



## CONSEIL SUPERIEUR DE LA PROPRIETE LITTERAIRE ET ARTISTIQUE

### MISSION INTELLIGENCE ARTIFICIELLE ET CULTURE

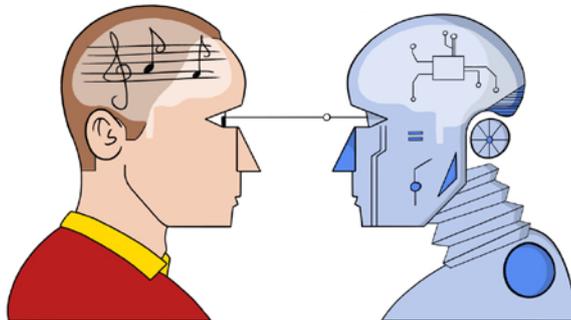
**Rapport final**  
**27 janvier 2020**

**Alexandra BENSAMOUN, professeure à l'Université de Rennes 1**  
**Joëlle FARCHY, professeure à l'Université Paris 1 Panthéon-Sorbonne**

*Personnalités qualifiées au CSPLA*  
*Présidentes de la mission*

**Paul-François SCHIRA, auditeur au Conseil d'Etat**

*Rapporteur de la mission*





## *Intelligence artificielle et Culture*

Le contenu de ce rapport, fruit d'une réflexion menée à partir d'auditions, de contacts informels et de nombreuses lectures, n'engage que ses auteures.

Les parties 1 et 4 ont fait l'objet de développements approfondis dans l'ouvrage *La culture des données*, Farchy, Denis, Presses des mines, à paraître 2020. Pour les parties 2 et 3, le lecteur pourra également se reporter au chapitre « Intelligence artificielle et propriété intellectuelle », in *Droit de l'intelligence artificielle*, dir. A. Bensamoun et G. Loiseau, LGDJ-Lextenso, coll. Les Intégrales, 2019, p. 235.

### **Sommaire**

Résumé .....	5
Introduction générale.....	7
1. Définition de l'IA .....	7
2. IA et culture, quels enjeux ?.....	8
Partie 1. – L'utilisation de l'IA dans le secteur culturel aux différents stades de la chaîne de valeur.....	12
1.1. La place de l'IA dans la science des données.....	12
1.2. La place des données culturelles dans les progrès de l'IA .....	13
1.2.1. Typologie des données culturelles .....	13
1.2.2. Les promesses de l'IA pour la production et l'extraction automatique de données .....	14
1.2.3. Les potentialités de l'apprentissage automatique pour la culture.....	15
1.3. La consommation : dialoguer avec les usagers et leur recommander des contenus..	16
1.4. L'investissement dans la production : connaître le marché pour prendre des décisions adaptées .....	19
1.4.1. La musique : s'informer pour promouvoir les futurs succès.....	19
1.4.2. L'audiovisuel : s'informer pour limiter les risques de la production .....	20
1.4.3. Le marché de l'art : s'informer pour mieux acheter ou investir.....	20
1.5. La création : l'émancipation limitée de la machine.....	21
1.5.1. Accompagner le processus (humain) de la création.....	23
1.5.2. S'émanciper (un peu) de la création humaine .....	23
Partie 2. – Le statut des produits culturels générés par l'IA « créative ».....	26
2.1. La piste du droit d'auteur .....	28
2.1.1. Les protections existantes.....	28
2.1.1.1. La non-protection de l'algorithme.....	28
2.1.1.2. La protection des autres composantes de l'IA .....	29
2.1.1.3. La nécessité d'une protection des « produits » générés par l'IA.....	30
2.1.2. L'application renouvelée du droit d'auteur .....	31
2.1.2.1. Les réalisations de l'IA sont-elles des « créations » ?.....	32
2.1.2.2. Les « œuvres » générées par l'IA sont-elles « originales » ? .....	34
2.1.2.3. Les « œuvres » générées par une IA ont-elles un auteur ?.....	35
2.1.3. La faisabilité de l'application du droit d'auteur aux produits générés par l'IA (questions de titularité).....	36
2.1.3.1. Première hypothèse de titularité : cibler le concepteur de l'IA.....	37
2.1.3.2. Deuxième hypothèse de titularité : cibler l'utilisateur de l'IA.....	39
2.1.3.3. Imposer légalement un régime de titularité.....	40

2.2. Les pistes alternatives au droit d’auteur classique .....	41
2.2.1. La création d’un droit d’auteur spécial.....	42
2.2.2. La création d’un droit d’auteur à la manière d’un « droit voisin » .....	42
2.2.3. La création d’un droit <i>sui generis</i> .....	44
2.3. L’absence de droit privatif.....	45
2.4. Quelle solution ?.....	47
Partie 3. – Le régime des œuvres permettant la production de créations culturelles par l’IA .	50
3.1. Le droit d’auteur s’applique-t-il aux œuvres dont se nourrit l’IA créative ? .....	50
3.2. Le difficile équilibre entre les objectifs d’innovation par l’IA et de protection légitime des auteurs .....	53
3.2.1. Les possibles exceptions au monopole conféré par le droit d’auteur.....	53
3.2.1.1. L’inapplicabilité des exceptions actuellement en vigueur .....	53
3.2.1.2. La nouvelle exception de fouille de données .....	54
3.2.2. Le mécanisme des licences comme solution adaptée ?.....	58
Partie 4. – Améliorer le partage de données pour répondre aux enjeux du développement de l’IA .....	60
4.1. Les enjeux de l’accès et du partage des données pour le secteur culturel.....	60
4.1.1. Les données d’usage : des enjeux de transparence et de maîtrise de la relation- client .....	60
4.1.2 Les métadonnées : des enjeux de qualification des contenus.....	63
4.2. Les voies futures du partage.....	63
4.2.1. Faire des données culturelles des infrastructures essentielles : une proposition séduisante mais inadaptée .....	64
4.2.2. La piste de la portabilité des données d’usage .....	66
4.2.3. La piste des métadonnées d’intérêt général.....	67
Liste des personnes ayant contribué à la mission.....	70
Contributions écrites .....	73

## Rapport *Intelligence artificielle et Culture*

### Résumé

L'intelligence artificielle (IA) a, de manière évidente, investi les secteurs culturels, sur la base d'algorithmes et de larges corpus de données qui les alimentent et offre déjà de multiples applications développées par des centres de recherche, des grandes entreprises ou des start-ups spécialisées. Selon les objectifs poursuivis, les techniques mais aussi les types de données utilisées ne sont pas identiques. Si la question de l'IA est aujourd'hui étroitement corrélée à celle des « données », ce terme générique recouvre, de fait, des réalités très hétérogènes. Pour des raisons historiques différentes, certaines catégories comme les données personnelles, les données publiques ou les données « œuvres » font l'objet d'une qualification juridique bien précise. Il n'en est pas de même d'autres **catégories de données** (métadonnées, données d'usage...) qui correspondent, dans les industries culturelles, à des pratiques professionnelles.

Grâce à ces données, **des applications concrètes de plus en plus nombreuses, apparaissent tout au long de la chaîne de valeur, tant au stade de la création, qu'à celui de la consommation, en passant par celui de la production.** Au niveau de la consommation, l'IA et les algorithmes sont massivement utilisés pour recommander des contenus aux internautes. Si la recommandation personnalisée algorithmique fait débat depuis les premiers travaux sur la « bulle de filtre », avec les risques d'enfermement du consommateur dans ses habitudes, les formes de recommandation sont, en réalité, plurielles. L'exploitation de données, grâce à des algorithmes, redonne également vigueur à l'ambition de prendre les décisions d'investissement adaptées et d'appuyer, voire de remplacer, les habituelles intuitions et expertises humaines par des analyses supposées objectives des déterminants du succès d'une œuvre ou d'un artiste. Outre l'analyse des tendances du marché, l'une des promesses de l'intelligence artificielle est de comparer, sur la base de l'exploitation de données historiques, les contenus ayant connu le succès avec ceux en cours de production afin d'analyser les clés de la réussite, et éventuellement de l'anticiper. Au niveau de la création, parmi les diverses expérimentations mises en œuvre dans les secteurs culturels, toutes n'ont pas le même degré de maturité ; certaines, qui se contentent d'accompagner le processus humain de la création, sont largement présentes dans l'audiovisuel et l'édition ; d'autres, qui tentent de s'en émanciper, se retrouvent plutôt dans la musique et le marché de l'art.

D'un point de vue juridique, l'IA intervient dans la phase de **création artistique**, ce qui suscite des interrogations en termes de droit d'auteur. En effet, alors que le marché de l'art reçoit des créations annoncées comme générées par une IA, se pose la **question de la qualification de ces nouvelles productions**. S'agit-il d'œuvres de l'esprit, à ce titre protégées par le droit d'auteur ? Dans cette hypothèse, qui est l'auteur et le titulaire de droit ? **Une analyse renouvelée des conditions d'accès à la protection (création, originalité, auteur) pourrait permettre de recevoir ces réalisations culturelles au sein du droit d'auteur.**

Mais d'autres solutions sont également proposées (droit spécial, absence de protection privative...). Aussi, il importe d'éprouver le droit positif et d'être prêt à intervenir si un éventuel besoin de régulation se révélait à l'avenir. Dans tous les cas, la démarche devrait s'opérer dans un cadre international, *a minima* européen.

**En outre, l'IA créative fonctionne, dans une démarche d'apprentissage, grâce à l'ingestion d'œuvres, décomposées, analysées, afin d'identifier des traits caractéristiques communs.** Ce processus permet la création d'un modèle d'inférence dont la mise en œuvre aboutit à la génération d'une création algorithmique. Ainsi, la réalisation du portrait Edmond de Bellamy a été permise grâce à l'élaboration d'une base d'entraînement de près de 15 000 portraits classiques du XIV<sup>e</sup> au XX<sup>e</sup> siècle. Doit-on considérer que ces actes amont constituent des actes d'exploitation donnant prise au droit d'auteur ? La mise en place d'une exception « fouille de données » dédiée aux usages de l'IA, à l'article 4 de la directive 2019/790 du 17 avril 2019, semble valider une réponse positive. Mais cette nouvelle limitation, très large, prévoit aussi l'exercice possible d'un *opt-out* par les titulaires de droit, qui permet un retour à la réservation. Dès lors, d'autres solutions doivent être avancées pour accéder à une utilisation facilitée des contenus protégés tout en assurant la protection des ayants droit. Dans ce cadre, des licences générales volontaires pourraient assurer un équilibre.

Enfin, **dans la mesure où la quantité et la qualité des données mobilisables pour alimenter le développement de l'IA devient un facteur de compétitivité, il est nécessaire de s'interroger sur les enjeux spécifiques du partage et de la circulation des données pour les entreprises culturelles.** Pour les données d'usage et les métadonnées, les enjeux de la circulation et du partage de données, sans toujours être directement reliés à des questions de propriété intellectuelle, n'en sont pas moins considérables. Au-delà de l'enjeu de transparence pour une juste répartition des revenus, au profit des ayants droit, ce sont, en réalité, tous les équilibres au sein de la filière qui s'avèrent susceptibles d'être remis en cause par l'accès aux données d'usage ou au contraire par la perte de la maîtrise de la relation-client. Pour les métadonnées, il s'agit essentiellement de qualifier les contenus afin d'accompagner les processus de création artificielle et de renouveler les formes de prévision ou de recommandation, proposées dans les industries culturelles. Une régulation tenant compte des spécificités sectorielles pourrait utilement accompagner l'accès des opérateurs aux grandes masses de données détenues par d'autres.

## Rapport *Intelligence artificielle et Culture*

### Introduction générale

#### 1. Définition de l'IA

L'IA dont il est question aujourd'hui a des origines anciennes puisque l'idée de machines « intelligentes » est ancrée dans la science-fiction, que les réflexions scientifiques remontent aux années 50 et que la plupart des algorithmes actuels ont été inventés dans les années 80. L'expression d'« intelligence artificielle » se trouve en effet formalisée pour la première fois dans la présentation du *Summer Camp* du Dartmouth College en 1956 par John McCarthy. Dans ses travaux, menés avec son collègue John McCarthy, le mathématicien américain Marvin Minsky, dix ans après l'apparition du premier ordinateur, suppose que celui-ci peut être utilisé pour mieux comprendre l'intelligence et définit l'IA comme « la science qui consiste à faire faire aux machines ce que l'homme ferait moyennant une certaine intelligence »<sup>1</sup>.

La poignée d'invités à la conférence de 1956 va dominer, durant les années suivantes, les recherches, qui se multiplient, autour de ce concept nouveau, d'intelligence artificielle. Au cours des dernières décennies, l'IA a donc connu des périodes de recherches plus ou moins intenses, dominées alternativement par deux approches, l'approche symbolique ou cognitiviste et l'approche connexionniste. Le symbolisme renvoie à un raisonnement déductif, basé sur les règles et la logique, tandis que le connexionnisme correspond à un raisonnement inductif, basé sur l'expérience. Le tournant des années 2010 est marqué par le triomphe du *deep learning*, après 60 ans de succès et de déconvenues de l'approche connexionniste. Bien que les algorithmes d'apprentissage automatique puissent utiliser des outils mathématiques très différents, au sein de ce vaste ensemble, les travaux se concentrent en effet sur les réseaux de neurones artificiels et l'apprentissage profond. Le terme neurone renvoie ici à une analogie fonctionnelle avec le comportement des neurones réels en biologie. Les nouvelles architectures de réseaux de neurones sont qualifiées de « profondes », car elles sont composées d'un nombre de couches supérieur aux réseaux de neurones traditionnels proposés dans les années 1960. De cette façon, l'apprentissage profond imite plus efficacement le fonctionnement du cerveau humain.

L'IA, discipline aux multiples facettes, revient donc sous les feux de l'actualité avec le *deep learning* dont le succès correspond à la conjonction de trois facteurs :

- l'amélioration des capacités de calcul des machines. Combiner un grand nombre de neurones formels sous la forme d'un réseau « profond » demeurerait une construction essentiellement théorique tant qu'aucune machine n'était suffisamment puissante pour calculer et simuler le fonctionnement de ces réseaux ;
- les progrès de la recherche sur les réseaux de neurones convolutifs (*Convolutional Neural Networks*, CNN), plus sophistiqués, développés dans les années 1980, qui suscitent une nouvelle vague d'intérêt à partir de 2012, amorcée par quelques succès d'applications pratiques ;

---

<sup>1</sup> CNIL, *Comment permettre à l'homme de garder la main ? Rapport sur les enjeux éthiques des algorithmes et de l'intelligence artificielle*, Paris, déc. 2017.

