



European
Writers'
Council

IA : BOITE A OUTILS POUR LE SECTEUR

Recommandations de l'European Writers' Council (EWC) pour les auteurs et les traducteurs, les éditeurs, les libraires, les organisateurs d'événements et autres parties prenantes du secteur du livre en matière d'accords bilatéraux et contractuels ainsi que d'exigences techniques, en ce compris, les procédures dites d'opt-out (droit d'opposition).



BOITE A OUTILS IA POUR LE SECTEUR DU LIVRE

Recommandations de l'European Writers' Council (EWC) pour les auteurs et les traducteurs, les éditeurs, les libraires, les organisateurs d'événements et autres parties prenantes du secteur du livre en matière d'accords bilatéraux et contractuels ainsi que d'exigences techniques.

Cette boîte à outils est destinée aux décideurs politiques ainsi qu'aux entreprises du secteur de l'IA devant mettre en place une documentation transparente sur leurs données d'entraînement (training data) et sur leurs obligations de rémunération.

Cette boîte à outils IA ne possède pas de caractère contraignant et ne vise pas à remplacer ou à prescrire les accords nationaux.

L'European Writers' Council (EWC) est le plus grand groupe international d'auteurs du secteur du livre tous genres confondus (fiction, non-fiction, littérature académique et de jeunesse, poésie, etc.). Fort de 50 organisations et fédérations professionnelles issues de 32 pays de l'UE, de l'EEE et des pays européens non membres de l'UE, l'EWC représente 220 000 écrivains et traducteurs. Ces personnes écrivent et publient en 34 langues, dans le monde entier.

Les écrivains, les illustrateurs, les concepteurs de couvertures, les traducteurs, les interprètes de livres audio, ainsi que les éditeurs, les rédacteurs en chef et les organismes de gestion collective (OGC) et de gestion de droits (ODR) sont directement et immédiatement concernés par les répercussions de ce que l'on présente comme l'arrivée de l'IA dans le secteur du livre, et en particulier par la production et l'utilisation de systèmes informatiques génératifs. L'EWC et son groupe d'experts ont élaboré cet ensemble de recommandations afin de contribuer à l'établissement d'un socle de pratiques équitables pour le secteur du livre et pour les relations entre, d'une part, les auteurs avec leurs éventuels agents, traducteurs, éditeurs et, le cas échéant, les libraires ou les organisateurs d'événements ; et, d'autre part, les développeurs d'IA (génératives).

Vous trouverez dans cette boîte à outils :

- Les juridictions concernées et les exceptions controversées relatives à la fouille de textes et de données (TDM pour *Text and Data Mining*)
- Toutes les IA ne sont pas concernées : définition des types d'applications visées par cette boîte à outils
- 10 recommandations pratiques pour des relations commerciales équitables
- Fondements juridiques et mise en œuvre des 10 recommandations

Remarque : *Ces recommandations visent les territoires où la directive 2019/790 (UE), relative au droit d'auteur dans le marché unique numérique (directive CDSM), s'applique et prévoit en conséquence un régime d'exception concernant le TDM (art. 4). Elles s'appliquent également aux traductions publiées dans les États membres de l'UE. Tous les autres points abordés peuvent également s'appliquer à l'international.*

Veillez noter que la question de savoir si la fouille de textes et de données (ou « TDM ») telle qu'elle est décrite dans l'article 4 de la directive CDSM qui couvre l'utilisation subséquente d'œuvres à des fins de développement d'IA génératives reste juridiquement controversée. L'IA Act (Règlement européen sur l'IA) du 12 juillet 2024 fait référence à la directive CDSM dans ses considérants. Pour autant, elle ne clarifie pas formellement cet aspect. Nous anticipons donc des contentieux sur ce point et/ou une clarification de la part de la Commission lors de l'évaluation de la directive CDSM dans le courant 2025-2027. La position de l'EWC est de recommander, par mesure de précaution, l'exercice du droit d'opposition, dit TDM opt-out pour empêcher l'utilisation de tout contenu par une IA générative (IAg). Toutefois, en parallèle, nous jugeons également très probable l'apparition d'un nouveau droit exclusif concernant l'exploitation pour le développement d'une IA générative. Ainsi, nous recourons à l'opt-out comme protection temporaire, jusqu'à ce que la situation juridique et la question du TDM n'incluant pas le développement d'IAg aient été formellement clarifiées.

PARTIE I

TOUTES LES IA NE SONT PAS CONCERNEES : DEFINITION DES TYPES D'APPLICATIONS VISEES PAR CETTE BOITE A OUTILS

Pléthore de prétendus « modèles économiques fondés sur l'IA » ont d'ores et déjà émergé dans le secteur du livre, avec leur lot de faux auteurs, de faux livres, mais aussi de faux lecteurs. On peut supposer que les bases des grands modèles de langage (LLM), tels que GPT, Meta, StableLM ou BERT, ont été générées à partir d'ouvrages protégés par le droit d'auteur, pour partie extraits de bibliothèques clandestines telles que Library Genesis (LibGen), Z-Library (Bok), Sci-Hub et Bibliotik, autrement dit : des sites Internet pirates. En l'absence de réglementation, les technologies génératives accélèrent et favorisent l'exploitation éhontée des droits d'auteur et font croire à la légitimité de leur violation, en plus d'encourager la distorsion de l'information et de la communication, et de favoriser les fraudes aux redevances et à la rémunération des licences collectives.

En parallèle, un examen approfondi et une évaluation sont indispensables à la catégorisation et à la régulation des différents aspects de l'« informatique avancée ». Puisque les logiciels peuvent être "intelligents" sans employer l'IA, toutes les applications ne présentent pas un risque égal.

L'EWC distingue tout d'abord trois types de systèmes :

- Les logiciels et systèmes informatiques d'assistance, n'étant considérés ni comme de l'IA, ni comme un risque ;
- Les systèmes informatiques d'analyse, étant considérés en partie comme de l'IA et comme un risque potentiel ;
- Les systèmes informatiques génératifs, qui forment la catégorie d'IA considérée comme un risque dans les domaines du texte, de la voix et de l'image : l'intelligence artificielle générative, en abrégé : « IAg ».

Ce document se concentre sur les aspects juridiques, administratifs et techniques de ce que l'on appelle l'« IA générative » (IAg) et sur les pratiques qui s'y rapportent.

