

## Contribution à la mission du CSPLA sur l'exception de fouille de textes et de données en vue de la transposition en droit français de la directive 2019/790

*EARE, 29 octobre 2020,*

[L'Alliance européenne pour l'excellence en recherche](#) (EARE) est une coalition d'entreprises et d'organismes de recherche formée en 2017 qui s'engage pour l'avenir de l'innovation et de la R&D en Europe. Les membres de la coalition incluent [BSA | The Software Alliance](#), [Allied for Startups](#), [Research Libraries UK](#), [SCONUL](#) (Society of College, National and University Libraries), [UK Libraries and Archives Copyright Alliance](#) (LACA) et [UCL Library](#) (University College London). L'Alliance a plaidé en faveur d'une réglementation européenne en matière de propriété intellectuelle permettant une utilisation équitable et efficace de la fouille de textes et de données afin de garantir la compétitivité et la prospérité future de l'Europe.

La présente contribution constitue la réponse d'EARE à la consultation publique lancée par le Conseil Supérieur de la propriété littéraire et artistique sur la transposition des articles 3 et 4 de la directive 2019/790 dite « droit d'auteur dans le marché unique numérique ».

### I. Questions générales

La Commission européenne a reconnu à juste titre que l'intelligence artificielle, l'apprentissage artificiel ('machine learning') et l'analyse des données offrent un vaste potentiel de croissance économique et de compétitivité accrue en Europe. Cependant, leur adoption reste limitée et il est urgent de promouvoir les pratiques de fouille de textes et de données pour que l'Europe devienne un leader mondial de l'intelligence artificielle, comme l'a déclaré le député français Cédric Villani dans son [rapport](#) de 2018 "Donner un sens à l'intelligence artificielle".

La fouille de textes et de données représente en effet une nouvelle vague d'innovation qui permet aux gouvernements européens, aux chercheurs et aux petites entreprises de libérer le pouvoir des données. La fouille de textes et de données génère une intelligence exploitable à partir d'ensembles de données qui étaient autrefois trop volumineux pour être analysés. Elle est à l'origine de progrès révolutionnaires dans l'analyse des données, l'apprentissage et l'intelligence artificiels, qui contribuent à relever certains des défis les plus pressants de notre société.

Les chercheurs, les start-ups, les entreprises de toutes tailles, les gouvernements, les écoles et les particuliers forment par exemple des partenariats uniques pour utiliser la fouille de textes et de données afin de tirer parti de la quantité toujours croissante d'informations produites chaque jour. Dans le seul domaine de la recherche, plus de 28 000 revues scientifiques actives publient collectivement près de deux millions d'articles par an évalués par des pairs. Grâce à la fouille de textes et de données, les chercheurs peuvent exploiter les informations et les idées contenues dans ces documents pour faire des découvertes révolutionnaires dans des domaines essentiels comme la santé, la sécurité publique, les transports et l'écologie.

D'autres pays dans le monde ont facilité, sur le plan juridique, la fouille de textes et de données pour contribuer à stimuler le rythme de l'innovation et de la recherche, conduisant certains innovateurs européens à travailler à l'étranger. Les tribunaux américains, par exemple, ont reconnu à plusieurs reprises que l'utilisation « non expressive » de matériels telle que réalisée par la fouille de textes et de données (c'est-à-dire un acte de reproduction n'ayant pas pour but de permettre à l'homme d'apprécier

ou de comprendre cette reproduction comme expression créative) contribue à promouvoir les objectifs du droit d'auteur, qui consistent à élargir l'accès à l'information et à l'apprentissage sans porter atteinte aux intérêts légitimes des détenteurs de droits d'auteur. Le Japon a également [mis à jour](#) sa "loi sur le droit d'auteur" en 2018 afin de promouvoir le développement de l'intelligence artificielle et de services numériques innovants. Avec cette mise à jour, le Japon a réussi à équilibrer ses règles en matière de droit d'auteur pour soutenir ses ambitions technologiques, permettant ainsi aux chercheurs et aux entreprises privées de mener des activités d'apprentissage artificiel, principalement en levant l'ambiguïté quant à l'utilisation d'œuvres protégées par le droit d'auteur pour la compréhension et l'analyse. De même, Singapour a modifié ses lois sur le droit d'auteur pour permettre la fouille de textes et de données par toutes entités afin de développer la capacité de Singapour à participer et à être compétitif dans le secteur émergent de l'analyse de données. La Chine vise également à devenir le principal centre d'innovation en matière d'intelligence artificielle dans le monde d'ici 2030. Elle est en bonne voie pour y parvenir en investissant massivement dans le développement de ces technologies et en offrant un environnement juridique accueillant aux entreprises d'analyse de données.