- l'accroissement de la masse des données, variées et pertinentes, disponibles. Pour les modèles d'apprentissage automatique qui connaissent actuellement les développements les plus importants, la performance de l'IA est, en effet, directement corrélée aux données collectées.

A la promesse d'une IA « forte », capable d'égaliser, comme le souhaitaient les pionniers, l'intelligence humaine, s'opposent des IA faibles, non conscientes et spécialisées dans un domaine bien particulier, qui ne peuvent résoudre que les problèmes pour lesquelles elles ont été conçues et entraînées. Les IA aujourd'hui sur le marché, intégrées notamment aux produits grand public, correspondent en réalité uniquement à cette seconde catégorie. De manière très schématique, il est possible de répartir les principaux systèmes d'IA existants actuellement (IA faibles) en deux grandes catégories issues des courants symboliques et connexionnistes :

- Les systèmes déductifs, à base de règles prédéfinies, des plus simples aux plus sophistiquées (anciennement appelés « systèmes experts »), sont capables de résoudre des problèmes précis et prédéterminés à partir de faits et de règles connus. Ces systèmes d'IA peuvent être qualifiés de déterministes, dans la mesure où la réponse qu'ils apportent à une question peut être précisément déterminée en fonction des données d'entrée et de faits et de règles objectifs intégrés dans le système.
- Les systèmes inductifs, basés sur l'apprentissage automatique ; ce terme de « *machine learning* » renvoie à la capacité des machines d'apprendre sans être explicitement programmées, à partir de données d'entraînement. Il ne s'agit plus de spécifier chaque action du programme, mais de concevoir un programme capable d'adapter son comportement à son environnement. Ces systèmes, qui s'appuient sur des modèles probabilistes, sont donc capables de résoudre des problèmes complexes à l'aide d'algorithmes paramétrés à partir de données d'apprentissage.

Les progrès en matière d'apprentissage automatique ont rendu les systèmes à bases de règles prédéfinies obsolètes dans beaucoup de domaines, les systèmes basés sur l'apprentissage automatique étant généralement beaucoup plus performants pour une même tâche. L'apprentissage automatique, qui est donc la grande voie de développement de l'IA contemporaine, marque le passage d'une logique de programmation à une logique d'auto-configuration. En pratique, les applications métier de l'IA combinent souvent ces deux systèmes et les approches connexionnistes et symboliques sont plus complémentaires que concurrentes.

## 2. IA et culture, quels enjeux ?

La Commission européenne a dévoilé, le 25 avril 2018, sa stratégie dans une Communication *Une intelligence artificielle pour l'Europe*. Le groupe d'experts nommé par la Commission a proposé des actions communes : accroître les investissements, rendre davantage de données disponibles, favoriser les talents et développer une éthique de l'IA pour garantir la confiance. Par ailleurs, à la suite de la remise du rapport de Cédric Villani, la France a décidé de soutenir le développement de l'IA via un plan d'investissement, tandis que la Commission européenne prévoit d'investir près de 2,5 milliards d'euros entre 2021 et 2027. Dans les recommandations adoptées par l'OCDE le 22 mai 2019 sur l'IA ou dans la communication de la Commission européenne du 8 avril 2019, on trouve quelques grands principes de régulation de l'IA, comme l'absence de discrimination, la transparence, l'explicabilité, le respect de la vie privée et de l'autonomie humaine...

Si le rapport Villani met l'accent sur quelques secteurs en se concentrant sur ceux dans lesquels la France peut développer les avantages compétitifs les plus importants, il note que l'IA touche en réalité tous les domaines, y compris ceux où on ne l'attend pas forcément, comme le sport ou la culture. Les enjeux de l'IA dans le secteur de la culture sont en effet nombreux. Le grand public découvre l'intelligence artificielle à travers des expositions comme en témoignent, en 2019, l'exposition interactive au Barbican Centre, à Londres, *AI: More than Human*, qui explore les innovations créatives et scientifiques de l'intelligence artificielle, ou encore, en 2018, l'exposition *Artistes et Robots*, au Grand Palais à Paris.

Les initiatives visant à faire émerger des réflexions et recommandations dans la sphère publique se multiplient également. L'UNESCO s'est notamment intéressée à la question à travers le prisme de la diversité des expressions culturelles<sup>2</sup> et l'OIF a organisé en 2019 une série de rencontres sur les enjeux et les opportunités de l'IA pour le secteur des industries culturelles et créatives<sup>3</sup>. Le sujet de la création artistique et de son statut est également un point d'attention pour de nombreuses autorités et institutions. Ainsi, l'OMPI fait état, fin septembre 2019, de l'organisation d'un « dialogue sur la propriété intellectuelle et l'intelligence artificielle »<sup>4</sup>, qui a notamment abordé les questions juridiques étudiées dans le rapport de la mission (mais le dialogue étant « informel », aucun document écrit n'a pour l'heure été publié à l'exception du programme). La réflexion est aussi en cours aux États-Unis<sup>5</sup>. Le Parlement européen (*Legal Service*) s'interroge également, dans le cadre de l'établissement de son programme de travail à venir, sur la question de la protégeabilité des créations réalisées par une IA<sup>6</sup> et la Commission européenne a évoqué le sujet de la propriété intellectuelle dans ses deux communications sur l'IA<sup>7</sup>. En octobre 2019, le *Global Forum on AI for Humanity*, sous le haut patronage du Président de la République, proposait un *workshop* intitulé « *Rethinking cultural and ethical issues in AI* »<sup>8</sup>. En France, une mission franco-québécoise est en cours sur les aspects de « découvrabilité » liée à l'utilisation de l'IA dans le secteur de la culture<sup>9</sup>.

---

<sup>2</sup> O. Kulesz, *Culture, machines et plateformes : l'intelligence artificielle et son impact sur la diversité des expressions culturelles*, Comité intergouvernemental pour la protection et la promotion de la diversité des expressions culturelles, 12<sup>e</sup> session, déc. 2018 :

[https://en.unesco.org/creativity/sites/creativity/files/12igc\\_inf4\\_fr.pdf/](https://en.unesco.org/creativity/sites/creativity/files/12igc_inf4_fr.pdf/).

<sup>3</sup> <https://www.francophonie.org/node/170/>.

<sup>4</sup> [https://www.wipo.int/meetings/fr/details.jsp?meeting\\_id=51767/](https://www.wipo.int/meetings/fr/details.jsp?meeting_id=51767/).

<sup>5</sup> <https://www.uspto.gov/about-us/events/artificial-intelligence-intellectual-property-policy-considerations/>. – V. la partie « *Copyright: Who is the author of AI-generated content? Are such works copyrightable? What policy implications arise from the use of copyrighted works for the purposes of machine learning?* ». Des contacts ont été pris par la mission. Il est prévu de transmettre le rapport.

<sup>6</sup> Audition au Parlement européen, séminaire *Legal consequences of the technological revolution* : A. Bensamoun, « Droit d'auteur et intelligence artificielle : un nouvel enjeu pour la création », 6 déc. 2018, Bruxelles.

<sup>7</sup> V. en dernier lieu, Communication de la Commission européenne, *Artificial Intelligence for Europe*, COM(2018) 237 final, 25 avr. 2018, pt 3.3. et note 52.

<sup>8</sup> V. la table ronde 3 « *Cultural politics in the era of AI: creativity and discoverability* », avec V. Guèvremont, « *The new measures in favour of discoverability* » ; O. Kulesz, « *Artificial intelligence and the cultural sector: opportunities and challenges* » ; A. Bensamoun, « *The protectability of AI creations: a legal and ethical issue* » ; P.-L. Déziel, « *Use personal data in a context of valuing cultural diversity* ». – <https://gfaih.org/>.

<sup>9</sup> V. Guèvremont *et al.*, *Agir en faveur de la découvrabilité des contenus culturels dans l'environnement numérique : arguments, réflexions et pratiques émergentes*, rapport commandé par le Ministère de la Culture et des Communications du Québec dans le cadre de la *Mission franco-québécoise sur la découvrabilité des contenus culturels francophones en ligne* lancée le 3 avril 2019, à venir.

Prenant acte de cet intérêt croissant pour l'IA et la culture, le CSPLA a confié aux professeures Alexandra Bensamoun et Joëlle Farchy une mission sur les enjeux juridiques et économiques de l'IA dans les secteurs culturels, afin d'identifier l'ensemble des problématiques soulevées (v. lettre de mission du 1<sup>er</sup> juillet 2018<sup>10</sup>).

**Comme prévu dès le lancement de cette mission, ce rapport constitue une simple étape dans la réflexion, qui devra sans aucun doute être poursuivie à moyen terme.** Il aspire principalement à permettre une compréhension de la technique et de ses conséquences juridiques et économiques, et à dresser un état des lieux des solutions au vu de l'état de la technique, elle-même appelée à évoluer.

Si la question de l'IA est aujourd'hui étroitement corrélée à celle des « données », ce terme générique recouvre, de fait, des réalités très hétérogènes. Pour des raisons historiques différentes, certaines catégories comme les données personnelles, les données publiques ou les données « œuvres » font l'objet d'une qualification juridique bien précise. Il n'en est pas de même d'autres catégories de données (métadonnées, données d'usage...) qui correspondent, dans les industries culturelles, à des pratiques professionnelles. Sur la base d'algorithmes et de larges *corpus* de données qui les alimentent, l'IA intervient dans l'ensemble de la chaîne de valeur des industries culturelles. Les enjeux ne se posent cependant pas dans les mêmes termes sur les plans juridiques et économiques.

Sur le plan juridique, c'est au niveau de la création que surgissent les questions de propriété littéraire et artistique, développées dans le présent rapport. Sur le plan économique, ce n'est pas seulement au niveau de la création, mais également au niveau de la consommation, en passant par la production, que l'IA est mise en œuvre. Un des grands défis à relever au cours des prochaines années sera de concilier le respect de grands principes de régulation, dont la propriété intellectuelle n'est qu'un aspect, avec des modèles économiques qui font du traitement et de la valorisation des données le cœur de leur métier et qui imposent une large circulation de ces données.

Afin d'appréhender l'ensemble des problématiques, dans un premier temps, le rapport développe les utilisations actuelles de l'IA dans les secteurs culturels aux différentes étapes de la chaîne de valeur. L'IA offre déjà de multiples applications développées par des centres de recherche, des grandes entreprises ou des start-ups spécialisées. Ayant connu une dématérialisation précoce, les industries culturelles et les médias font figure de laboratoire pour les modèles économiques émergents mobilisant l'IA (Partie 1).

D'un point de vue juridique, les œuvres elles-mêmes constituent des données entrantes nourrissant les algorithmes et l'IA produit de nouvelles « créations » artificielles. Deux questions principales se posent donc au niveau de la création. La première – aussi la plus médiatique sans doute –, en aval, vise à déterminer le statut de la production générée par une IA créative. Est-ce une œuvre de l'esprit, à ce titre protégeable par le droit d'auteur ? Et dans cette hypothèse, qui en est l'auteur ? (Partie 2).

La seconde question juridique, en amont, consiste à identifier le régime applicable aux œuvres qui alimentent l'IA permettant la production de ces créations artificielles (Partie 3). Dans le modèle de création par apprentissage automatique, l'IA est nourrie de données, des plus simples au plus sophistiquées, dont les œuvres de l'esprit font partie. Outre la question du

---

<sup>10</sup> <http://www.culture.gouv.fr/Thematiques/Propriete-litteraire-et-artistique/Conseil-superieur-de-la-propriete-litteraire-et-artistique/Travaux/Missions/Mission-du-CSPLA-sur-l-intelligence-artificielle/>.

statut des nouvelles créations, se pose donc celle de l'appréhension, par le droit d'auteur, de ce phénomène d'analyse, d'utilisation et de décomposition d'œuvres protégées.

Enfin, dans la mesure où la quantité et la qualité des données mobilisables pour alimenter le développement de l'IA devient un facteur de compétitivité, il est nécessaire de s'interroger sur les enjeux spécifiques du partage et de la circulation des données pour les entreprises culturelles et sur les pistes futures d'évolution (Partie 4).

**Partie 1. - L'utilisation de l'IA dans le secteur culturel aux différents stades de la chaîne de valeur**

**Partie 2. – Le statut des produits culturels générés par l'intelligence artificielle « créative »**

**Partie 3. – Le régime des œuvres permettant la production de créations culturelles par l'IA**

**Partie 4. – Améliorer le partage de données pour répondre aux enjeux du développement de l'IA**

## **Partie 1. – L'utilisation de l'IA dans le secteur culturel aux différents stades de la chaîne de valeur**

Si l'apprentissage automatique nécessite, toujours, de grandes quantités de données, réciproquement, l'analyse de données mobilise des disciplines ou champs de recherche d'une grande variété qui ne font pas toujours appel à l'apprentissage automatique (1.1). Selon les objectifs poursuivis, les techniques mais aussi les types de données utilisées ne sont pas identiques (1.2). Outre les données « œuvres », d'autres catégories de données, qui ne correspondent pas à des qualifications juridiques précises, comme les données d'usage ou les métadonnées, sont utilisées par les professionnels et les applications concrètes, tout au long de la chaîne de valeur, apparaissent de plus en plus nombreuses, au niveau de la consommation (1.3), de la production (1.4) ou de la création (1.5).

### ***1.1. La place de l'IA dans la science des données***

Les différentes techniques développées ne poursuivent pas les mêmes objectifs et ne s'adressent pas aux mêmes communautés.

#### *La finalité explicative des techniques quantitatives*

Les économistes font depuis longtemps appel aux techniques quantitatives d'analyse de données. Ces techniques, développées autour du principe de l'inférence causale (explicitement la relation liant  $y$  aux variables  $x$ ), supposent une modélisation *a priori* où les données viennent valider des hypothèses ou des intuitions. Dans cette démarche, l'interprétation de chaque variable explicative est prépondérante.