- Nous y précisons la situation juridique relative à l'amont (procédures contractuelles et techniques) et à l'aval (en ce qui concerne les questions d'identification et la transparence requise).
- Ce document aborde la robotique textuelle automatisée (par exemple, GPT), les machines de traduction automatique (DeepL), la génération d'images (Midjourney), et le clonage synthétique de la voix humaine ou autre utilisation de voix générées par l'intelligence artificielle générative.

Ce travail tient compte, d'une part, de l'IA Act et, d'autre part, des cadres juridiques, européens ou non, relatifs à la fouille de textes et de données ainsi qu'aux questions de propriété intellectuelle encore en suspens : notamment la définition de l' "apprentissage machine" (*machine learning*) dans lequel nous voyons plutôt une forme de "programmation algorithmique" et qui comprend les processus de reproduction préalables au développement d'IA génératives tels que : le *scraping*, la conversion temporaire de fichiers .pdf, .mobi, ou .epub au format .xml, la copie continue visant à former un (ou des) corpus lexicaux, le dépôt de fichiers source à des fins de reproductibilité ou de vérifiabilité, ou encore la copie, le stockage et la décomposition contextuelle ainsi que la reproduction de l'expression individuelle au sein de grands modèles linguistiques artificiels, y compris les aspects de proximité et d'imitation stylistique, les modèles de diffusion d'images et le clonage vocal synthétique.

Remarque : ne sont pas couverts par cette boîte à outils les technologies et logiciels d'analyse ou d'assistance tels que les programmes d'analyse sémantique et de relecture (par exemple, Word Editor), de retouche d'image (Photoshop), de gestion de base de donnée, de classement, de conversion, de répertoire de citations, de story-boarding, de synthèse textuelle pour la génération de métadonnées, de mixage et de montage audio, de processus d'inventaire automatisés, pas davantage que les outils de TAO utilisés par les traducteurs.

PARTIE II

10 RECOMMANDATIONS PRATIQUES POUR DES RELATIONS COMMERCIALES ÉQUITABLES

1. **Autorisation d'exploitation pour la fouille de textes et de données ou pour le développement d'IAg :** Le consentement éclairé et la permission écrite des auteurs, des artistes plasticiens et des interprètes de livres audio posent les fondements du respect de leurs droits en matière de propriété intellectuelle pour toute utilisation de leurs œuvres textuelles, plastiques ou sonores pour (a) le TDM, (b) le *scraping*, (c) toute étape associée à la programmation d'algorithmes pour le développement d'IA génératives. Par conséquent, il est nécessaire de développer des procédures liées au contrat et à la communication concernant l'opt-out face au TDM. De plus, les développeurs d'IA doivent obtenir des licences individuelles (via les éditeurs) ou collectives (via les organisations de gestion collective / CMO). Les auteurs qui souhaitent protéger leurs œuvres (publiées ou inédites) contre le *scraping*, le TDM et toute autre forme d'exploitation pour l'IAg doivent réclamer à leurs éditeurs, par écrit, une clause d'opt-out.

2. **Rémunération** : Les auteurs, artistes et interprètes qui autorisent, en toute connaissance de cause et sans contrainte, l'exploitation de leurs œuvres ou interprétations à des fins de fouille de textes et de données ou de programmation d'IAg doivent recevoir une rémunération appropriée et proportionnelle, dans le respect des modèles de licence, à durée et à portée limitées. Ces modèles de licence doivent être transparents concernant leur objectif et l'usage qui sera fait des œuvres, et respecter des obligations claires en matière de transparence sur ces usages, afin de garantir une rémunération adéquate et régulière. Il revient aux législateurs et aux parties prenantes de chaque pays de décider si cette gestion s'effectue par le biais de licences individuelles ou collectives, par l'intermédiaire des OGC et des ODR.
3. **Transparence (en amont)** : L'IA Act impose aux *scrapers* et aux *crawlers*, ainsi qu'aux concepteurs de corpus et, enfin, aux développeurs d'IA, de fournir un résumé suffisamment détaillé des titres, des auteurs et des sources auxquels ils ont légalement eu accès, ainsi que de leurs méthodes d'acquisition d'œuvres protégées, en incluant toute donnée pertinente relative à la propriété intellectuelle. Le secteur du livre peut à cet effet développer et mettre en œuvre des normes adaptées et harmonisées. Actuellement, les solutions techniques suivantes sont en usage ou en cours de développement, pour permettre de déclarer un opt-out ou de fournir des informations de licence : les métadonnées, l'International Standard Content Code (ISCC) associé à une déclaration de droits (par exemple, mais pas exclusivement, à l'aide des logiciels de Liccium ou de Creators' Credentials), le Digital Object Identifier (DOI), l'ISBN (International Standard Book Number) et le TDM Reservation Protocol du W3C (World Wide Web Consortium) dans ONIX. Ces systèmes permettent non seulement de déclarer des opt-outs, mais facilitent également le suivi des œuvres pour lesquelles les développeurs d'IAg ont acquis une licence, et notamment des sources (Internet et autres) auxquelles ils ont eu accès.
4. **Transparence (en aval)** : Tout texte automatisé, y compris ceux obtenus par traduction automatique, tout produit visuel généré par IA, ainsi que tout produit audio synthétique devraient, lors de leur publication, être identifiés comme des produits générés par IA. On pourrait également envisager d'identifier les œuvres 100 % humaines, sur le modèle des « *Trusted Shops* » (Marque de Confiance).
5. **Communication claire et respect du droit moral des auteurs** concernant l'intégrité de leur œuvre. Les éditeurs et autres parties contractantes doivent obtenir l'accord des auteurs avant de recourir à l'IA générative pour leurs œuvres et établir une entente mutuelle quant à l'utilisation des différents types de logiciels et outils informatiques avancés, tels que, par exemple : les voix de synthèse dans les livres audio, la traduction automatique, l'illustration générative de couvertures, et toute autre adaptation des œuvres par IA générative. Les auteurs devraient être en droit de préférer recourir au travail humain en refusant les couvertures générées par IA, les adaptations de livres audio IA ou les traductions IA de leurs œuvres sans subir de préjudice ni de conséquences négatives, telles qu'une diminution de leurs droits d'auteur. Les éditeurs devraient également pouvoir obtenir la garantie de savoir si une œuvre contient des éléments générés par IA.
6. **Respect des choix méthodologiques personnels des auteurs, traducteurs, artistes et interprètes** : Aucun auteur, traducteur, interprète ou illustrateur ne devrait être contraint d'employer une IAg ou de travailler à partir d'un texte généré par IA, ni de traduire automatiquement un texte ou d'utiliser des images générées par IAg.

