est appréhendée comme un produit susceptible d'entrer en concurrence avec les œuvres ou services existants, et non comme un simple outil technique neutre.

La différence majeure réside toutefois dans le traitement de la phase d'entraînement elle-même. Là

où le juge américain raisonne principalement à partir de l'effet final sur le marché et sur la substituable des produits, le Tribunal de Munich s'attache à la qualification juridique des actes intermédiaires, en retenant que la mémorisation constitue déjà une reproduction illicite, indépendamment même de la démonstration d'un préjudice économique immédiat. Cette approche est révélatrice de la tradition européenne de protection forte des droits patrimoniaux, fondée sur un droit d'interdiction préalable, là où le droit américain accepte plus volontiers certaines atteintes dès lors qu'elles ne bouleversent pas l'équilibre économique global.

Ces trois décisions, mises en regard, dessinent ainsi un paysage jurisprudentiel fragmenté. Hambourg illustre une lecture permissive, voire optimiste, des exceptions de fouille de données, au risque de sous-estimer les effets cumulatifs de l'entraînement massif des modèles. Le Delaware marque un tournant du *fair use*, en remplaçant la concurrence et la substitution au centre de l'analyse. Munich, enfin, propose une voie européenne alternative, consistant à réaffirmer la centralité de la notion de reproduction et à refuser toute dilution des droits exclusifs au stade même de l'entraînement.

Il n'en demeure pas moins que ces décisions laissent subsister une incertitude profonde. En Europe, l'articulation entre la directive DSM, l'interprétation des exceptions de *data mining* et les obligations de transparence issues de l'*AI Act* demeure largement inachevée. Aux États-Unis, l'avenir du *fair use* appliqué à l'IA dépendra de l'orientation que prendront les juridictions fédérales dans les litiges pendants, notamment ceux opposant les éditeurs de presse aux acteurs de l'IA générative. Dans ce contexte, le jugement *GEMA c/OpenAI* ne clôt pas le débat, mais il en redéfinit les termes, en rappelant que l'entraînement des IA n'est pas un simple préalable technique, mais bien un acte juridiquement signifiant, susceptible de porter atteinte aux fondements mêmes du droit d'auteur.