#### *La finalité performative de l'apprentissage automatique*

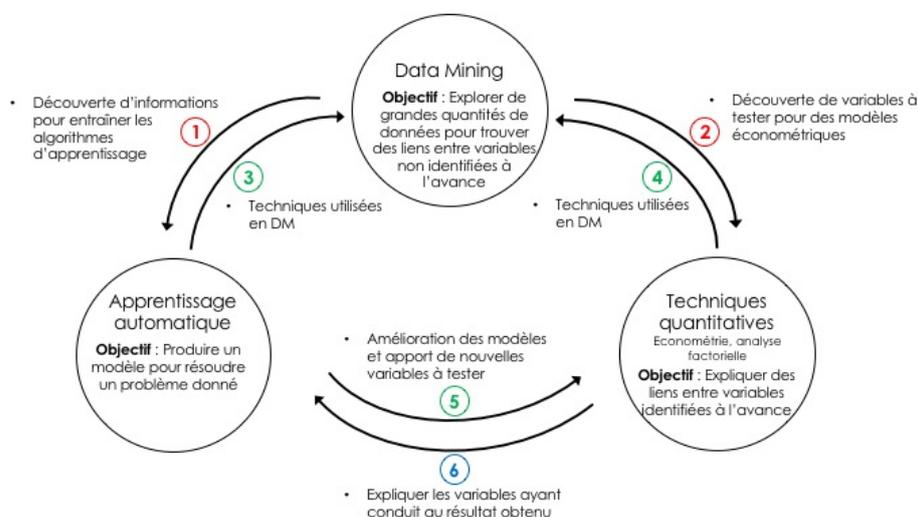
Parallèlement aux outils développés par les économistes, la performance des traitements automatisés, basés sur l'apprentissage automatique, a connu de nombreux progrès ; ceux-ci sont des traitements inductifs (voir *supra*) intrinsèquement conservateurs, qui n'anticipent pas des évolutions mais reproduisent le passé dans le présent ou l'avenir pour produire un résultat, des tendances ou estimations. La rupture avec les techniques quantitatives classiques repose sur le fait que la modélisation est issue des données elles-mêmes. La performance du modèle ne tient pas à son explicabilité des résultats mais à sa capacité à fournir rapidement et efficacement une réponse au problème posé. Plus l'algorithme d'apprentissage automatique est complexe, moins le processus qui a conduit au résultat pourra être exprimé en termes compréhensibles par l'homme.

#### *La finalité exploratoire du data mining*

Le terme de *data mining*, « exploration » ou « fouille » des données, émerge à la fin des années 1970 à l'initiative de la communauté marketing. Peu à peu, le *data mining* est devenu non plus seulement une technique de recherche d'informations en masse, mais également de mise en lumière de structures et d'associations, *a priori* difficiles à détecter ou à imaginer, entre des données. L'approche *data mining* fait appel aux techniques de l'analyse quantitative comme de l'apprentissage automatique. Elle se concentre sur la modélisation exploratoire et la découverte de relations précédemment inconnues entre les données.

Les techniques quantitatives, l'apprentissage machine et le *data mining* qui, historiquement, diffèrent par leurs objectifs, leurs communautés et leurs méthodologies, tendent aujourd'hui à devenir complémentaires et sont utilisées, tour à tour, par les nouveaux « *data scientists* » selon la nature de la question à résoudre. L'apprentissage automatique réalise des prévisions opérationnelles d'autant plus précises que le système est alimenté par de grandes quantités de données. Le *data mining* découvre des corrélations totalement inattendues entre des données, pouvant servir d'entraînement à l'apprentissage automatique, ou être testées par des techniques quantitatives. Ces dernières mettent l'accent sur des liens de causalité au-delà de simples corrélations liées au hasard et apportent des éléments de stratégie sur le long terme aux entreprises comme aux décideurs publics.

Le schéma ci-dessous illustre de façon synthétique les liens entre les diverses approches scientifiques d'analyse des données.



Farchy, Denis, 2020

## 1.2. La place des données culturelles dans les progrès de l'IA

Dans les industries culturelles, outre les données « œuvres », une typologie des données mobilisées par les acteurs conduit à distinguer données d'usage, métadonnées et données hybrides. Grâce aux traitements de l'ensemble de ces données, l'IA améliore diverses tâches qui trouvent des applications nouvelles dans les secteurs culturels.

### 1.2.1. Typologie des données culturelles

Les **données d'usage** résultent de la rencontre d'UN utilisateur et d'UNE œuvre. Les données relatives à l'usage de contenus culturels et ludiques sont devenues un « *input* » déterminant permettant d'analyser les goûts et habitudes des usagers et de leur proposer une offre personnalisée. Les données relatives au temps et à la fréquence d'écoute ou de lecture, aux préférences et achats réalisés, pour n'en citer que certaines, peuvent tomber sous le coup de la qualification de données personnelles dès lors qu'elles permettent d'établir le profil précis d'un individu.

Nous retiendrons comme définition des **métadonnées**, toutes les informations décrivant un contenu ou un créateur précis selon des caractéristiques objectives ou plus ou moins construites. Ces métadonnées, qu'elles soient ou non visibles pour les utilisateurs finaux, sont au cœur de nombreux dispositifs. Elles peuvent être classées, en distinguant celles juridiques (titulaires des droits), descriptives (titre, durée) ou enrichies (genre, analyse sémantique du langage utilisé dans un scénario, rythme d'un morceau de musique...).

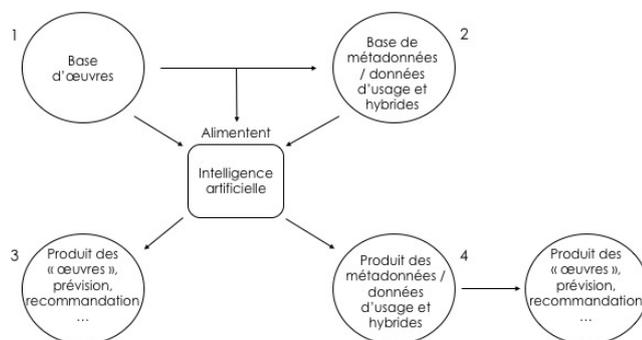
Enfin, à l'intersection des données d'usage et des métadonnées, se construisent des données que nous qualifions d'**hybrides** : informations produites par la comparaison ou l'agrégation de données d'usage et/ou de métadonnées (analyse des échanges réalisés par les internautes sur les réseaux sociaux associant des termes significatifs à un contenu donné).

### 1.2.2. Les promesses de l'IA pour la production et l'extraction automatique de données

Sans toujours les révolutionner, l'IA prolonge et automatise les possibilités de production et d'extraction de données faisant émerger des recherches qui trouvent de plus en plus d'applications pratiques. L'IA mobilise en effet les données de deux manières.

Tout d'abord, en amont, les algorithmes d'apprentissage automatique « piochent » dans des bases d'œuvres qu'ils décomposent ou analysent (1) ou dans des bases de métadonnées préexistantes (2). Selon les situations, l'IA est en effet alimentée par des métadonnées sur les œuvres, ou, dans des cas plus spécifiques, directement par les œuvres elles-mêmes (similarité d'une œuvre musicale par rapport à une autre par exemple) grâce, notamment, aux progrès de l'apprentissage profond. Dans ce second cas, il est nécessaire d'avoir une IA capable de « comprendre » tous les aspects de l'œuvre (pour un film à la fois la vision et le son), ce qui conduit à un coût élevé d'entraînement des algorithmes sur une quantité considérable de données. Par ailleurs, mobiliser des bases de métadonnées enrichies plutôt que les œuvres directement accroît le pouvoir explicatif des résultats proposés par l'IA en termes de prévision ou de recommandation. Selon les objectifs recherchés, le traitement conjoint de métadonnées et de données d'usage alimente diverses applications.

Ensuite, en aval, l'IA produit automatiquement des créations artificielles (3) et de nouvelles métadonnées, fournissant un travail de plus en plus précis de qualification des contenus (4) ou de nouvelles données hybrides, afin d'offrir divers services aux internautes ou aux professionnels. Le schéma ci-après récapitule ces circuits complexes.



### 1.2.3. Les potentialités de l'apprentissage automatique pour la culture

L'une des grandes promesses de l'intelligence artificielle est donc l'extraction et l'analyse de larges bases de données non structurées comme des œuvres et la production de nouvelles métadonnées compréhensibles par les humains.

Parmi les diverses tâches améliorées par l'apprentissage profond, beaucoup trouvent des applications dans les secteurs culturels. On retiendra notamment les exemples suivants.

L'analyse du **langage naturel** (NLP, *Natural Language Processing*), qui se réfère à la fois à une meilleure compréhension du langage naturel et à une meilleure restitution de résultats en langage naturel, a été largement améliorée par l'IA. D'un côté, sur les réseaux sociaux par exemple, l'analyse de données en langage naturel permet de suivre non seulement le volume des discussions autour d'un contenu mais également les émotions exprimées dans ces discussions. De l'autre côté, la génération automatique de textes simples en langage naturel (*Natural Language Generation*) vise à développer des programmes informatiques capables de produire des textes dans un langage compréhensible par l'humain.

Dans le domaine musical, une discipline académique de recherche à part entière a émergé à la fin des années 90, le **MIR** (*Music Information Retrieval*) dont l'objectif est de rassembler une multitude d'informations permettant de qualifier un contenu musical (on parle aussi d'indexation musicale). Des techniques MIR ont été mises au point pour résoudre des problèmes tels que la classification des genres, l'identification des artistes et la reconnaissance *in situ* d'un morceau.

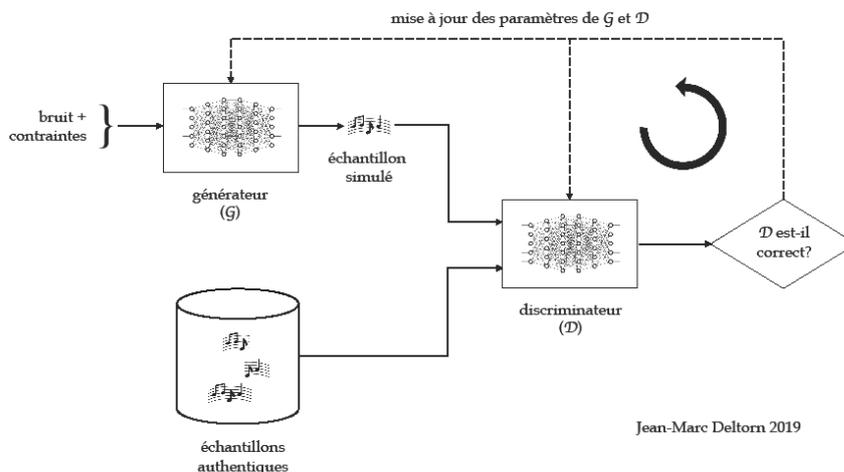
Dans l'audiovisuel, le **catalogage vidéo automatique**, devenu un sujet de recherche important, prend ses sources dans les années 1990 lorsque la quantité d'images disponibles explose. Le catalogage vidéo a pour objectif de structurer les vidéos et d'extraire de l'information sémantique pour assurer que le contenu puisse être retrouvé rapidement. Qualifier et cataloguer efficacement chaque contenu permet un cycle de vie plus long et augmente les possibilités de valorisation. En outre, l'apprentissage profond apporte de nouveaux outils capables de percevoir, en temps réel, le contenu vidéo offert. Il serait par exemple possible lors du suivi d'un événement sportif en direct par un spectateur de choisir de regarder le compétiteur de son choix, lui-même identifié automatiquement par une intelligence artificielle utilisant une technique de reconnaissance faciale et d'étiquetage automatique.

Sur le marché de l'art, les nouvelles techniques de « **vision par ordinateur** » visent à obtenir automatiquement la description d'une image et une recherche efficace. Certains travaux de recherche ont proposé des algorithmes qui combinent la reconnaissance d'objets et des modèles de langage pour générer des descriptions de l'image basées sur le langage naturel. La description du contenu de l'image à l'aide du langage naturel est particulièrement utile pour présenter la création de légendes ou pour effectuer des recherches basées sur le contenu à partir de requêtes textuelles.

Enfin, la génération de contenus artificiels a été renouvelée par les modèles *Generative Adversarial Networks* (**GAN**) permettant de modifier des images et des sons et d'en produire de nouveaux, très réalistes. Les GANs opposent deux réseaux de neurones ; le premier (Générateur – ci-dessous G –), sorte de « faussaire », génère un contenu qui tente d'imiter les

données d'entraînement fournies à l'IA (se rapprocher d'une œuvre de Van Gogh par exemple si ce sont des tableaux de Van Gogh qui ont été fournis à l'entraînement) ; le deuxième (Discriminateur – ci-dessous D –) tente de discriminer les données d'entraînement (une œuvre de Van Gogh) et l'*output* fourni par G (un contenu ressemblant à un tableau de Van Gogh). De manière itérative, l'information est analysée à chaque retour, jusqu'au point optimal où D n'est plus en mesure de distinguer si le contenu nouveau proposé est issu des données d'entraînement ou généré artificiellement. Lorsque toutes les données d'entraînement sont des œuvres produites par des humains, D pourra ainsi faire la différence entre ces œuvres humaines et l'*output* généré artificiellement.

### Séquence d'entraînement d'un réseau adverse génératif



Ces chantiers sont essentiels, comme le montrent les multiples applications possibles de l'IA dans le secteur, qu'il s'agisse, comme nous allons le voir, des processus de création artificielle, de décisions stratégiques d'investissement ou d'orientation des internautes vers des contenus précis.

#### 1.3. La consommation : dialoguer avec les usagers et leur recommander des contenus

Pour le grand public, une des applications d'IA les plus répandues correspond aux agents conversationnels (*chatbots*) ; le système est capable de reconnaître et d'interpréter le texte tapé par l'internaute pour extraire les caractéristiques de sa question, et les fournir comme données d'entrée à un moteur de recherche. Outre ce dialogue direct avec les usagers, l'IA et les algorithmes sont massivement utilisés pour recommander des contenus aux internautes. Dans un univers d'abondance informationnelle où l'intérêt du consommateur est la ressource rare, la recommandation désigne l'ensemble des dispositifs visant à orienter l'internaute vers un contenu ou un ensemble de contenus particuliers. D'un point de vue économique, la recommandation facilite l'appariement de l'offre et de la demande sur des marchés de prototypes où existent de fortes asymétries d'information. En fonction de l'angle adopté, diverses typologies de recommandation peuvent être proposées, qui se recoupent partiellement.

Une première distinction oppose, selon la nature des traitements opérés, les recommandations fondées sur des traitements purement humains (comme la recommandation éditoriale) aux recommandations basées sur des traitements automatisés algorithmiques. Ces derniers peuvent être personnalisés grâce à des données d'usage ou non personnalisés grâce à des analyses sémantiques de similarités de contenus. Tandis que pour l'élaboration d'analyses de marché ou de prédictions, les acteurs ont uniquement besoin des données de consommation agrégées, pour les outils de recommandation personnalisée, il est nécessaire de connaître les données d'usage de chacun des utilisateurs.