7. **Bonne traçabilité de l'information sur les licences et les opt-outs :** Les éditeurs doivent préciser aux tierces parties, parmi lesquelles on trouve les plateformes, les agrégateurs, les banques de contenus ou les distributeurs commerciaux, si elles ont le droit ou non de délivrer des sous-licences, de reproduire et/ou d'utiliser l'œuvre de quelque manière que ce soit à des fins de fouille de textes et de données ou de programmation de technologies génératives (IAg). Un protocole de droit de réservation (« opt-out TDM / l'IA / IAg ») doit être communiqué pour chaque fichier de titre, par exemple dans ses métadonnées ou par d'autres procédures de gestion des droits numériques telles que, par exemple : la déclaration de droits ISCC+ ou le protocole de réservation de droits relatif au TDM (TDMRep)+ONIX.
8. **Si nécessaire, les sites web personnels des auteurs, artistes et interprètes doivent pouvoir** déclarer leur droit de réservation en matière de TDM, en vertu de l'art. 4 de la directive 2019/790 du CDSM, en précisant clairement qu'ils n'accordent pas d'autorisation pour le TDM, ni pour le *scraping* et l'exploitation à des fins de *machine learning* ou de programmation d'algorithmes d'IA générative. Ceci doit être rendu possible et conforme aux exigences par l'inscription dans les conditions générales d'utilisation (CGU) ou dans les mentions légales d'un site, mais surtout par une indication lisible par une machine, par exemple dans le fichier robots.txt de l'URL du site, ou bien via le protocole de réservation de droit relatif au TDM (TDMRep). Les éditeurs doivent apposer sur le site Internet de leur société un avertissement d'opt-out relatif à la fouille de textes et de données, soit pour chaque titre concerné, soit pour l'ensemble de leur catalogue. Il en va de même pour les distributeurs et les libraires recourant à des sites de vente en ligne. Au-delà des sites Internet, c'est-à-dire pour l'œuvre elle-même, on peut recourir à d'autres normes techniques et lisibles par machine, par exemple en liant une déclaration de droits ISCC+ aux ressources, ou bien à l'aide des métadonnées des fichiers de l'œuvre.
9. **Vérification des CGU des logiciels employés :** Les sociétés d'édition ou d'agrégation ainsi que les agents, les éditeurs et les traducteurs doivent vérifier les logiciels qu'ils utilisent pour s'assurer qu'ils ne déclarent pas, dans leurs CGU, être autorisés à scraper, utiliser, copier ou stocker tout contenu en vue du développement, de l'amélioration ou de l'augmentation d'une IA, notamment générative. Ceci s'applique également aux plateformes, aux réseaux sociaux et aux portails où des captations vidéo de lectures publiques ou de débats d'experts peuvent être publiées.
10. **Chacun se doit d'être conscient de ses propres responsabilités éthiques.** La majeure partie des grands systèmes d'IAg actuels sont présumés avoir été construits sur la base d'une violation des droits d'auteur. Les œuvres et les investissements des auteurs et des éditeurs ont été exploités à leur insu, sans autorisation, rémunération ou transparence, du fait notamment des sites de piratage, et bien du fait de la non-rétroactivité des exceptions relatives au TDM fixées par la directive 2019/790 du CDSM. Le secteur devrait avoir pour ambition d'assurer son avenir en menant un effort commun pour dénoncer ces dommages et obtenir une compensation financière, ainsi que pour faire pression en vue de la neutralisation de ces systèmes d'IAg, si cela s'avère nécessaire. On prétend souvent à tort que l'« apprentissage » des machines est équivalent à la lecture humaine et qu'il s'agit donc d'un « droit ». Il faut que les entités nationales, en coopération notamment (mais pas seulement) avec les OGC et les parties prenantes, procèdent à l'évaluation des dommages économiques déjà causés par le développement incontrôlé de l'IAg. **Des actions collectives, tant sur le plan européen que national, ainsi que des réglementations internationales visant à responsabiliser les développeurs d'IA, sont une nécessité.**

Une coopération en faveur d'un « code de bonne conduite de l'IA et de la propriété intellectuelle » ainsi que d'une réglementation permettant d'assurer un avenir équitable, éthique et réglementé garantira la protection de la puissance innovatrice et de la créativité humaines et préservera les connaissances et la culture du bouleversement technologique.

En tant que professionnels du livre, nous sommes à l'orée des discussions dont la poursuite déterminera l'avenir des générations futures. Puissent ces recommandations de l'EWC en être l'impulsion initiale.

PARTIE III

FONDEMENTS JURIDIQUES ET MISE EN ŒUVRE DES 10 RECOMMANDATIONS

1. Questions contractuelles

Présentation de la base juridique des recommandations :

- 1.1. Droits de propriété intellectuelle des auteurs (copyright, droits d'auteur) liés à la fouille de textes et de données (TDM) à des fins commerciales et aux nouvelles formes d'exploitation actuelles : le *scraping*, la copie, le stockage et toute autre forme d'exploitation en vue de l'« apprentissage » machine, autrement dit : de la programmation d'algorithmes pour l'IA générative.
- 1.2. Droits moraux des auteurs : l'intégrité de l'œuvre, incluant sa traduction, couverture et narration audio.
- 1.3. Accords pratiques en matière de transparence et d'évaluation de la juste rémunération des auteurs.