L'EARE a donc [demandé](#) aux parties prenantes européennes d'adopter et de garantir la mise en œuvre d'une exception générale pour la fouille de textes et de données, facile à comprendre et à appliquer en Europe, qui permettrait l'émergence d'une industrie de l'intelligence artificielle de haute qualité et florissante en Europe - et plus généralement d'exceptions pour la fouille de textes et de données qui reflètent les réalités de la recherche du 21<sup>e</sup> siècle, où l'utilisation croissante d'un grand nombre de données ('big data') et d'outils d'intelligence artificielle dans la recherche et l'innovation résulte désormais d'une collaboration fluide entre les organisations publiques et privées.

L'EARE a [salué](#) l'adoption des articles 3 et 4 de la directive 2019/790 considérant qu'ils apporteront une clarté bien nécessaire pour toutes les entités qui sont à la pointe de l'écosystème de la recherche et de l'innovation en Europe. L'article 4 sera particulièrement utile pour soutenir le développement et la croissance de l'intelligence artificielle dans l'Union Européenne. Il promouvra les applications d'analyse de données et de 'big data' et ouvrira ainsi de nombreuses nouvelles opportunités commerciales pour les innovateurs en Europe.

## II. Transposition des articles 3 et 4 de la directive 2019/790

Afin que la transposition des articles 3 et 4 de la directive en France réponde aux objectifs des institutions européennes, nous souhaitons souligner les points suivants, qui nous semblent essentiels pour garantir un accès et une utilisation efficaces, sûrs et équitables des données en Europe :

- Les nouvelles exceptions créées par les articles 3 et 4 devraient permettre expressément la reproduction d'œuvres auxquelles l'accès est légal pour faciliter la fouille de textes et de données, à des fins et par des entités commerciales ou non.
- Des limitations strictes et claires aux droits réservés des propriétaires de contenus devraient être mises en place. Le maintien d'un niveau de protection suffisant pour les propriétaires de contenus ne devrait pas exclure le droit, pour les bénéficiaires de ces exceptions (qu'il s'agisse d'entités commerciales ou non commerciales) de :
  - Stocker des copies temporaires d'une œuvre réalisées dans le cadre d'activités de fouilles de données, aussi longtemps que cela est nécessaire à ces fins de fouilles de texte et de données ;
  - Commercialiser les produits de la recherche développés grâce aux exceptions prévues aux articles 3 et 4 de la directive ;

- Mener des activités de fouilles de données sur des contenus légalement accessibles sans risquer d'en être empêché par l'utilisation disproportionnée de mesures techniques de protection ou par la nécessité d'interpréter des clauses contractuelles ambiguës. Par conséquent, lorsque les propriétaires de contenus souhaitent exercer leurs droits et empêcher la fouille de textes et de données sur des œuvres légalement accessibles, nous souhaiterions vivement que la transposition de ces articles exige l'utilisation par ces derniers de standards technologiques reconnus en matière de lecture par une machine ('machine-readable standards'), tels que robot.txt – et ce en plus de dispositions claires dans leurs contrats disponibles en ligne.
- Enfin, exiger de tiers qu'ils détiennent des données issues de la fouille de textes et de données créerait des obstacles inutiles au développement de la recherche européenne. Les universités et les bibliothèques dépensent chaque année des milliards d'euros en abonnements et en acquisitions et on leur fait confiance pour préserver des documents protégés par le droit d'auteur. On devrait également leur faire confiance pour conserver les données dérivées issues de la fouille de textes et de données.