On distingue par ailleurs, parmi les traitements automatisés algorithmiques, trois types de recommandation basés sur la nature des données entrantes :

- Le filtrage fondé sur la comparaison entre la description précise des caractéristiques d'un contenu (métadonnées) et les préférences propres à chaque utilisateur (c'est un film d'action, j'aime les films d'action, donc je vais aimer ce film).
- Le filtrage collaboratif ne s'appuie nullement sur les caractéristiques intrinsèques d'un contenu mais sur celles de groupes d'utilisateurs aux goûts proches ; reprenant un principe développé avec succès par la plateforme Amazon, les recommandations sont effectuées en comparant le profil de l'internaute avec ceux d'autres utilisateurs ayant acheté les mêmes contenus : « les internautes qui ont acheté le livre que vous achetez ont également acheté ceux-ci... ».
- Des formes hybrides sont souvent mises en œuvre afin d'améliorer la performance de la recommandation. Le filtrage fondé sur le contenu est en effet peu efficace lorsque les métadonnées manquent de précision. Or la qualification des contenus nécessite souvent une forme d'intervention humaine pour en vérifier la qualité en complément des traitements automatisés. Le filtrage collaboratif est quant à lui inopérant lorsqu'il n'y a que trop peu de données existantes (démarrage à froid de l'algorithme) et peu performant pour les contenus peu populaires pour lesquels des tendances ne peuvent être repérées.

Dans l'**audiovisuel**, les outils de recommandation ont des objectifs différents selon qu'il s'agit d'attirer le spectateur dans une salle de cinéma pour un film précis à un moment donné, ou de fidéliser, de manière récurrente, le téléspectateur sur une chaîne ou sur un service de vidéo par abonnement. La plateforme Netflix s'est notamment fait connaître par la puissance de son modèle de recommandation algorithmique personnalisée. Pourtant, au-delà de l'argument marketing, le dispositif proposé par Netflix est plus hybride qu'il n'y paraît puisqu'il mixe une recommandation algorithmique automatisée basée sur des données d'usage et une intense intervention humaine de qualification des contenus afin de produire des métadonnées pertinentes ; en interne, chaque contenu est regardé d'un bout à l'autre par un ou deux humains et fait l'objet d'une centaine de tags afin de pouvoir proposer une catégorisation susceptible d'être croisée avec les comportements des utilisateurs. Dès lors, la puissance du modèle économique de la plateforme n'est pas liée à la seule force technique de son algorithme, mais à l'intensité et à la qualité des données collectées sur les utilisateurs et sur les contenus, que seul un *leader* du marché, disposant d'une immense base de données, est en capacité de mettre en œuvre.

Les chaînes de télévision, pour développer des systèmes de recommandation pertinents, souffrent d'un double handicap par rapport aux plateformes « *digital natives* » : l'accès aux

données d'usage et le manque de compétences internes pour exercer une activité qui ne constitue pas leur cœur de métier. Des applications et l'écran d'accueil des téléviseurs intelligents, des tablettes et des boîtiers OTT facilitent une meilleure collecte des données, étape préalable aux services de recommandation, mais ces données ne sont pas toujours partagées par les fabricants de matériels ou les opérateurs de télécommunications et nécessitent la mise en place de relations contractuelles complexes. Pour pallier le manque de compétence en interne, des *pure players* se sont positionnés sur le marché de la recommandation audiovisuelle. En France, deux acteurs, Cognik et Spideo proposent des systèmes de recommandation clés en main, destinés aux chaînes de télévision et éditeurs de VOD.

Le marché de la **musique en ligne** constitue une application majeure des outils de recommandation et la connaissance des utilisateurs a été largement enrichie par l'exploitation de données de contexte. De nombreuses études ont en effet démontré que les auditeurs sont des êtres complexes aux multiples facettes, qui souhaitent écouter de la musique correspondant au contexte dans lequel ils se trouvent à un moment précis. Par ailleurs, les recherches MIR (*Music Information Retrieval*) dont l'objectif est de rassembler une multitude d'informations afin d'analyser, de plus en plus finement, un contenu musical ont trouvé dans les services de recommandation un domaine d'application majeur. Au-delà de la recommandation basée sur les métadonnées descriptives, l'analyse du contenu audio tente de modéliser la similitude entre des chansons avec des éléments tels que le tempo, le rythme ou encore la mélodie. Les récentes avancées, en traitement du signal audio, fournissent des informations pouvant être incorporées dans les systèmes de recommandation de manière complémentaire aux méthodes de filtrage collaboratif ; alors que ces derniers favorisent par essence les contenus les plus populaires au détriment de la « longue traîne », la recommandation fondée sur le signal audio permet de favoriser la découverte musicale.

Les plateformes de *streaming*, *leaders* du marché, n'ont d'autres choix que de proposer des systèmes de recommandation de plus en plus performants pour se différencier de la concurrence. Si la plupart ont d'abord testé des formes de recommandation exclusivement éditorialisée (Deezer) ou fondée sur les usages passés de chacun (Spotify), elles se sont finalement orientées vers une approche hybride. Ainsi, elles analysent aussi bien le comportement des utilisateurs que les données de blogs ou réseaux sociaux par le traitement du langage naturel, les données de contexte ou encore les métadonnées intrinsèques (signal audio) pour construire leurs systèmes de recommandation.

Sur le **marché de l'art**, la visibilité sur le web est devenue un enjeu crucial aussi bien pour les maisons de ventes aux enchères que pour les galeries, opérateurs historiques, ou pour de nouveaux intermédiaires qui se positionnent en mettant en relation les collectionneurs avec les galeries ou directement avec les artistes. Les outils de recommandation tendent à devenir de plus en plus hybrides afin d'associer l'efficacité de préconisations personnalisées sur la base des usages de l'internaute et une qualification fine des contenus proposés. Les stratégies adoptées sont différentes selon les entreprises. Les plateformes dominantes disposent à la fois de la base utilisateurs et des moyens financiers nécessaires pour internaliser le développement d'outils de recommandation. D'autres vont piocher dans des bases de données existantes, s'appuient sur des développeurs extérieurs ou sur le rachat de *pure players*.

Si la recommandation personnalisée algorithmique fait débat depuis les premiers travaux sur la « bulle de filtre » avec les risques évoqués d'enfermement du consommateur dans ses habitudes, les formes de recommandation sont, en réalité, plurielles. Face aux tendances

homogénéisatrices, les algorithmes de recommandation peuvent également être utilisés afin de faire découvrir une diversité inégalée de contenus.

#### ***1.4. L'investissement dans la production : connaître le marché pour prendre des décisions adaptées***

L'économie de la culture étant une économie de prototype, où l'incertitude sur la demande est particulièrement élevée, identifier les outils d'aide à la décision permettant de limiter la prise de risque a toujours été au cœur de l'attention des professionnels. Un petit nombre de productions attire en effet l'essentiel de la demande selon le modèle bien documenté de l'économie des superstars. L'exploitation de données, grâce à des algorithmes, redonne vigueur à l'ambition de prendre les décisions d'investissement adaptées et d'appuyer, voire de remplacer, les habituelles intuitions et expertises humaines par des analyses supposées objectives des déterminants du succès d'une œuvre ou d'un artiste. Outre l'analyse des tendances du marché, l'une des promesses de l'intelligence artificielle est de comparer, sur la base de l'exploitation de données historiques, les contenus ayant connu le succès avec ceux en cours de production afin d'analyser les clés de la réussite, et éventuellement de l'anticiper.

Qu'il s'agisse d'analyser les tendances du marché ou de prévoir le futur, les outils d'aide à la décision des professionnels se font selon des mécanismes spécifiques à chaque secteur. Dans l'industrie musicale, fortement remodelée par le numérique, les professionnels disposent de nombreux outils de tendance en temps réel et le succès dépend largement des plateformes de *streaming*. Dans le cinéma, où les données de tendance sont moins accessibles, les travaux académiques sur la prévision sont nombreux, apportant de multiples indications aux professionnels afin de développer leurs outils. Enfin, sur le marché de l'art, l'objectif essentiel est de permettre une évaluation objective du rendement et du prix des œuvres comme appui aux décisions d'achat ou d'investissement.

##### **1.4.1. La musique : s'informer pour promouvoir les futurs succès**

Analyser des morceaux musicaux sur la base de leurs caractéristiques objectives, à l'aide d'algorithmes d'apprentissage automatique, est devenu un champ de recherche scientifique à part entière avec le développement des systèmes MIR (voir *supra*). Cependant, malgré des progrès, il n'existe pas d'application commerciale assez mature pour prévoir avec certitude un probable « tube » sur la seule base de l'analyse de ses caractéristiques propres et ceux qui s'y sont essayé n'ont guère abouti. D'autres chercheurs tentent de prévoir le succès d'un morceau en s'appuyant non plus sur les seules métadonnées intrinsèques, mais en intégrant des premières données d'usage afin d'anticiper un éventuel succès, de l'accompagner notamment sur les réseaux sociaux et le faire durer.

Pour les majors, labels et artistes, la restitution de données d'usage en temps réel constitue donc une information précieuse afin d'éclairer leurs prises de décisions. Certains proposent, moyennant un abonnement, des tableaux de bord et reports pour visualiser les tendances ou l'évolution d'artistes sur différentes plateformes ; Soundcharts est un des leaders de ce marché. D'autres constituent des systèmes purement internes ou accessibles uniquement aux entreprises musicales et artistes distribués sur une plateforme (Spotify for Artist, Deezer Backstage, etc.).

### **1.4.2. L'audiovisuel : s'informer pour limiter les risques de la production**

Lorsque le film est encore dans une phase où relativement peu d'argent a été investi, les producteurs sont tentés d'évaluer le potentiel commercial d'un nouveau scénario en prenant comme référence la performance au *box-office* de films jugés proches. Sur la base des qualifications réalisées, sont comparés des scénarios de films ayant connu le succès avec ceux en attente de production. Plusieurs *start-ups* (Vault, ScriptBook, StoryFit) se sont positionnées sur ce marché. Le bilan de l'analyse de scénario, activité en vogue, apparaît cependant mitigé. Les progrès incontestables dans l'analyse du langage naturel ont conduit à la multiplication des offres commerciales dont l'objectif est d'analyser (semi-) automatiquement un scénario sur la seule base des métadonnées ; la qualité des résultats fournis n'a cependant pas donné lieu à des évaluations solides. De plus, de nombreux professionnels voient avec réticence l'essor de ces nouveaux outils, craignant le formatage de leur travail. A court terme, il semble donc plus prometteur d'utiliser les modélisations de données, non pas pour analyser la qualité du scénario lui-même, mais pour déterminer le public potentiel, sur le plan à la fois quantitatif et qualitatif, d'un scénario spécifique.

Par-delà la comparaison d'un film à ceux qui lui sont similaires, les studios, distributeurs et exploitants ont recours à des outils de prévision intégrant les données d'usage. En dehors des salles, pour lesquelles les données d'audience sont disponibles et permettent l'exploitation des potentialités de l'IA à des fins de marketing, pour le reste de l'audiovisuel, les outils, issus de l'ère analogique, ne sont pas toujours adaptés aux possibilités du numérique. Contrairement à l'industrie musicale, les professionnels disposent d'un éventail limité de données d'usage sur lesquelles s'appuyer ; celles-ci sont soit agrégées et peu qualitatives, soit non partagées et peu accessibles par l'ensemble des professionnels.

Comme dans la musique cependant, les opérateurs qui disposent de la relation client à l'aval de la chaîne de valeur sont en capacité de traiter des quantités de données plus importantes pour proposer des prévisions performantes. Sur le marché de la SVOD, sur lequel Netflix s'est imposé, les données d'usage sont mal connues. Les mesures de sociétés spécialisées, issues de l'ère analogique, peinent à s'adapter à des offres délinéarisées, personnalisées et multi-supports. Quant à la plateforme elle-même, Netflix, après avoir choisi de ne proposer aucun outil accessible aux professionnels, a évolué et communique désormais, aux seuls ayants droit, des données générales sur les audiences des contenus diffusés.

Enfin, en raison de la convergence technologique de la télévision et d'Internet, les fournisseurs de téléviseurs intelligents et les opérateurs de télécommunications qui proposent des box et des services OTT (*Over The Top*) ont des capacités nouvelles de collecter des données d'usage en temps réel. Cependant, en France, seule SFR (groupe Atlice) propose le service *SFR Analytics Live* qui restitue les données d'audience en temps réel moyennant un abonnement annuel des chaînes de télévision. D'autres opérateurs, comme Orange, restituent des données d'usage agrégées selon des accords, contractuels et limités, signés avec les chaînes de télévision.

### **1.4.3. Le marché de l'art : s'informer pour mieux acheter ou investir**

L'essor du marché de l'art et l'arrivée de nouveaux collectionneurs ont accéléré la demande d'information transparente et organisée, y compris au niveau international, provoquant la prolifération des classements et autres indicateurs. Au cours des dix dernières années, plusieurs plateformes, tirant parti tant des avancées scientifiques que du matériau dont

certaines bénéficient naturellement du fait de leur activité d'agrégateur de données, ont créé leurs propres indicateurs de marché. La prolifération des indicateurs ne rend pas aisée la compréhension de ces informations par les acheteurs et vendeurs d'œuvres d'art. Afin d'éclairer ceux-ci dans leurs choix, des opérateurs se sont donc positionnés sur le marché de l'estimation des prix. D'autres tentent, grâce à des algorithmes, de suggérer dans quelles œuvres il est opportun d'investir. Sur un marché de l'art particulièrement spéculatif, plus encore qu'ailleurs, les frontières entre informations générales, estimation d'une œuvre particulière et prévision se révèlent poreuses d'autant que le manque de transparence des méthodologies de classement alimente des formes de spéculation, d'inflation des prix et de prédictions auto-réalisatrices, les œuvres classées les plus chères étant supposées être celles des artistes les plus cotés, ce qui augmente mécaniquement la cote des artistes.

Ainsi, les tentatives de parfaitement modéliser la réussite d'un livre, d'un scénario ou d'une peinture grâce aux algorithmes, n'ont, à ce jour, pas permis de découvrir les clés de ce secret si convoité. L'analyse de données ne fournit en effet que des probabilités de succès et les algorithmes sont fondamentalement conservateurs. Parce qu'ils apprennent en analysant ce qui a fonctionné dans le passé, ils sont incapables de tenir compte des changements qui se produiront dans l'avenir. De plus, la mise en place de modèles de prévision a un coût, notamment pour collecter et formater des données non structurées, qui ne peut se concevoir dans l'économie de tous les contenus. Enfin, les tentatives d'automatisation du succès portent en germe une forme de contradiction interne car des nouvelles productions basées largement sur la comparaison des résultats des productions passées pour en copier les critères du succès peuvent conduire à des situations absurdes puisque des contenus similaires éternellement proposés aux consommateurs ou des ouvrages peuplés de héros et d'idées de son choix finiront inexorablement par le lasser.