1.1. Droits de propriété intellectuelle des auteurs : (I) Opt-out (TDM) et (II) Réserve de droits concernant tous les processus relatifs à la programmation d'algorithmes visant à développer une IA :

En application des articles 9.1, 9.2 et 9.3 de la Convention de Berne, tout auteur, artiste et interprète audio a le droit de déterminer les modalités de publication, de distribution, de copie et d'utilisation de son œuvre, sous réserve d'exceptions, de limitations ou d'autres accords contraignants prévus par les législations nationales ou transnationales. Ceci inclut la décision de savoir si l'œuvre peut être copiée et utilisée (a) pour la fouille de textes et de données (à des fins commerciales ou non), (b) dans un but de *scraping*, de copie et de stockage, et (c) pour la programmation (« l'apprentissage ») de machines à des fins générales ainsi que pour le développement de programmes informatiques génératifs (texte, image, voix) et d'IA. L'EWC considère les utilisations mentionnées aux points b) et c) comme de nouveaux types d'utilisations précédemment non couvertes par les contrats d'auteur ou les exceptions légales. Toute exploitation d'une œuvre nécessite le consentement écrit de son auteur.

Dans l'UE, les exceptions contenues dans l'art. 3 (TDM non commercial) et l'art. 4 (TDM commercial) de la directive 2019/790 sur le droit d'auteur dans le marché unique sont entrées en vigueur le 7 juin 2021. L'article 4 autorise la fouille de textes et de données à des fins commerciales. Le seul moyen pour les auteurs et les éditeurs de s'opposer à cette exploitation non rémunérée est d'exercer leur droit de déclarer une réserve de droits dans un format lisible par une machine ou d'une autre manière suffisante. C'est ce qu'on appelle l'« opt-out TDM » ou la « réserve de droits TDM ».

Il reste encore de nombreuses questions à clarifier, notamment : comment traiter les œuvres antérieures à 2021, si elles doivent également faire l'objet d'un opt-out lisible par une machine ; comment prouver leur utilisation évidente par les sociétés du secteur de l'IA ; et enfin comment cette utilisation doit être répertoriée et, le cas échéant, rémunérée. Tant que le droit européen en vigueur n'aura pas clarifié, par l'intermédiaire de la Commission ou des Cours de justice par exemple, si la programmation algorithmique de substituts économiques d'œuvres d'auteurs (IAg) est couverte par l'exception de l'article 4 de la directive CDSM 2019/790 et par ses mises en œuvre nationales, **l'opt-out TDM contractuel ou autrement accordé par l'auteur** (écrivain, traducteur, illustrateur, artiste-interprète) à des fins commerciales constitue le principe général que les auteurs et éditeurs peuvent utiliser pour protéger leurs droits en attendant ladite clarification. En outre, l'opt-out confère à l'auteur le droit de concéder des licences s'il le souhaite, qui peuvent alors être gérées par les éditeurs ou les OGC.

→ Pour les contrats déjà existants, il est recommandé de faire un avenant couvrant la réserve de droits en matière de gestion du TDM. Néanmoins, modifier les contrats en vigueur (au nombre de 13,8 à 20 millions rien qu'en Europe) exigera des efforts administratifs et humains considérables. En l'occurrence, les éditeurs et les auteurs doivent définir des procédures pour les œuvres déjà présentes sur le marché, et notamment déterminer si, par exemple, les éditeurs devront systématiquement déclarer un opt-out, sous réserve d'une autorisation expresse du TDM par l'auteur en vertu de l'article 4 de la directive sur les droits d'auteur.

→ De leur côté, les auteurs devront exiger de leurs éditeurs qu'ils attachent une déclaration d'opt-out par écrit à tous les ouvrages publiés et à venir.

Les nouveaux contrats comprenant des accords portant sur les droits étrangers/ les accords de traduction devront inclure une clause, ou toute autre forme de confirmation écrite, à cet effet, afin de coordonner l'opt-out. Par exemple :

- Ce contrat ne transfère pas le droit à la fouille de textes et de données à des fins commerciales (droit de réserve TDM en vertu de l'art. 4 de la directive CDSM 2019/790 (UE)). L'auteur réclame que l'éditeur l'informe de manière obligatoire de cette déclaration d'exclusion du TDM.
- L'auteur se réserve expressément le droit exclusif de procéder à de la fouille de textes et de données, à du *scraping*, à de la copie et à du stockage, ainsi qu'à de la programmation d'algorithmes à des fins générales, y compris, mais sans s'y limiter, à des fins de production d'une IA générative. L'auteur réclame que l'éditeur l'informe de manière obligatoire de cette déclaration d'exclusion du TDM.

- L'éditeur prendra toutes les mesures nécessaires pour communiquer la réservation de droits TDM commercial de manière appropriée et efficace y compris, mais sans s'y limiter, dans les métadonnées et ONIX, avec des indications lisibles par machine sur les sites Internet ou les mentions légales des sites Internet, et s'assurera que les détenteurs de licences et/ou les distributeurs communiquent également toutes les réservations de droits associées.

Cette trace écrite permettra aux éditeurs d'agir en toute conformité et de faire figurer sur les livres à publier (livres électroniques, livres audio numérisés et livres imprimés) cet opt-out TDM, conformément aux exigences de l'article 4 de la directive CDSM 2019/790 (au sein des métadonnées et d'ONIX, des mentions légales, dans l'identifiant de la déclaration de droits ISCC+, ou via TDMRep).

Quelques remarques pour compléter ce tableau des réservations de droits TDM :

- Les traducteurs sont également tenus de respecter les droits de réservation relatifs à la fouille de textes et de données ou à la programmation d'IA génératives, et ne doivent pas télécharger de textes dans des logiciels de traduction automatique sans que l'auteur ou le titulaire des droits en ait connaissance et y ait consenti par écrit ;
- Les libraires ou autres intermédiaires tels que les bibliothécaires ne doivent pas non plus introduire de textes ou de fichiers dans GPT, par exemple pour générer des résumés, car cela constituerait une violation de la clause d'opt-out ;
- Les bibliothèques qui donnent accès à des œuvres numérisées en vertu de l'arrêt Ulmer c. TU Darmstadt (C-117/13, 2014) ne sont pas autorisées à octroyer des sous-licences ou à donner accès à ces œuvres à des fins commerciales ;
- Dans le cadre d'un système d'édition fondé sur l'informatique dématérialisée (« cloud »), il convient d'indiquer aux auteurs si et à quelles étapes du processus un éditeur introduit des textes ou des œuvres d'art dans des systèmes d'IAg, par exemple pour en extraire des mots-clés, des résumés ou d'autres informations, car ceci pourrait également être caractérisé comme une atteinte aux réservations de droits des auteurs. Dans ce contexte, toutes les parties prenantes doivent échanger leurs informations de manière ouverte et transparente et parvenir à un accord et à une compréhension mutuelle du travail éditorial ;
- Alors que l'art.8 (2019/790 (EU)) autorise certaines entités à numériser des œuvres ne relevant pas du domaine commercial, ce même article ne confère pas le droit de donner accès à ces œuvres pour le TDM commercial ou la programmation d'algorithmes (développement d'IAg.). Les entités habilitées doivent veiller à ce qu'une clause d'opt-out lisible à la machine soit appliquée à ces œuvres (par exemple, mais pas exclusivement, aux archives numérisées et rendues publiques en ligne) ; elles ne peuvent pas empêcher les titulaires de droits d'exercer leur réservation de droits.