Plus spécifiquement sur la mise en œuvre de l'article 3 :

- Nous pensons que les organisations de recherche devraient pouvoir partager les résultats de leurs recherches en utilisant l'apprentissage artificiel à des fins autres que de vérification. En effet, ces résultats sont utilisés pour beaucoup plus d'applications, y compris le développement et l'amélioration des algorithmes. En outre, ces ensembles de données devraient pouvoir être partagés pour de futurs projets de recherche, car la recherche s'appuie sur le travail que d'autres ont déjà accompli.
- Nous encourageons également l'adoption de dispositions claires sur les modalités de partage des résultats de la recherche en matière de fouille de données, car les chercheurs doivent pouvoir partager les données et les résultats de l'apprentissage artificiel sans craindre de violer la législation sur le droit d'auteur.
- L'EARE soutiendrait l'élaboration d'une "Charte de bonnes pratiques" entre les propriétaires de contenus, les organismes de recherche et les institutions culturelles par le gouvernement français.

Concernant la transposition de l'article 4, nous rejetons fermement l'introduction de licences prévoyant une rémunération financière comme contrepartie pour les propriétaires de contenus d'accepter la fouille de textes et de données telle que prévue par cet article. L'introduction d'une telle mesure irait totalement à l'encontre de l'objectif de la directive. Les licences ne sont pas une solution raisonnable pour la fouille de textes et de données à des fins commerciales. Tout d'abord, la fouille de textes et de données est similaire à la lecture et la lecture n'a jamais nécessité de licence de la part des titulaires de droits d'auteur, car le droit d'auteur ne couvre pas les faits, les idées et les concepts contenus dans les œuvres, et en pratique, très peu de ces licences sont disponibles. Il est important qu'une exception pour la fouille de textes et de données soit rédigée de manière à ne pas porter atteinte à la fouille de contenus librement disponibles.

Deuxièmement, les textes et les données proviennent de milliers de sources différentes, et il est très rare qu'il y ait une identification claire de ce qui est protégé et de qui en est propriétaire. Ainsi, imposer une obligation de licence pour la recherche utilisant la fouille de textes et de données nécessiterait la négociation de centaines de milliers de licences potentielles avec des propriétaires non identifiés autour de droits peu clairs. Cela paralyserait la recherche et créerait la possibilité de contentieux abusifs en

matière de droits d'auteur contre les personnes engagées dans la recherche, et plus grave encore, cela reviendrait à revenir sur d'importants objectifs de la politique du droit d'auteur – c'est à dire le transfert et le partage des connaissances sans nuire aux intérêts légitimes des titulaires de droits d'auteur.

Plus généralement, nous pensons que la grande majorité des titulaires de droits ne s'opposeront pas à l'utilisation de leurs œuvres, autorisée par une exception adoptée en vertu des articles 3 et 4. Plus important encore, toute transposition des articles 3 et 4 devrait reconnaître qu'il existe de multiples façons pour les titulaires de droits et les utilisateurs de tirer profit de matériels protégés par le droit d'auteur qui n'impliquent pas de traditionnelle licence restrictive contre paiement, comme l'utilisation de licences open source permissives dans le domaine du développement de logiciels.

Nous demandons instamment que la transposition des articles 3 et 4 encourage un éventail diversifié de d'utilisations possibles qui répondent aux besoins du plus grand nombre possible de titulaires de droits et d'utilisateurs.

Nous pensons que les points énumérés ci-dessus sont des conditions essentielles pour s'assurer que la capacité à fouiller des données en Europe ne sera pas limitée et que les meilleurs talents en matière d'intelligence artificielle et les capitaux d'investissement ne migreront pas vers des juridictions non européennes ayant des lois plus favorables en matière de droits d'auteur. Une législation claire, conforme aux objectifs fixés par les institutions européennes, favorisera le développement d'un écosystème juste et équilibré où les nouvelles idées sont construites par la communauté de recherche européenne au sein de l'Europe.

Nous avons beaucoup apprécié l'opportunité de contribuer à cette mission sur la transposition de la directive dite « droit d'auteur dans le marché unique numérique » et serions très heureux de participer à tout processus de dialogue et de consultation supplémentaire que votre gouvernement pourrait vouloir créer à l'avenir.