Au-delà de ce rêve, largement inatteignable, d'une modélisation globale, les progrès de l'apprentissage automatique dans le domaine de la reconnaissance visuelle, de l'analyse du langage naturel ou des variables audio, ainsi que les possibilités ouvertes pour le recueil et le traitement des données d'usage ont cependant bouleversé les pratiques professionnelles. Des algorithmes sont déjà largement utilisés pour repérer les tendances du marché, faciliter la prise de décision d'un éditeur, d'un producteur, d'un collectionneur ou encore préciser la stratégie commerciale d'un distributeur afin de cibler, d'élargir le public potentiel ou d'optimiser la présence d'une œuvre ou d'un artiste selon les supports de diffusion. Indicateurs complémentaires de l'expérience professionnelle, les algorithmes aident à mieux positionner un projet ou à anticiper les réactions du public. Le marché de l'art et celui de la musique privilégient l'efficacité de la prévision alors que dans l'audiovisuel, l'explication des combinaisons de variables du succès reste un enjeu prépondérant.

### ***1.5. La création : l'émancipation limitée de la machine***

Les machines sont depuis longtemps au cœur de la création artistique. Certains outils visent non seulement à accompagner mais aussi à encourager l'imaginaire propre à chaque créateur en passant, grâce à l'IA, par une étape de manipulation et d'appropriation de l'univers d'autres artistes ou encore d'intégration, grâce aux données d'usage, des réactions de son public. La plupart des grands acteurs du numérique ont intégré l'importance de la culture comme champ d'application de l'IA et développé des programmes de recherche dédiés : IBM avec son IA « maison » Watson, Microsoft avec les expériences du *Next Rembrandt* et autres outils de montage, ou encore Google et son équipe de recherche Google Brain qui propose des outils aux musiciens (Magenta) ou artistes du marché de l'art (DeepDream). En 2018, le

leader du *streaming* musical, Spotify a inauguré, à Paris, le *Creator Technology Research Lab* afin de développer des outils pour aider les artistes dans leur processus créatif.

Pour appréhender avec exactitude les aspects juridiques de cette création (Partie 2), il est essentiel de bien comprendre le **processus de gestation d'un objet par une IA créative** qui est technique. D'abord, des données de départ exogènes, sous forme de base d'entraînement (ou *training set*), nourrissent l'IA. Dans le cas d'une IA créative, ces sources (ou *inputs*) seront notamment des œuvres de l'esprit, lesquelles constituent une forme d'inspiration externe. Un effort de sélection est parfois fait, avec un degré de d'intervention plus ou moins appuyé, mais il peut aussi s'agir d'une agrégation automatique par un robot. Ce set d'entraînement prend la forme d'une base de données. Il est injecté dans un processus algorithmique (exprimé sous forme logicielle) qui va en identifier les corrélations, les correspondances, les similarités en termes de structures. Cette décomposition par l'algorithme permet aussi de représenter les données initiales d'une manière compréhensible par la machine.

En fonction de l'objectif identifié, une fonction objective d'utilité doit être définie : il s'agit ici par exemple d'indiquer quelles caractéristiques doivent être mises en exergue. Cette fonction représente la mesure d'efficacité de la solution produite par l'algorithme, *i.e.* une mesure de distance entre les attentes et le résultat du processus algorithmique. La fonction de l'apprentissage sera alors de diminuer la distance, de minimiser l'erreur entre le produit de l'algorithme et les résultats escomptés. Le résultat de ce travail de minimisation est de déterminer les paramètres d'un modèle d'inférence<sup>11</sup>, lequel est constitué de l'ensemble des valeurs qui caractérisent le réseau de neurones dans ce type de modèle, à savoir les valeurs caractérisant les liens entre les neurones. L'organisation de ces valeurs peut prendre la forme d'une base de données, mais elle est gérée de manière automatique.

Ce modèle d'inférence vient cristalliser des éléments fondamentaux de la volonté du donneur d'ordre : choix de la base d'entraînement (le modèle d'inférence dépend de la matrice initiale), choix du modèle et de la fonction d'utilité. C'est cet objet, ce modèle stabilisé (une fois l'apprentissage réalisé), qui va donner lieu à une production en accomplissant sa fonction. Réceptacle de l'apprentissage, il est donc le point de départ de la création. Le modèle d'inférence peut ensuite être utilisé, y compris par un tiers qui n'a en aucun cas participé à sa conception. L'utilisateur n'a alors plus qu'à « cliquer » pour générer la production d'une création<sup>12</sup>. Le modèle d'inférence peut aussi, en aval, être modulé, soit en pénalisant les éléments de sortie (ou *outputs*) considérés comme mauvais et en valorisant ceux au contraire évalués comme bons<sup>13</sup>, soit en réinjectant des produits créés dans la base d'entraînement<sup>14</sup>.

Parmi les diverses expérimentations mises en œuvre dans les secteurs culturels, toutes n'ont pas le même degré de maturité ; certaines, qui se contentent d'accompagner le processus humain de la création, sont largement présentes dans l'audiovisuel et l'édition ; d'autres, qui tentent de s'en émanciper, se retrouvent plutôt dans la musique et le marché de l'art. Dans de nombreux cas en effet, il ne s'agit pas de se substituer au processus humain de création, mais

---

<sup>11</sup> V. J. Deltorn, « Quelle(s) protection(s) pour les modèles d'inférence ? », Cahiers Droit, Sciences & Technologies, 7/2017, p. 127, spéc. p. 133 : les modèles d'inférence sont « constitués par un ensemble de paramètres déterminés lors de la phase d'apprentissage et organisés en une structure de données. »

<sup>12</sup> C'était le modèle proposé par Jukedeck (IA musicale), qui mettait à la disposition (contre rémunération) un modèle d'inférence pour créer des jingles musicaux après des sélections mineures comme le style ou les caractéristiques émotionnelles.

<sup>13</sup> Ce travail de sélection et d'assemblage a notamment été réalisé par François Pachet pour la production de *Daddy's car*, titre à la manière des Beatles, produit grâce à une IA.

<sup>14</sup> C'est ainsi que David Cope crée de la musique classique.

de faciliter le travail et l'inspiration du créateur en réduisant les contraintes. Plus ambigus sont les travaux visant à imiter, à créer « à la manière » d'autres artistes, voire à faire émerger des œuvres complètement nouvelles qui limitent fortement l'intervention humaine.

### **1.5.1. Accompagner le processus (humain) de la création**

Dans l'**audiovisuel**, la production d'un film ou d'une série étant complexe et coûteuse, les outils d'intelligence artificielle visent ainsi à simplifier l'ensemble des étapes et à rationaliser les prises de décision lors de la pré-production, du tournage ou de la post-production (simulation de foules, production d'effets spéciaux ou animations). L'IA est également utilisée pour la réalisation de productions audiovisuelles spécifiques comme les spots publicitaires, les bandes-annonces ou les clips pour les réseaux sociaux. Pour d'autres types de productions, les professionnels se montrent plus réservés. En 2016, une campagne *Kickstarter* pour un film d'horreur intitulé *Impossible Things* a permis de financer la production d'un long métrage dont le scénario a été co-écrit avec une machine. Il faut cependant relativiser la part de la machine dans cette création, puisque l'écriture proprement dite a été réalisée par un humain. Au contraire, la même année, *Sunspring*, court-métrage créé par le réalisateur Oscar Sharp et le chercheur en intelligence artificielle Ross Goodwin pour le *48-Hour Film Challenge* du festival du film Sci-Fi de Londres, est rédigé entièrement par la machine. L'expérience n'a cependant pas donné les résultats attendus ; après avoir implémenté une dizaine de scénarios de science-fiction dans une machine d'intelligence artificielle surnommée *Benjamin*, celle-ci a bien proposé un scénario complet mais tout à fait incohérent, avec des dialogues et des scènes invraisemblables et sans aucune logique narrative.

Les progrès en analyse du langage « naturel » ont trouvé diverses applications dans l'**édition**, conduisant à l'émergence d'outils d'automatisation de certaines tâches répétitives ou à la génération de textes simples notamment dans la presse. Au-delà, produire de la littérature en imitant les compétences de l'humain est un des objectifs de l'IA depuis ses débuts. Le générateur de poésie WASP (*Wishful Automatic Spanish Poet*), le programme SPAR (*Small Poem Automatic Rhymers*), le narrateur de récits d'événements réels *Raconteur* ou encore le générateur narratif *PropperWryter*, utilisé dans la comédie musicale *Beyond The Fence* sont quelques exemples d'application possibles. Des chercheurs s'emploient également à permettre d'écrire un ouvrage « à la manière de », en reprenant l'univers d'un auteur célèbre. Le finaliste du concours littéraire Nikkei Hoshi Shinichi, au Japon, a présenté un court roman produit en collaboration avec une IA. *Just This Once* (1993), écrit dans le style de Jacqueline Susann par IA en collaboration avec son programmeur, Scott French, a suscité des réflexions sur le partage de la part de créativité entre la machine et les autres acteurs impliqués. A l'heure actuelle, la génération automatique de contenus textuels en langage naturel est effective essentiellement pour la création d'articles ou communiqués de presse simples et informatifs. Composer des œuvres littéraires capables de tenir le lecteur en haleine pendant des centaines de pages, grâce à un langage habilement manié et une orchestration d'ensemble des événements, n'est pas encore à la portée de l'intelligence artificielle.

### **1.5.2. S'émanciper (un peu) de la création humaine**

Des expériences en matière de créativité associée à la machine ont été menées de longue date dans les arts graphiques ; dès 1973, Harold Cohen développe le programme évolutif AARON, qui combine une IA logicielle avec un équipement de peinture robotique pour permettre à la machine de produire automatiquement des peintures dans un style donné. Loin du robot-

artiste primitif initié par Harold Cohen, les systèmes qui ont recours à l'apprentissage profond et aux réseaux de neurones suscitent désormais l'attention dans leurs applications aux images artistiques. Un projet a été initié par l'agence de publicité J. Walter Thompson Amsterdam pour le compte de la banque néerlandaise ING avec l'appui technologique de Microsoft, *The Next Rembrandt*. Dévoilée à Amsterdam le 5 avril 2016, cette création d'intelligence artificielle « à la manière de » a été réalisée à partir de la base de données extraites de l'entière collection (346 portraits) de l'œuvre de Rembrandt.

D'autres programmes cherchent à faire émerger des créations inédites. Google a ainsi développé un programme d'intelligence artificielle, DeepDream, qui s'inscrit dans la lignée des algorithmes neuronaux CNN (*Convolutional Neural Networks*). Il a conduit à la production de créations nouvelles, exposées en galerie à San Francisco en 2016 et vendues jusqu'à 8 000 \$. D'autres types d'algorithmes neuronaux, les **GANs**, ont été expérimentés pour produire de nouvelles œuvres d'art. Le collectif artistique **Obvious**, fondé par trois jeunes français, a ainsi intégré les données de près de 15 000 portraits classiques du XIV<sup>e</sup> au XX<sup>e</sup> siècle à un algorithme et obtenu des créations inédites représentant une famille imaginaire, les **Belamy**. Dans cette série, un des tableaux, intitulé *Edmond de Belamy* a été vendu plus de 430 000 dollars en octobre 2018 par la Maison de vente aux enchères Christie's à New-York<sup>15</sup>. En réalité, la nouveauté ne réside pas tant dans la prouesse technique ou artistique, ni même dans la vente (d'autres portraits de la famille Belamy avaient auparavant été vendus à des prix plus modestes) que dans la capacité du marché de l'art à communiquer sur la première peinture issue d'une intelligence artificielle vendue aux enchères, par des commissaires-priseurs. Le prix élevé témoigne donc surtout de la fascination du marché pour cette nouvelle technique de création. Pourtant, les membres du collectif reconnaissent eux-mêmes que leur expérience n'a rien d'original. La technique des GANs avait été développée par Ian Goodfellow et son équipe antérieurement ; de plus, le code source de l'algorithme est repris de celui de Robbie Barrat disponible en *open source*, ce dernier s'en étant d'ailleurs ému par la suite compte tenu des finalités commerciales du projet.

Les chercheurs en IA considèrent les fondements de ce travail comme déjà éprouvés. Cependant, l'effervescence médiatique provoquée par cette vente a mis en avant des recherches comme celles de l'artiste Mario Klingemann, un des pionniers de l'utilisation des réseaux neuronaux et des GANs qui a développé un « neurographe » pour créer de nouvelles formes artistiques. Le 6 mars 2019, c'est une œuvre évolutive en temps réel proposée par cet artiste qui a été vendue à Londres aux enchères par Sotheby's au prix de 51 000 livres, devenant ainsi la seconde création issue de l'IA, vendue aux enchères.

Comme pour les arts graphiques, le processus de **création musicale** a été confronté précocement à la machine. Dès les années 80, le compositeur et scientifique américain David Cope fait réaliser à son outil EMI (*Experiments in Music Intelligence*) des centaines de productions à la manière de Bach, Mozart ou Rachmaninov. Aujourd'hui, la plupart des systèmes fonctionnent sur la base de réseaux d'apprentissage profond et repèrent, parmi de grandes quantités de données, des schémas rythmiques ou mélodiques afin de produire de nouveaux morceaux. Si les différents outils sont capables de générer une composition musicale « finie », tous ne laissent pas la même place à l'humain dans le processus de création. Certains outils ne nécessitent aucune connaissance préalable (Amper, Jukedeck, etc.), tandis que d'autres reposent sur une certaine connaissance du codage informatique et de l'écriture musicale (IBM, Google).

---

<sup>15</sup> Pour voir le portrait : <https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx>.

Certaines applications commerciales fournissent des solutions « **clés en main** » grâce auxquelles le néophyte, ou le « pressé », obtient une **composition prête à l'emploi**, comme pour des jingles, maquettes, musiques à l'image de reportage... Elles permettent de créer, sur une plateforme simple en libre-service, une composition correspondant à des paramètres compréhensibles et définis par l'utilisateur traduisant par exemple une émotion (joyeux, inquiétant, exotique) ou un genre (rock, funk...). Les entreprises concernées attirent des clients avec des compositions qui sans être toujours de haute qualité ont une valeur commerciale, en facilitant l'accès à de la musique à bas coût laissant craindre des effets de substitution par rapport aux compositions protégées par le droit d'auteur. Pour autant, s'il est fort probable que l'IA occupera dans les années à venir une part du marché (musiques d'ascenseur, « au kilomètre »...), elle aura aussi pour conséquence de renforcer la créativité des auteurs sur des segments plus qualitatifs de musiques que l'IA reste incapable de produire.