1.2. Droits moraux des auteurs

Les droits moraux des auteurs comprennent le droit au respect de l'intégrité de leur œuvre – i.e. « droit à l'intégrité ». En pratique, ce droit est déjà respecté depuis des décennies, par exemple lorsqu'un auteur vérifie des épreuves et les approuve avant qu'elles ne soient imprimées, ou lorsqu'il donne son consentement écrit à une version abrégée de son œuvre pour produire un livre audio ou une édition abrégée. De la même manière, aucun auteur n'aurait jamais pu imaginer que

ses œuvres et sa façon spécifique de s'exprimer seraient dispersées en millions d'éléments pour le développement des systèmes d'IAg. Cette utilisation numérique imprévue de leur œuvre dans un but autre que celui de transmettre l'art constitue une atteinte au droit à l'intégrité des auteurs.

Dans le domaine de l'IAg, le droit moral de l'auteur sur l'intégrité de son travail s'étend :

- Aux éditions audio
- Aux traductions
- Au matériel supplémentaire : notamment les couvertures et les illustrations
- A la transformation et à la présentation de leurs œuvres

Le droit moral comprend également le droit de paternité (c'est-à-dire le droit de faire figurer le seul nom de l'auteur sur l'œuvre). Les logiciels d'IAg n'ont aucune légitimité à figurer dans les mentions légales au même titre qu'un auteur humain.

Accords sur la production de livres audio

- Les auteurs et les éditeurs doivent consentir à ce que l'auteur ait le droit de refuser qu'une licence soit accordée pour l'édition d'un livre audio de son œuvre faisant appel à des voix artificielles et/ou synthétiques. En principe, l'auteur doit donner son autorisation par écrit et ne doit pas craindre d'être désavantagé s'il refuse les voix produites par IAg ou s'il insiste pour que des narrateurs humains soient engagés dans la production d'un livre audio sous licence. Certains éditeurs et producteurs de livres audio auront un point de vue différent, par exemple sur la base de considérations économiques. Espérons que l'intérêt commun pour l'appréciation du travail humain et des compétences culturelles continuera à prévaloir.
- Remarque : Les exigences du Règlement Européen sur l'Accessibilité (EAA, Directive 2019/882), qui entrera en vigueur en 2025, permettent des adaptations de texte vers la parole des livres électroniques (« *text to speech* »), qui seront effectuées par des voix générées par IA sur les appareils des utilisateurs. Ceci ne peut vraisemblablement pas être exclu contractuellement, mais peut être limité par une définition spécifique, telle que : « la synthèse vocale par IA n'est permise que dans le contexte d'une synthèse vocale légalement autorisée en vertu de la Directive 2019/882. Toute utilisation supplémentaire du texte mis en audio de cette manière pour le TDM ou la programmation algorithmique est prohibée. » Sur le plan politique, il serait souhaitable de parvenir à un accord avec les législateurs pour que toute édition d'un livre électronique utilisant l'IA pour la synthèse vocale ne concurrence pas les éditions originales de livres audio effectuées par des narrateurs humains.
- Si les auteurs enregistrent eux-mêmes un livre, il doit être stipulé dans un accord contractuel distinct que tout enregistrement à des fins de clonage vocal n'est pas autorisé, sous réserve de leur consentement explicite assorti d'une rémunération.
- En conséquence, les éditeurs de livres audio doivent répondre aux souhaits de l'auteur de se soustraire au TDM audio et à l'utilisation de l'interprétation du narrateur à des fins de reproduction vocale synthétique, et prévoir à ces fins une clause d'opt-out lisible par une machine.

Accords sur les traductions en cas de transfert de droits étrangers

- Afin de préserver l'intégrité de l'œuvre traduite et de réserver le droit à son utilisation pour le TDM et pour le développement ou l'amélioration de toute IA_g, l'auteur a le droit de refuser que son œuvre soit traduite, que ce soit intégralement ou partiellement, par une traduction automatique, et d'exercer son droit d'approbation sur la traduction automatique, y compris le droit de refuser la pré-traduction par une machine et la post-édition par un traducteur. Ce point est particulièrement important si l'auteur veut exercer son droit légitime de ne *pas* voir ses textes utilisés pour le développement de l'IA_g, ce qui se produit dès lors que les manuscrits sont introduits dans un logiciel de traduction automatique. Malheureusement, les auteurs peuvent constater que le refus de la traduction automatique peut avoir pour effet qu'un éditeur décide de ne pas poursuivre la traduction de l'œuvre. Nous espérons que le secteur de l'édition restera fidèle à ses principes fondamentaux et continuera à valoriser avant tout le travail humain.
- Parallèlement, les traducteurs jouissent aussi du droit d'auteur (et donc de droits moraux), et peuvent refuser de post-éditer un texte traduit à l'avance par une machine. Ici, des accords et une communication claire au sein du triangle auteur-éditeur-traducteur sont essentiels, notamment en ce qui concerne la responsabilité de la traduction résultante et les exigences en matière de transparence. Il est regrettable que les auteurs risquent de constater que tout refus d'utiliser l'IA_g de leur part et/ou de celle du traducteur risque d'inciter l'éditeur de ne pas poursuivre l'exploitation d'une œuvre donnée.
- Les traducteurs ont également le droit de réserver leurs droits en matière de TDM et d'exploitation pour le développement d'IA_g concernant leur version traduite. L'auteur et le traducteur devraient, dans l'idéal, convenir que l'œuvre finale ne peut être utilisée pour le TDM, le *scraping*, la copie et le stockage, ou la programmation de machines et toute production d'informatique générative, et exiger de l'éditeur qu'il déclare cet opt-out dans leur contrat ou sous toute autre forme écrite contraignante. La maison d'édition étrangère doit prendre les mesures nécessaires pour déclarer une réservation de droits lisible par machine, et pour s'assurer que toute personne bénéficiant ultérieurement d'une sous-licence ou d'un accord de distribution soit informée du protocole d'opt-out, le respecte et l'applique.
- L'auteur et le traducteur ne devraient pas craindre de refuser l'utilisation de la traduction automatique, et un tel refus ne devrait pas entraîner de désavantages pour eux.

Accords sur les couvertures et autres créations (éléments graphiques, illustrations, photos)

- Les éditeurs ne devraient pas utiliser des couvertures, graphismes, illustrations ou autres éléments de conception créés exclusivement ou dans une large mesure par la (re)production d'images générative (IA_g, *text to image*) pour compléter une œuvre sans avoir reçu une autorisation écrite de son auteur.
- Ni les concepteurs de couvertures, ni les illustrateurs et artistes plasticiens, ni les auteurs ne devraient avoir à craindre d'être désavantagés, par exemple à travers une rémunération ou des droits d'auteurs diminués, s'ils refusent l'utilisation de l'IA_g.

- De manière générale, les parties prenantes doivent faire preuve de transparence sur leur degré d'utilisation de l'informatique de pointe en tant qu'outil d'assistance, comme l'augmentation de la qualité visuelle pour les œuvres d'art créées par l'humain, ou la génération automatique de texte alternatif (« *ALT text* ») pour produire des formats d'e-books accessibles. Lorsqu'ils produisent du texte alternatif, les éditeurs doivent s'assurer que les images qu'ils chargent dans leur logiciel de description IA ne soient pas exploitées par le développeur du logiciel, par exemple pour entraîner une IA.
- De telles clauses peuvent d'ores et déjà être incluses dans les nouveaux contrats.

Respect du choix des méthodes de travail des auteurs, traducteurs ou des artistes plasticiens.

Les écrivains, les traducteurs, les artistes plasticiens et les interprètes ne devraient pas être requis ou contraints d'utiliser une IA générative ou de travailler à partir de textes générés par IA ou d'images générées par IA contre leur gré.

1.3. Déclarations personnelles des auteurs et exigences des éditeurs en matière de transparence

Les produits issus de l'IA ne sont pas des créations humaines et ne sont donc pas protégés par le droit de la propriété intellectuelle. Par conséquent, les textes ou les images des machines ne possèdent aucun droit légal à une rémunération et ne peuvent pas faire l'objet d'une licence d'exploitation. Les productions d'une machine peuvent être copiées et réutilisées par quiconque.

Il en va de même pour un auteur qui propose un manuscrit ; si celui-ci est produit par IA, il n'a pas le droit d'accorder des licences d'exploitation ou de percevoir une rémunération. Par conséquent, les agents et les éditeurs doivent savoir si l'auteur a utilisé une IA. Ainsi, une chaîne transparente peut être établie et se poursuivre jusqu'au lecteur et au public : il devient important de savoir si une œuvre en voie de publication ou déjà publiée – livre, livre audio ou élément graphique, par exemple – a été entièrement réalisée par des humains ou si elle a été générée (intégralement ou non) par un logiciel lorsque des considérations financières entrent en jeu, par exemple sous la forme d'une rémunération des organismes de gestion collective (pour les copies effectuées par des services de reproduction, le prêt de livres imprimés via le droit de prêt public en bibliothèque, les taxes sur les appareils, mais aussi les droits des artistes-interprètes), lorsque les redevances doivent être partagées, mais aussi en ce qui concerne les obligations en matière de traçabilité, conformément à l'IA Act. Les institutions qui décernent des prix ou des bourses doivent également s'assurer qu'elles continuent à honorer des créations humaines. En outre, les productions des machines ne doivent pas bénéficier d'une TVA réduite, d'un prix fixe pour les livres ou autres subventions destinées aux œuvres humaines et aux biens culturels protégés.

Concernant l'obligation ou non, pour l'auteur ou le traducteur, de fournir une auto-déclaration correspondante : à proprement parler, la clause habituelle des contrats actuellement en vigueur stipulant que l'auteur ou le traducteur est le (seul) créateur de l'œuvre, conformément à la législation nationale sur le droit d'auteur couvrant le niveau de création (= 100 % de son propre travail humain), est suffisante. Toutefois, les éditeurs, en raison de préoccupations légitimes

concernant, par exemple, les atteintes au droit d'auteur, le plagiat, ou leur propre obligation de traçabilité en matière de productions IA, et désireux de garantir leurs droits de sous-licence, peuvent souhaiter que l'auteur signe une déclaration à titre personnel concernant toute utilisation d'IA(g).

Afficher pour toute œuvre publiée si elle a été entièrement ou partiellement produite par IA est nécessaire en termes de responsabilité : au premier chef, ceci permet de déterminer clairement quels auteurs et autres détenteurs de droits doivent être rémunérés. Deuxièmement, cette information permet de savoir quels droits de propriété intellectuelle sont engagés et, troisièmement, de déterminer toute attribution de responsabilité (en cas de contrefaçon, de plagiat, d'atteinte aux droits de la personne et de désinformation). Dans ce contexte, il est également important d'identifier les produits IA d'une manière lisible par un humain. D'une part, pour permettre au lecteur de décider en toute connaissance de cause ce qu'il souhaite acheter ; d'autre part, pour ne pas offrir à des produits issus d'une machine des privilèges dont jouit le livre en tant que bien culturel dans de nombreux États membres, tels que la TVA réduite, les subventions à l'édition ou les bourses et prix. De même, les agrégateurs et les distributeurs de livres insistent pour que les produits IA soient identifiés. Toutefois, on peut s'attendre à ce que l'industrie de l'édition ait un point de vue divergent sur l'obligation d'étiqueter toutes les technologies génératives utilisées.

L'EWC part du principe que seule une transparence complète et fiable associée à une véritable confiance dans les capacités de l'humain permettront à notre secteur de se prémunir et de continuer à inspirer confiance aux lecteurs.