D'autres outils développés par des grands groupes ou laboratoires de recherche **accompagnent les utilisateurs « aguerris »** (musiciens, compositeurs, programmeurs) en leur permettant d'intervenir sur un grand nombre de paramètres et donc d'obtenir une forme de complexification / personnalisation de l'œuvre. L'IA poursuit ce processus de décomposition d'un morceau en isolant chaque caractéristique bien plus précisément. Ainsi, les chercheurs de l'IRCAM développent des outils et algorithmes personnalisés pour chaque artiste qui choisit, outre le jeu de données d'entraînement, le réglage des paramètres de l'algorithme. L'informaticien, qui intervient dans le choix des paramètres et possède également des connaissances musicales, a un rôle important dans le processus de création. Flow Machines, un projet d'intelligence artificielle créé en 2012 par une équipe autour de François Pachet, alors directeur du *Sony Computer Science Laboratories Paris*, se présente comme un outil d'optimisation de la composition musicale. Les développeurs de Flow Machines ont ainsi dévoilé en 2016 deux morceaux « à la manière de » : *DeepBach* façon Bach et *Daddy's Car* façon Beatles<sup>16</sup>. La finalité n'est cependant pas seulement d'imiter un style, mais bien de créer une nouvelle œuvre qui sera le résultat d'un travail homme-machine.

Dans ces hypothèses, l'IA est un **outil** – certes très perfectionné – entre les mains de l'auteur. Le principe doit alors rester celui d'une **prééminence de l'être humain**. Les qualifications juridiques ne sont donc pas appelées à être modifiées pour ce qu'on peut appeler les créations *assistées* par une IA.

Le raisonnement, d'un point de vue juridique, devrait changer si la réalisation est cette fois *générée* par une IA sans pouvoir identifier un auteur, au sens classique du terme.

---

<sup>16</sup> Pour écouter : [https://www.youtube.com/watch?v=LSHZ\\_b05W7o](https://www.youtube.com/watch?v=LSHZ_b05W7o).

## Partie 2. – Le statut des produits culturels générés par l’IA « créative »

L’intérêt des juristes pour le croisement entre l’informatique et le droit d’auteur n’est pas nouveau<sup>17</sup>, dans une discipline fortement marquée par l’influence de la technique.

L’hypothèse ici retenue est celle d’une création *générée* par une IA, pour laquelle la désignation classique d’un auteur-créateur personne physique est rendue complexe par le processus créatif. On peut alors parler d’une « **IA créative** », comme visant une IA qui permet la production d’une réalisation culturelle qui *ressemble* à une œuvre. Cette IA créative (l’outil donc) est elle-même un objet juridique **hybride** (sans doute aussi une œuvre complexe, comme l’est la création multimédia). La qualification de cet élément pluriel ne sera abordé que rapidement (et partiellement), car il se situe en amont de la création produite et dont on cherche le statut.

En aval, d’évidence, la **frontière** entre la création *assistée* et la création *générée* par une IA est difficile à tracer<sup>18</sup>, d’autant que l’emploi d’une machine pour créer est sans conséquence sur la qualification, au nom du principe de l’indifférence du genre<sup>19</sup> et aussi de celui, plus général, de neutralité technologique. L’expérience malheureuse des errements législatifs au sujet de la protection de la photographie doit servir de leçon. Qu’un créateur utilise un pinceau, un burin, un appareil photo ou un instrument technologique de pointe ne modifie pas le raisonnement. Aussi, dès lors que l’IA reste un outil entre les mains d’un créateur-auteur, il faut convenir d’une sorte de **principe de prééminence de l’humain**, qui efface la machine et permet l’application classique des règles du droit d’auteur. Ainsi, dès lors qu’une personne humaine intervient de manière créative dans la création, l’outil utilisé importe peu. Ce n’est que lorsque l’intervention humaine s’éloigne de l’approche classique du fait d’une génération autonome par une IA que le raisonnement proposé pourra trouver à s’appliquer.

La prudence est d’autant plus importante que de nombreux acteurs sont susceptibles d’intervenir aux différentes étapes de la chaîne de création algorithmique, de l’amont, par la sélection de la base d’entraînement, à l’aval, dans la phase de post-production ou de curation, en passant par les étapes intermédiaires de sélection du mode de représentation du signal ou encore le choix du modèle, comme le montre le schéma ci-après.

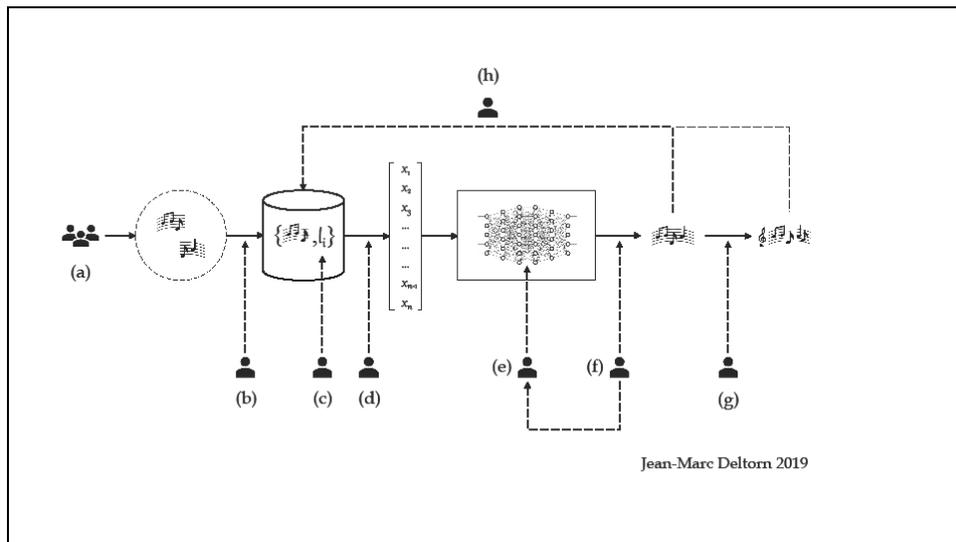
---

<sup>17</sup> V. déjà *L’informatique et le droit d’auteur*, Actes du 57<sup>e</sup> Congrès de l’ALAI, Éd. Yvon Blais Inc., Canada, 1990.

<sup>18</sup> Y. Gaubiac, « Objet du droit d’auteur. Œuvres protégées. Œuvres créées avec un ordinateur (CPI, art. L. 112-2) », J.-Cl. Propriété littéraire et artistique, fasc. 1164.

<sup>19</sup> CPI, art. L. 112-1.

## Schéma de la production d'une création artistique par une IA



*Légende (interventions humaines) :*

(a) La création des œuvres premières

(b) La sélection des œuvres (en vue de former une base d'entraînement)

(c) L'assignation éventuelle de labels<sup>20</sup> (li) à chacune des entités sélectionnées (dans le cas d'un apprentissage supervisé)

(d) La sélection d'un mode de représentation du signal est opérée (les données d'entraînement sont transcrites en un format aisément manipulable par la machine, ici un vecteur de paramètres "x1" ... "xn")

(e) Le choix du type du modèle (ici un réseau de neurones), les détails de son architecture et ses paramètres sont choisis, ainsi que les modalités de la phase d'entraînement (fonction de coût, etc.)

(f) Lors de l'entraînement, les entités produites peuvent être observées, sélectionnées et servir à modifier soit le modèle, soit les caractéristiques de l'apprentissage

(g) Une phase de « post-production » peut permettre de sélectionner et modifier les entités issues du modèle génératif.

(h) Certaines des entités produites peuvent éventuellement être réintroduites dans la base d'entraînement (pour orienter un apprentissage ultérieur. David Cope a par exemple procédé de la sorte, voir *supra*).

Certes, la totalité de ces acteurs n'intervient pas systématiquement dans le processus de création par apprentissage automatique. Mais ce diagramme présente l'intérêt majeur de révéler **le nombre et la variété des intervenants humains** potentiels dans la chaîne de création. A ces programmeurs, entraîneurs, curateurs, on pourrait encore ajouter les investisseurs, les donneurs d'ordre, les utilisateurs ou détenteurs d'une licence d'exploitation... L'émancipation de la création humaine reste donc relative. Pour autant, ces acteurs de la création – ou certains d'entre eux, à tout le moins – peuvent-ils être considérés comme des auteurs, au sens du droit d'auteur ?

Plusieurs questions se posent donc, celle du statut de cette production et de sa protection, celle de l'identification d'un auteur. En droit d'auteur, ces interrogations se rejoignent souvent dans la mesure où on mélange la définition de l'objet du droit d'auteur (l'œuvre), son caractère protégeable (l'originalité), son sujet (l'auteur) et même parfois son titulaire. Il faut dire que la structure circulaire du droit d'auteur français, dans le lequel les notions-clés se renvoient les unes aux autres, y invite.

<sup>20</sup> Les labels sont des mots-clés, des métadonnées, qui permettent d'instruire le modèle dans le contexte de l'apprentissage supervisé.

*A priori*, la conception française personnaliste semble exclure de la protection la production d'un contenu culturel généré par une IA, au point d'ailleurs que certains auteurs de doctrine rejettent son application à cet endroit (2.3). Cependant, l'appropriation de ces produits par la voie du droit d'auteur a été suggérée par le Parlement européen, qui propose de définir des « critères de “création intellectuelle propre” applicables aux œuvres protégeables par droit d'auteur créées par des ordinateurs ou des robots »<sup>21</sup>. Ce qui justifie qu'on examine dans un premier temps la piste d'une protection par le droit d'auteur (2.1). Pour autant, l'analyse n'est pas évidente et d'autres voies d'appropriation pourraient émerger (2.2). Cette diversité d'options implique de s'interroger sur la solution à privilégier (2.4).

## **2.1. La piste du droit d'auteur**

Si l'IA peut faire l'objet de protections, en tant qu'outil, ces protections amont ne déteignent pas sur les productions avalées (2.1.1). D'où la nécessité d'envisager la protection du droit d'auteur, pour peu que les conditions d'accès à la protection puissent être adaptées à la création artificielle, dans une lecture renouvelée (2.1.2), et qu'un titulaire initial puisse être identifié (2.1.3).

### **2.1.1. Les protections existantes**

Plus l'intelligence artificielle se perfectionne, plus les investissements auxquels elle donne lieu la mettent à la portée d'un nombre important d'utilisateurs potentiels, et plus il importe de séparer le régime de l'intelligence artificielle, d'une part, du régime des produits qu'elle est susceptible de générer, d'autre part.

Or les protections actuellement assurées ne concernent guère que la matrice, c'est-à-dire l'outil ou l'innovation que constitue l'intelligence artificielle elle-même – et non les produits qu'elle génère pour la consommation de tous. Certaines composantes sont protégées par un droit de propriété intellectuelle (2.1.1.2) pendant que d'autres en sont exclues (2.1.1.1).

#### **2.1.1.1. La non-protection de l'algorithme**

L'algorithme en lui-même, comme « description d'une suite finie et non ambiguë d'étapes (ou d'instructions) permettant d'obtenir un résultat à partir d'éléments fournis en entrée »<sup>22</sup>, en langage compréhensible par la machine, n'est pas protégé en tant que tel par le droit d'auteur. Il est assimilé à une **méthode** ou une **simple idée**, « par essence et par destination de libre parcours » (Desbois). En ce sens, l'article 9.2 de l'Accord sur les aspects des droits de propriété intellectuelle qui touchent au commerce (ADPIC) précise que « La protection du droit d'auteur s'étendra aux expressions et non aux idées, procédures, méthodes de fonctionnement ou concepts mathématiques en tant que tels ». La règle est reprise à l'article 1.2 de la directive 2009/24/CE<sup>23</sup>, en vertu duquel « Les idées et principes qui sont à la base de quelque élément que ce soit d'un programme d'ordinateur, y compris ceux qui sont à la base de ses interfaces, ne sont pas protégés par le droit d'auteur en vertu de la présente

---

<sup>21</sup> Rapport du Parlement européen contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique, 27 janv. 2017, 2015/2103(INL), p. 32.

<sup>22</sup> CNIL, *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*, déc. 2017, p. 15.

<sup>23</sup> Directive 2009/24/CE du Parlement européen et du Conseil du 23 avril 2009 concernant la protection juridique des programmes d'ordinateur (version codifiée).

directive »<sup>24</sup>. La jurisprudence française<sup>25</sup> comme européenne<sup>26</sup> a en outre rappelé l'exclusion des **fonctionnalités** de la protection du droit commun ; or, la fonctionnalité d'un logiciel est exprimée par un algorithme, qui en est l'énoncé logique.

De la même façon, l'algorithme, comme simple méthode mathématique, n'est pas, en tant que tel, brevetable<sup>27</sup>.

### 2.1.1.2. La protection des autres composantes de l'IA

L'article 10.1 de l'accord ADPIC stipule que « Les **programmes d'ordinateur**, qu'ils soient exprimés en code source ou en code objet, seront protégés en tant qu'œuvres littéraires en vertu de la Convention de Berne ». La formalisation de l'algorithme en expression codée d'un programme est ainsi réservée par un droit spécial au sein du droit d'auteur, notamment issu de la directive 2009/24<sup>28</sup> et déjà introduit en France par la loi du 11 juillet 1985<sup>29</sup>. L'intelligence artificielle pourra ainsi bénéficier de cette protection, mais sans que celle-ci puisse s'étendre à la méthode elle-même (l'algorithme) ou aux idées sous-jacentes (du fait de l'existence d'une exception d'ingénierie inverse, ouverte à l'utilisateur légitime, en vue d'accéder aux fonctionnalités du programme<sup>30</sup>).

Certaines intelligences artificielles seront même susceptibles d'être brevetées en tant qu'invention nouvelle impliquant une activité inventive et susceptible d'application industrielle » (art. L. 611-10 du CPI). Certes, les programmes d'ordinateur sont exclus de la brevetabilité. Mais les rapports d'examen préliminaire international récemment rendus en vertu du Traité de coopération en matière de brevets ont laissé entendre que plusieurs applications développées par la société Deep Mind, rachetée par Google, seraient susceptibles de faire l'objet d'une protection par le droit des brevets<sup>31</sup> – ils sont actuellement en phase d'étude à l'Office européen des brevets (OEB). De manière générale, les algorithmes d'apprentissage automatique peuvent constituer des solutions techniques innovantes,

---

<sup>24</sup> *Adde* considérant 11, excluant « les idées et principes qui sont à la base de la logique, des algorithmes et des langages de programmation ».

<sup>25</sup> Cass. 1<sup>re</sup> civ., 13 déc. 2005, not. JCP 2006, I, 103, obs. C. Caron ; JCP E 2006, 1895, § 1, obs. M. Vivant, N. Mallet-Poujol et J.-M. Bruguière ; P. Belloir, « L'exclusion de la protection des fonctionnalités d'un logiciel par le droit d'auteur », RLDI 2006/14, n° 396 ; Cass. 1<sup>re</sup> civ., 14 nov. 2013, n° 12-20.687, RIDA janv. 2014, p. 493 et p. 399, obs. P. Sirinelli ; LEPI 2014, 174, obs. S. Chatry ; Propr. intell. 2014, p. 56, obs. A. Lucas.