Par conséquent, des pratiques contractuelles et de communication bilatérale doivent être instaurées entre les auteurs et les éditeurs :

- Dans les contrats en vigueur, l'auteur déclare déjà qu'il a créé l'intégralité de l'œuvre en utilisant son pouvoir créatif, conformément aux dispositions nationales relatives aux droits d'auteur et au droit de la propriété intellectuelle, avec un niveau de protection de la création suffisant (par exemple, voir en droit allemand : §2(2) UrhG, « Schöpfungshöhe ») et qu'il détient l'intégralité des droits de cession et des droits d'exploitation. Strictement parlant, cette clause standard exclut déjà les œuvres contenant de l'IA générative.
- À l'avenir, les éditeurs pourraient préférer une auto-déclaration plus spécifique de l'auteur indiquant qu'il n'a pas utilisé de texte, d'image ou de traduction automatique dans l'ouvrage à publier, ou qu'il n'a pas eu recours à des technologies génératives. La question de savoir si une « petite » quantité d'IA devrait être tolérée est controversée et loin d'être réglée ; aux États-Unis, par exemple, on considère qu'une « proportion négligeable d'IA générative » ne peut dépasser 5 % du total d'une œuvre donnée pour qu'elle puisse continuer à être acceptée comme une œuvre humaine protégée par le droit d'auteur. Quant à savoir si cette question peut être résolue en spécifiant des pourcentages et/ou quelles applications d'assistance ou d'analyse IA peuvent être utilisées, il faudra encore un certain temps pour que les associations et l'ensemble du secteur du livre trouvent un consensus. Par principe, toutes les parties prenantes devraient faire preuve de compréhension mutuelle et de transparence.

- L'utilisation de logiciels *d'assistance ou d'analyse* (par exemple : d'indexation automatique de citations, de suggestions automatisées de synonymes ou de reformulation, Photoshop en tant que logiciel d'assistance pure, Word Editor), ou encore le fait de s'inspirer en observant une image IA ou en lisant des « poèmes » générés par IA, par exemple, ne doivent pas faire l'objet d'une déclaration. En revanche, les « œuvres » ou parties d'œuvres produites de manière générative doivent être indiquées sans exception.

Loyauté entre collègues : en utilisant ChatGPT, Bert ou d'autres robots générateurs de texte, d'image ou de voix, vous risquez d'enfreindre les droits d'auteur de vos collègues auteurs et interprètes.

Il est allégué que les éléments de base des grands modèles de langage (« *large language models* » ou LLM) ont été élaborés à partir de collections composées de plus de 4 millions d'œuvres protégées par le droit d'auteur. 194 000 ouvrages contenues dans Pile, Books1, Books2, Books3 ont déjà été identifiées. Leurs sources : des sites pirates proposant des liens *BitTorrent*. Les logiciels d'IAg copient et « mémorisent » des chaînes de mots et des expressions individuelles à partir d'œuvres existantes et produisent souvent des textes proches des originaux, voire des paragraphes entiers qui reproduisent un original mot pour mot. Quiconque utilise effectivement ces applications risque de porter atteinte à la propriété intellectuelle des auteurs dont les œuvres ont été exploitées et sont maintenant reproduites, de manière manifeste ou non, dans ce processus, d'autant plus que ces applications n'assurent pas la traçabilité des œuvres qui ont été exploitées pour générer le texte qu'elles délivrent.

Les logiciels de contrôle du contenu utilisés par les grands distributeurs d'e-books analysent chaque livre avant d'en autoriser le téléchargement, et de plus en plus de logiciels anti-plagiat sont employés, dans le but (a) d'éliminer les produits IA et (b) de détecter les infractions au droit d'auteur.

2. Quelques mots sur les TDM et les techniques d'opt-out : du B2B (« Business to Business ») aux conditions générales d'utilisation de votre logiciel

2.1. Tout le monde utilise des logiciels dans le secteur du livre : votre logiciel est-il en train de vous scraper ?

En 2023 et 2024, presque tous les développeurs de logiciels ont étendu leurs CGU. Cela concerne les logiciels de traitement de texte, d'image et de gestion, les outils de travail collaboratif, le stockage en ligne, les prestataires de services cloud, les services de messagerie, les réseaux sociaux, etc. Désormais, ces nouvelles CGU prévoient des clauses qui leur permettent de copier, de stocker, de reproduire et d'utiliser tout texte, image et autre contenu pour le développement ou l'optimisation d'une IA, y compris générative. Malgré l'interdiction de cette pratique en vertu de la législation européenne sur le droit d'auteur, la protection des données et la vie privée, il est souvent difficile, voire impossible, de faire valoir un opt-out. En cas de refus de consentement, le fabricant restreint l'ensemble des fonctionnalités du logiciel. C'est le cas pour Microsoft, Adobe, Apple, Google, Meta et de nombreuses autres sociétés : « soit vous nous permettez l'accès, soit vous cessez de pouvoir profiter de l'intégralité des services ou des fonctionnalités disponibles. ».

Par conséquent, il incombe à chacun, des auteurs aux agents en passant par les rédacteurs en chef et les éditeurs, de vérifier les logiciels qu'il utilise afin de ne pas ouvrir par inadvertance une faille permettant l'accès aux œuvres, aux données de travail et à d'autres informations commerciales sensibles.

2.2. Développeurs web au service des maisons d'édition, librairies et auteurs : **refusez clairement le TDM, le *scraping* et l'exploitation des œuvres à des fins de développement IA(g)**

- **Opt-out TDM au format texte** : {votre société ou le nom de l'auteur} se réserve expressément le droit exclusif d'utiliser l'ensemble du contenu du site web pour la fouille de textes et de données à des fins commerciales, dans le cadre de {votre législation nationale sur le TDM}. De même, {votre entreprise ou le nom de l'auteur} se réserve expressément tous les droits d'accorder le *scraping* et l'apprentissage machine à des fins telles que (entre autres), le développement d'IA et d'IA génératives. Toute personne souhaitant obtenir une licence d'utilisation de ce matériel est priée de contacter {votre courriel}.