<sup>26</sup> CJUE, 22 déc. 2010, *BSA*, aff. C-393/09, v. pts 42 à 44, RIDA janv. 2011, p. 417 et 213, obs. P. Sirinelli ; Comm. comm. électr. 2011, comm. 42, note C. Caron ; Propr. intell. 2011, p. 205, obs. V.-L. Benabou ; D. 2011, p. 2364, obs. C. Le Stanc ; LEPI 2011, 35, obs. A. Lucas. – CJUE, gr. ch., 2 mai 2012, *SAS Institute*, aff. C-406/10, v. spéc. pt 32, RIDA juill. 2012, p. 341 et p. 181, obs. P. Sirinelli ; Comm. comm. électr. 2012, comm. 105, note C. Caron ; D. 2012, p. 2836, obs. P. Sirinelli ; A & M 2012/6, p. 547, note B. Michaux ; Expertises 2012, p. 265 ; RLDI 2012/83, n° 2768, obs. C. Castets-Renard ; Europe 2012, comm. 284, obs. L. Idot ; Propr. industr. 2012, comm. 61, note N. Bouche ; Propr. intell. 2012, p. 423, obs. V.-L. Benabou ; RTD com. 2012, p. 536, obs. F. Pollaud-Dulian.

<sup>27</sup> CPI, art. L. 611-10, 2, et CBE, art. 52 (2).

<sup>28</sup> Précit. ; version codifiée de la Directive 91/250/CEE, du 14 mai 1991, concernant la protection juridique des programmes d'ordinateur.

<sup>29</sup> Loi n° 85-660 du 3 juillet 1985 relative aux droits d'auteur et aux droits des artistes-interprètes, des producteurs de phonogrammes et de vidéogrammes et des entreprises de communication audiovisuelle.

<sup>30</sup> Ce droit spécial autorise explicitement l'utilisateur légitime du logiciel à observer, étudier ou tester le fonctionnement ou la sécurité du logiciel « afin de déterminer les idées et principes qui sont à la base » de ce logiciel (art. L. 122-6-1, III du CPI, reprenant l'article 5.3 de la directive précitée).

<sup>31</sup> V. par exemple l'application W02018048934 – « *Generating Audio Using Neural Networks* ». <http://ipkitten.blogspot.com/2019/01/full-speed-ahead-for-deepminds-ai.html>.

comprises dans une catégorie plus large et en pleine expansion, que l'OEB nomme les « **inventions mises en œuvre par ordinateur** » (IMOO)<sup>32</sup>.

De même, il n'est pas contesté que les **bases de données** utilisées par l'intelligence artificielle sont susceptibles de recevoir une protection spécifique. La directive 96/9/CE du 11 mars 1996<sup>33</sup> pose en effet un double régime de protection : par le droit d'auteur pour le contenant de la base (ses principes d'organisation, son architecture, notamment)<sup>34</sup> ; par le droit *sui generis* pour son contenu<sup>35</sup>. De fait, l'IA utilisée dans un contexte créatif mobilisera deux types de bases de données<sup>36</sup> : la base d'entraînement (*training set*) composée de données collectées et assemblées et l'encodage du modèle d'inférence lui-même, en tant que collection de valeurs, qui permet le paramétrage de l'IA.

### 2.1.1.3. La nécessité d'une protection des « produits » générés par l'IA

Toutes les protections évoquées ci-dessus – droit d'auteur, droit *sui generis*, droit des brevets – ne disent rien du régime juridique applicable aux produits qui sont générés par la machine intelligente. On pourrait être tenté, en invoquant le modèle des œuvres plurales, d'octroyer la qualité d'auteurs du produit généré par l'IA aux différents créateurs amont. Mais cette voie est fragile car les différents éléments ne se retrouvent pas dans la création finale<sup>37</sup> (alors que c'est le cas par exemple pour l'œuvre complexe qu'est le jeu vidéo). Aussi, ces protections amont **ne suffisent pas** à protéger et sécuriser les investissements en matière d'intelligence artificielle<sup>38</sup>, qui peuvent naturellement être guidés par le gain retiré de l'utilisation de la machine.

En outre, l'absence de protection des produits générés par l'intelligence artificielle laisserait les **interprètes** de ces créations de l'IA sans droits, puisque la protection que confère le régime des droits voisins implique, pour ces acteurs, l'existence d'un droit d'auteur sur l'œuvre interprétée (art. L. 211-1 du CPI).

---

<sup>32</sup> OEB, *Directives relatives à l'examen pratiqué*, , nov. 2018, G-II, 3.6 : « L'expression "invention mise en œuvre par ordinateur" couvre les revendications qui font intervenir des ordinateurs, des réseaux informatiques ou d'autres dispositifs programmables dans lesquels au moins une caractéristique de l'invention revendiquée est réalisée à l'aide d'un programme d'ordinateur. » - V. les exemples donnés par J.-M. Deltorn, « La brevetabilité des applications de l'intelligence artificielle et de l'apprentissage automatique : la pratique de l'Office européen des brevets », Propri. industr. n° 3, 2019, dossier 4.

<sup>33</sup> Directive 96/9/CE du Parlement européen et du Conseil, du 11 mars 1996, concernant la protection juridique des bases de données.

<sup>34</sup> Dir. 96/9/CE, art. 3.1 : « Conformément à la présente directive, les bases de données qui, par le choix ou la disposition des matières, constituent une création intellectuelle propre à leur auteur sont protégées comme telle par le droit d'auteur. Aucun autre critère ne s'applique pour déterminer si elles peuvent bénéficier de cette protection. »

<sup>35</sup> Dir. 96/9/CE, art. 7.1 : « Les États membres prévoient pour le fabricant d'une base de données le droit d'interdire l'extraction et/ou la réutilisation de la totalité ou d'une partie substantielle, évaluée de façon qualitative ou quantitative, du contenu de celle-ci, lorsque l'obtention, la vérification ou la présentation de ce contenu attestent un investissement substantiel du point de vue qualitatif ou quantitatif. »

<sup>36</sup> V. J.-M. Deltorn, « *Disentangling deep learning and copyrights* », AMI - tijdschrift voor auteurs-, media- en informatierecht 2018/5, p. 172.

<sup>37</sup> V. en ce sens la position du groupe français de l'AIPPI, « Droit d'auteur sur les œuvres générées artificiellement », prés. E. Treppoz, avr. 2019. – Pour une analyse internationale, v. le *Summary Report – 2019 – Study Question: Copyright in artificially generated works*.

<sup>38</sup> En ce sens, sur le site de l'OMPI, A. Guadamuz, « L'intelligence artificielle et le droit d'auteur », [https://www.wipo.int/wipo\\_magazine/fr/2017/05/article\\_0003.html](https://www.wipo.int/wipo_magazine/fr/2017/05/article_0003.html)

Enfin, cette absence de protection des réalisations de l'IA serait sans doute à l'origine de **contournements**. En effet, des protections détournées pourraient alors, en pratique, se développer, car il n'est nullement requis de révéler les processus créatifs. Ainsi, une personne physique pourrait bien taire la génération d'une création grâce à un outil d'IA et bénéficier de la présomption d'autorat<sup>39</sup> en accolant son nom à la création. De même, une entreprise, obtenant la qualification d'œuvre collective (lui permettant d'accéder à la titularité initiale des droits)<sup>40</sup> ou agissant en contrefaçon au bénéfice de la présomption jurisprudentielle de titularité<sup>41</sup>, pourrait s'attribuer les droits en gardant le silence sur la manière exacte dont les réalisations en cause ont été produites. D'autant que la frontière entre créations assistées par une IA et créations générées par une IA est poreuse : elle relève plus d'un *continuum* que d'une séparation nette.

Pour ces raisons, il importe de mettre à l'épreuve **le droit d'auteur qui pourrait, par une lecture renouvelée<sup>42</sup>, recevoir en son sein les créations générées par une IA.**

### 2.1.2. L'application renouvelée du droit d'auteur

Le droit d'auteur français est humaniste : à travers le droit d'auteur, c'est aussi la personne physique de l'auteur qui est protégée. Cette conception **personnaliste** a dessiné un champ extrêmement souple de la protection, d'où le mérite esthétique ou la nouveauté de l'œuvre protégée sont absents mais où la personne de l'auteur est centrale.

L'article L. 111-1 du Code de propriété intellectuelle prévoit ainsi que « l'auteur d'une œuvre de l'esprit jouit sur cette œuvre, *du seul fait de sa création*, d'un droit de propriété incorporelle exclusif et opposable à tous » et l'article L. 112-1 impose un principe général de non-discrimination entre les œuvres<sup>43</sup>. La législation française ne fixe cependant aucune définition positive de l'œuvre de l'esprit – à l'inverse des dispositions relatives au droit des brevets, des marques, des dessins et modèles. En cela, elle ne s'écarte pas significativement des définitions très prudentes données dans la Convention de Berne (« Les termes "œuvres littéraires et artistiques" comprennent toutes les productions du domaine littéraire, scientifique

---

<sup>39</sup> CPI, art. L. 113-1.

<sup>40</sup> CPI, art. L. 113-2, al. 3 (définition) et L. 113-5 (régime).

<sup>41</sup> Cass. 1<sup>re</sup> civ., 24 mars 1993, *Aréo* (2<sup>e</sup> arrêt), *RTD com.* 1995, p. 418, obs. A. Françon ; *JCP G* 1993, II, 22085, note F. Greffe ; *RIDA* oct. 1993, p. 191, obs. A. Kéréver : « [en] l'absence de toute revendication de la part de la ou les personnes physiques ayant réalisé les clichés, ces actes de possession étaient de nature à faire présumer à l'égard des tiers contrefacteurs que [la société] était titulaire sur ces œuvres, quelle que fût leur qualification, du droit de propriété incorporelle de l'auteur ». – Formulant la règle dans un attendu de principe : Cass. 1<sup>re</sup> civ., 9 janv. 1996, *RIDA* juill. 1996 ; *D.* 1996, somm. p. 285, obs. J.-J. Burst ; *JCP* 1996, II, 2643, note X. Daverat.

<sup>42</sup> V. not. J. Larrieu, « Le robot et le droit d'auteur », in *Mélanges A. Lucas*, LexisNexis, 2014, p. 465 ; B. Michaux, « Singularité technologique, singularité humaine et droit d'auteur », in *Droit, normes et libertés dans le cybermonde, Mélanges Y. Pouillet*, Larcier, 2018, p. 401 ; P.-Y. Gautier, « De la propriété des créations issues de l'intelligence artificielle », *JCP G* 2018, 913 ; J.-M. Deltorn, « Droit d'auteur et créations des algorithmes d'apprentissage », *Propr. intell.* 2016, n° 58, p. 4 ; F. Macrez et J.-M. Deltorn, « *Authorship in the age of machine learning and artificial intelligence* », [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3261329](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3261329) ; A. Cruquenaire, A. Delforge, J.-B. Hubin, M. Knockaert, B. Michaux, T. Tombal, « Droit d'auteur et œuvres générées par machine », in *L'intelligence artificielle et le droit*, dir. H. Jacquemin, A. de Steel, Bruxelles, Larcier, 2017, p. 189 ; A. Bensamoun, « Les créations générées par une intelligence artificielle ou la tentation du droit d'auteur », in A. Bensamoun et F. Labarthe (dir.), *Culture et numérique, Rencontre franco-québécoise*, Mare et Martin, coll. PUS, 2020, à paraître ; A. Bensamoun, « Intelligence artificielle et propriété intellectuelle », in *Droit de l'intelligence artificielle*, dir. A. Bensamoun et G. Loiseau, LGDJ-Lextenso, coll. Les Intégrales, 2019, p. 235.

<sup>43</sup> « Les dispositions du présent code protègent les droits des auteurs sur toutes les œuvres de l'esprit, quels qu'en soient le genre, la forme d'expression, le mérite ou la destination. »

et artistique, quel qu'en soit le mode ou la forme d'expression »). Nul ne doute cependant que l'œuvre de l'esprit vise, en droit français, une **création de forme originale**. Quoique le critère essentiel de la protection soit, en France comme en droit international, l'originalité de la création, il n'en existe pas davantage une définition statutaire.

Ce caractère elliptique de la notion a permis sa **souplesse** – et partant une plasticité du droit –, laquelle explique le succès du droit d'auteur et justifie qu'on envisage d'y avoir recours pour assurer la protection de créations qui, jusque-là, n'étaient pas susceptibles d'exister. Mais si le droit d'auteur se révèle une catégorie juridique malléable, l'idée d'intégrer les produits générés par l'intelligence artificielle dans son régime pourrait toutefois se heurter à plusieurs obstacles, pris dans des critères dont il faudrait alors renouveler la lecture : la création (2.1.2.1), l'originalité (2.1.2.2) et l'auteur (2.1.2.3).

### 2.1.2.1. Les réalisations de l'IA sont-elles des « créations » ?

La première question qui se pose est celle de savoir s'il est possible, quand on parle de produits générés par l'intelligence artificielle, de les qualifier de « créations » : y a-t-il réellement un processus créatif à l'œuvre lorsque l'intelligence artificielle fonctionne ? Pour répondre, il faut s'intéresser à la notion de création – peut-être d'ailleurs jusqu'à la réinventer ?<sup>44</sup>

La création, définie comme une transformation du réel, de l'existant, renvoie traditionnellement à la notion de **conscience**. On considère en ce sens que la création doit être consciente, comme impliquant « un minimum de maîtrise intellectuelle du processus créatif »<sup>45</sup>, ce qui exclut de la protection les créations des déments et des *infans*. Dans ce contexte, on peut affirmer sans hésitation que l'IA n'a aucune conscience. Et on pourrait arrêter là la démonstration et tout simplement nier la capacité créative des IA<sup>46</sup>.

Toutefois, le critère de conscience de la création n'est pas unanimement admis<sup>47</sup>, certains préférant évaluer l'œuvre dans son résultat et non dans le processus qui a conduit à son émergence<sup>48</sup>. En outre, les œuvres anonymes ou pseudonymes<sup>49</sup> sont bien protégées (CPI, art. L. 113-6), alors qu'il est impossible d'aller sonder le degré de conscience de leurs auteurs. Le juge s'est d'ailleurs toujours désintéressé de telles interrogations. Les arguments tirés de ce que l'œuvre aurait été créée dans un état d'hallucination, ou par accident, n'ont jamais été retenus, de même que celui d'une création sous l'emprise d'un « *agent psychique* », sorte d'esprit convoqué par un médium<sup>50</sup>. L'argument n'est pas sans portée : comme l'explique

---

<sup>44</sup> M. Vivant, « Intelligence artificielle et propriété intellectuelle », *Comm. comm. électr.* 2018, étude 18, n° 10.