→ Remarque : Il est très probable que les robots d'indexation ne seront pas en mesure de « comprendre » le langage humain, mais seulement de lire des portions de code, cryptogrammes ou autres options lisibles par une machine, par exemple :

- **Un opt-out intégré manuellement au code du fichier robots.txt**:
<https://www.iubenda.com/en/help/137640-block-openai-crawlers>
- **Le protocole de droit de réservation TDM du W3C** intégré aux fichiers EPUB et PDF d'un livre, ainsi qu'à chaque URL où le titre d'un livre est répertorié :
<https://www.w3.org/2022/tdmrep/>
- **Opt-Out via la déclaration de droits ISCC+ (International Standard Content Code Identifier) (non nécessairement contraignante) attachée aux fichiers**. L'EWC recommande au secteur du livre d'utiliser l'ISCC + la déclaration des droits. Ce code peut être appliqué à des œuvres de tous formats (imprimé, numérique, audio, image) et peut être associé à une déclaration qui contient toutes les informations essentielles telles que les réservations de droits (ou même les informations de licence), les données relatives à l'œuvre, les détails sur l'auteur, etc. sous une forme non substituable, puisque la déclaration n'est pas intégrée à l'œuvre, mais constitue un élément externe, sauvegardé dans une base de données d'œuvres enregistrées auprès de l'ISCC. De cette manière, non seulement les opt-outs peuvent être déclarés efficacement, mais les développeurs d'IA sont également en mesure d'utiliser le code ISCC pour documenter l'obtention légale de licences pour des œuvres et pour compiler facilement les listes de titres en vue de prouver leur utilisation en conformité avec les exigences de transparence définies par l'IA Act. Pour les illustrateurs, l'utilisation complémentaire du logiciel de Creators' Credentials est également un bon moyen de démontrer le caractère original (et humain) d'une image.

- **Informations supplémentaires :**

- <https://iscc.codes>
- <https://iscc.io/>
- <https://www.youtube.com/watch?v=S1vK8LMK0f4>
- <https://docs.tdmai.org/>

REMARQUE : l'EWC organisera des webinaires pour ses membres sur l'ISCC et l'opt-out TDM / IA / IAg pour les œuvres du secteur du livre, ainsi que sur les routines d'opt-out robots.txt pour les sites Internet fin 2024/début 2025.

2.3. Éditeurs : **Empêchez le contournement par sous-licence et les tentatives d'ignorer votre opt-out TDM**

Les éditeurs sont tenus de déclarer aux tierces parties, y compris les plateformes, les agrégateurs (dont les bibliothèques de prêt électronique) ou les distributeurs (dont les services d'impression à la demande), que ceux-ci n'ont PAS le(s) droit(s) nécessaires à la concession de sous-licences ou à la reproduction et/ou à l'utilisation de l'œuvre de quelque autre manière que ce soit à des fins de TDM en vertu de l'art. 4 (2019/790 (EU)) pour le développement de technologies d'intelligence artificielle capable de générer du texte, des images ou des voix. Le TDMRep peut être utile à cette fin. Il convient également de sensibiliser le secteur du livre au fait que les libraires, bibliothécaires, critiques journalistiques et attachés de presse, traducteurs, scouts littéraires, agents, évaluateurs (pairs), etc., ne doivent pas saisir de textes protégés dans des systèmes logiciels tels que ChatGPT ou Llama pour produire des résumés, des mots-clés ou d'autres informations, dans la mesure où il s'agit déjà d'un processus de copie et de stockage destiné à développer ou à affiner des IA génératives, ce à quoi l'auteur s'est opposé.

2.4. Éditeurs : **Adaptez vos métadonnées à l'ère du *scraping* et du *crawling*.**

Les éditeurs sont tenus de rendre l'opt-out manifeste et explicite en ayant recours à des métadonnées lisibles par machine et à ONIX, aux CGU sur leur site Internet, à la déclaration de droits l'ISCC+, à TDMRep ou aux mentions légales – bien que cette dernière option ne soit en aucun cas suffisante pour signaler l'opt-out aux robots d'indexation, n'étant utile que lorsqu'elle est lue par un humain. En outre, un protocole de déclaration doit être défini conjointement par les éditeurs, les distributeurs en ligne et les portails de vente de livres, qui sont également tenus d'indiquer à tout robot, crawler ou scraper que l'opt-out s'applique à toutes les œuvres mises en ligne, par exemple par le biais du TDMRep, reconnu par ONIX pour les livres. Par conséquent, une norme harmonisée applicable au flux bilatéral d'informations entre éditeurs et libraires (une « chaîne d'opt-out ») doit être établie, et les interfaces des librairies en ligne être mises à jour.

Deux sources pour approfondir la question :

- <https://www.w3.org/community/reports/tdmrep/CG-FINAL-tdmrep-20240202/>
- <https://iscc.codes>

2.5. Acteurs événementiels : **Soyez prudents lors de la captation et de la diffusion vidéo d'une lecture, d'une table ronde ou d'une conférence.**

Qu'il s'agisse d'une lecture, d'un colloque ou d'une discussion lors d'un salon du livre ou d'une table ronde, les événements sont souvent captés en vidéo et publiés ultérieurement sur des sites web, des chaînes YouTube ou des réseaux sociaux. YouTube a commencé à accorder des sous-licences pour des vidéos à des développeurs d'IA comme Open AI, leur permettant de transcrire des paroles en texte pour alimenter des modèles d'IAg. Cette pratique n'est ni autorisée ni rémunérée et constitue une atteinte notable aux droits d'auteur. Il est important de rechercher/vérifier le plus tôt possible les conditions générales d'une plateforme que l'on envisage d'utiliser afin de savoir si le contenu qu'on souhaite y publier sera réutilisé pour le développement d'un type quelconque d'IA – notamment en ce qui concerne les voix, les images, et les informations présentées lors de l'événement. Il faut en outre conclure des accords bilatéraux avec les artistes et les auteurs pour savoir s'ils acceptent l'enregistrement et la republication de leurs œuvres. De leur côté, les auteurs et les artistes devraient avoir la possibilité de refuser l'utilisation de leur image, de leur voix et du contenu qu'ils ont présenté.

Ressources complémentaires :

- Une clause modèle proposée par la Guilde américaine des auteurs (US Authors' Guild)
- Considérations pratiques de la Société britannique des auteurs (UK Society of Authors)
- Une analyse de l'EWC : <https://europeanwriterscouncil.eu/gai-is-based-on-theft/>
- EWC dictionary on advanced informatics
- Les 10 principes de l'EWC pour réguler l'IAg
- L'entraînement d'IA porte atteinte au droit d'auteur : Étude tandem technologique et juridique