<sup>45</sup> A. Lucas, A. Lucas-Schloetter, C. Bernault, *Traité de la propriété littéraire et artistique*, *op. cit.*, n° 57 (et n° 59 sur les créations générées par une IA).

<sup>46</sup> L. Devillers, *Des robots et des hommes, Mythes fantasmes et réalités*, Plon, 2017, spéc. p. 79.

<sup>47</sup> M. Vivant et J.-M. Bruguière, *Droit d'auteur et droits voisins*, Dalloz, 4<sup>e</sup> éd., 2019, n° 119 ; J. Larrieu, « Le robot et le droit d'auteur », in *Mélanges A. Lucas*, LexisNexis, 2014, p. 465 s., spéc. p. 467, et B. Michaux, « Singularité technologique, singularité humaine et droit d'auteur », in *Droit, normes et libertés dans le cybermonde, Mélanges Y. Pouillet*, Larcier, 2018, p. 401 s., spéc. p. 410.

<sup>48</sup> Not. N. Enser, *Conscience et création en droit d'auteur*, thèse, univ. Paris-Sud/Paris-Saclay, n° 373 : « la création ne peut être appréhendée que comme un résultat qui se constate, sans qu'il ne soit utile de s'attacher au processus qui l'a conduite à voir le jour ».

<sup>49</sup> CPI, art. L. 123-3.

<sup>50</sup> Il existe des ex. étrangers : v. A. Bridy, « *The Evolution of Authorship: Work Made by Code* », 39 *Colum. J.L. & Arts* 395 (2016), p. 18 s., citant les arrêts *Bell v. Catalda* 191 F 2d 99 (2d Cir 1951) ; *Penguin Books v. New Christian Church*, 55 USPQ 2d 1680 (SDNY 2000) ; *Cummins v. Bond* [1927] 1 Ch 167 ; *Leah v. Two Worlds Pub'g* [1951] 1 Ch 393, concernant des œuvres supposément créées par des esprits, des voix, ou un Dieu.

Jacques Larrieu, il ne viendrait à l'idée de personne de distinguer les œuvres de Van Gogh selon qu'elles ont été peintes à des stades différents de sa vie, lorsque son degré de conscience du monde qui l'entourait n'était pas des plus clairs (songeons à l'autoportrait à l'oreille coupée...). On pourrait en dire autant des *Fleurs du mal*, de Baudelaire, dont certains passages ont été écrits sous l'emprise des « paradis artificiels ».

Mais si la conscience n'est pas une exigence impérieuse, la création renvoie encore à la notion de **créateur en tant que personne physique**. L'acte de création protégé par le droit d'auteur s'oppose au caractère mécanique de la « production » de contenu par la machine. Comme l'expliquent les scientifiques, les logiciels seront à jamais « dépourvus d'émotion »<sup>51</sup>, celle-ci revenant à l'artiste qui opère l'IA et qui sélectionne et aménage les modèles qu'elle lui propose. Il y aurait donc systématiquement un humain derrière la création culturelle.

Juridiquement, l'exigence découle, en France, de la nature même du droit d'auteur, humaniste et personnaliste, et de l'esprit de la loi. On peut noter à ce titre que la loi évoque l'auteur avant même de préciser l'œuvre (CPI, art. L. 111-1 et -2). Par ailleurs, l'attribution des droits à une personne physique est mentionnée pour certains types d'œuvres<sup>52</sup> et on comprendrait mal pourquoi le particulier dérogerait sur ce point. Pour sa part, la Cour de cassation a jugé qu'« une personne morale ne peut avoir la qualité d'auteur »<sup>53</sup>, impliquant *a contrario* que seule une personne physique peut être auteur. Ce lien, entre création et personne physique, semble d'ailleurs partagé au niveau international<sup>54</sup>. Aux États-Unis, l'enregistrement d'une œuvre au *Copyright Office* n'est autorisé que si elle a été réalisée par un être humain<sup>55</sup>. En Australie, la Cour suprême a refusé la protection à une base de données générée automatiquement par une IA<sup>56</sup>.

La création semble donc l'apanage de l'humain, le fruit de son imagination. Pourtant, elle n'est pas un « acte magique » : elle répond à des règles, y compris lorsqu'il est décidé de s'écarter de ces dernières. Or, l'IA est par nature **biomimétique** – elle simule, imite l'humain ; elle peut donc aussi imiter l'arbitraire des choix humains, au sens d'activité créatrice. Aussi, les règles de création pourraient sans doute être informatiquement codées<sup>57</sup>, au moins en partie.

---

<sup>51</sup> H. Bersini : <https://lactualite.com/techno/lintelligence-artificielle-ne-sera-jamais-artiste/>

<sup>52</sup> CPI, art. L. 113-2 pour l'œuvre de collaboration, art. L. 113-7, al. 1<sup>er</sup>, pour l'œuvre audiovisuelle, art. L. 113-8, al. 1<sup>er</sup>, pour l'œuvre radiophonique.

<sup>53</sup> Cass. 1<sup>re</sup> civ., 15 janv. 2015, n° 13-23.566, D. 2015, p. 206 et p. 2215, obs. C. Le Stanc ; RTD com. 2015, p. 307, chron. P. Gaudrat ; Légipresse 2015, p. 223, note N. Binctin ; Comm. comm. électr. 2015, n° 19, obs. C. Caron ; Propr. industr. 2015, n° 3, comm. 25, obs. N. Bouche.

<sup>54</sup> A. Dietz, « Le concept d'auteur selon le droit de la Convention de Berne », RIDA janv. 1993, n° 155, p. 3, spéc. p. 11-13 ; S. Ricketson, « *People or Machines: The Berne Convention and the Changing Concept of Authorship* », 16 Columbia-VLA Journal of Law & the Arts, 1 (1991).

<sup>55</sup> *Compendium of U.S. Copyright Office practices*, § 306 et § 313.2.

<sup>56</sup> Cour suprême australienne, *Telstra Corporation Limited v Phone Directories Company Pty Ltd* [2010], FCAFC 149 : « *Following the High Court's reasoning in the 2009 decision of IceTV v Nine Network, the Full Federal Court recognised that to be original, a work must:*

*1. Not be copied;*

*2. Originate from a human author; and*

*3. Be the result of independent intellectual effort.*

*Secondly, the Full Court found that the extraction phase effort did not originate from an individual or group of individuals, but rather from a computerised process of storing, selecting, ordering and arranging the data to produce the directories in their published form. The Court emphasised that it was the extraction phase work that constituted the relevant effort for the purposes of originality. »*

<sup>57</sup> V. la démonstration de M. A. Boden, « *Computer Models of Creativity* », Association for the Advancement of Artificial Intelligence, 2009, p. 23.

Reste qu'il faudra identifier un créateur, personne physique, mais sans doute plus éloigné de la création que dans l'approche traditionnelle du droit, par exemple l'humain dont l'imagination a pu être codée. L'acte de création pourrait donc être caractérisé pour autant qu'on accepte l'idée d'une **création différente** et d'un **créateur plus indirect** (v. *infra* 2.1.2.3).

### 2.1.2.2. Les « œuvres » générées par l'IA sont-elles « originales » ?

La protection par le droit d'auteur suppose surtout que la création dont la protection est sollicitée soit jugée « originale ». Les créations générées par IA peuvent-elles remplir ce critère ?

En France règne traditionnellement une conception **subjective** de l'originalité, définie depuis Desbois, comme l'**empreinte de la personnalité de l'auteur**. La jurisprudence et la doctrine ont adopté une conception anthropocentrique du droit d'auteur, qui protège le travail créateur de l'être humain. L'œuvre doit dès lors pouvoir être retracée jusqu'à son auteur, qui en est l'*origine*, qui a laissé en elle des **traces** de sa personne – même si l'exécution personnelle n'est pas une condition requise pour cela<sup>58</sup>.

Mais cette approche classique est doublement remise en cause : en droit interne, par une certaine **objectivation du critère d'originalité**, comme la « marque de l'apport intellectuel »<sup>59</sup> ; en droit européen, depuis que la CJUE en a fait une **notion autonome de droit de l'Union**, entendue comme « une création intellectuelle propre à son auteur »<sup>60</sup>, même si la notion semble conserver en droit européen une part de subjectivité<sup>61</sup>.

---

<sup>58</sup> Cass. 1<sup>re</sup> civ., 13 nov. 1973, n° 71-14.469 : Bull. civ. I, n° 302 ; D. 1974, jur. p. 533, note C. Colombet ; JCP G 1975, II, 18029, note M.-Ch. Manigne ; v. également B. Edelman, « La main et l'esprit », D. 1980, chron. p. 7.

<sup>59</sup> Cass. Ass. plén., 7 mars 1986, n° 83-10477, *Babolat c/ Pachot*, v. not. JCP E 1986, II, 14713 et JCP G 1986, II, 20631, note J.-M. Mousseron, B. Teyssié et M. Vivant ; JCP E 1986, I, 15791, obs. M. Vivant et A. Lucas ; D. 1986, p. 405, conclusions de l'avocat général Cabannes, et p. 412, note B. Edelman ; RIDA, juill. 1986, n° 129, p. 136, note A. Lucas ; RTD com. 1986, p. 399, obs. A. Françon.

<sup>60</sup> CJUE, 16 juill. 2009, *Infopaq*, aff. C-5/08 ; 22 déc. 2010, *BSA*, aff. C-393/09 ; 4 oct. 2011, *Premier League*, aff. jointes C-403/08 et C-429/08 ; 1<sup>er</sup> mars 2012, *Football Dataco*, aff. C-604/10 ; 2 mai 2012, *SAS Institute*, aff. C-406/10 ; 13 nov. 2018, *Levola*, aff. C-310/17.

<sup>61</sup> Dans les arrêts précités, certaines expressions jurisprudentielles pouvaient faire croire que la cour avait favorisé la conception subjective française. Cependant, lorsque la cour faisait référence aux choix créatifs ou à la touche personnelle, elle visait des photographies, pour lesquelles cette approche subjective est présente dans le texte (v. cons. 16 de la Directive 2006/116/CE du 12 déc. 2006 relative à la durée de protection du droit d'auteur et de certains droits voisins - version codifiée). On pouvait dès lors considérer que l'approche subjective était cantonnée à la photographie. Cette analyse ne peut cependant plus être tenue dans la mesure où la cour utilise désormais ces expressions subjectives pour d'autres types d'œuvres : CJUE, 12 sept. 2019, aff. C-683/17, *Cofemel*, cons. 54 : « D'autre part, il est certes vrai que des considérations d'ordre esthétique participent de l'activité créative. Toutefois, il n'en reste pas moins que la circonstance qu'un modèle génère un effet esthétique ne permet pas, en soi, de déterminer si ce modèle constitue une *création intellectuelle reflétant la liberté de choix et la personnalité de son auteur*, et satisfaisant donc à l'exigence d'originalité évoquée aux points 30 et 31 du présent arrêt » (nous soulignons) ; CJUE, 29 juill. 2019, aff. C-469/17, *Funke*, cons. 19 : « Ainsi qu'il ressort d'une jurisprudence bien établie, pour qu'un objet puisse être qualifié d'« œuvre », il importe que soient réunies deux conditions cumulatives. D'une part, l'objet concerné doit être original, en ce sens qu'il constitue une création intellectuelle propre à son auteur. Pour qu'une création intellectuelle puisse être considérée comme étant propre à son auteur, celle-ci doit refléter la personnalité de celui-ci, ce qui est le cas si l'auteur a pu exprimer ses capacités créatives lors de la réalisation de l'œuvre en effectuant des choix libres et créatifs (voir, en ce sens, arrêt du 1<sup>er</sup> décembre 2011, *Painer*, C-145/10, EU:C:2011:798, points 87 à 89). »

La critique vient aussi de la doctrine. Certains auteurs estiment ainsi que la valeur reconnue à l'œuvre par le public doit rentrer en ligne de compte dans l'appréciation de l'originalité : cette objectivation éloignerait le droit d'auteur de la personne de l'auteur et permettrait d'envisager que tout produit qui aurait été reconnu, en termes mesurables, comme étant une œuvre de l'esprit par son public puisse recevoir la protection du droit d'auteur<sup>62</sup>. D'autres auteurs suggèrent une approche plus objective de la condition – peut-être aussi plus conforme à la réalité judiciaire –, proposant par exemple que l'originalité soit appréhendée comme la « nouveauté dans l'univers des formes »<sup>63</sup>.

Dans tous les cas, l'originalité *de* l'œuvre ne devrait-elle pas être recherchée *dans* l'œuvre justement, et non par référence à son auteur ? D'ailleurs, comment faire la démonstration lorsque l'auteur est décédé ou anonyme ? Aussi, le critère d'originalité devrait rester **intrinsèque à l'œuvre**, comme une sorte de manifestation créative minimum<sup>64</sup> inscrite dans la forme. Précisément, on pourrait imaginer qu'une forme intrinsèquement originale devrait, en quelque sorte, passer le **test de Turing**<sup>65</sup> et **pouvoir tromper sur son origine**. Il s'agirait alors d'une création *assimilable* à une œuvre, en tant qu'elle lui ressemble et qu'elle remplit les mêmes fonctionnalités.

Même sans aller jusqu'à ce renouvellement de la notion d'originalité, on pourrait aussi accepter de rechercher une personnalité **plus indirecte**, plus éloignée, celle du concepteur de l'IA notamment, qui délimite la cadre de la création algorithmique en façonnant le modèle d'inférence<sup>66</sup>. Il ne s'agit pas ici de s'intéresser uniquement au logiciel – il est évident que l'originalité du logiciel n'est pas reproduite au stade de l'œuvre produite –, mais à l'IA en tant qu'œuvre complexe dont les caractéristiques – les biais – contraignent la création générée.

Si l'on admet une certaine réinterprétation des critères d'accès à la protection, il devient possible d'accueillir les produits générés par l'intelligence artificielle au sein du droit d'auteur. Reste à vérifier le rattachement originaire, à un auteur.

### 2.1.2.3. Les « œuvres » générées par une IA ont-elles un auteur ?

La conception contemporaine qui prédomine largement dans le droit positif, même le plus éloigné de l'humanisme continental (comme au Royaume-Uni ou en Nouvelle-Zélande), implique très nécessairement que l'auteur d'une œuvre ne puisse qu'être une **personne humaine**. D'où la nécessité d'identifier une personne en lien avec la création générée par l'IA, et c'est la troisième difficulté que pose le rattachement de ces produits générés par l'IA sous le boisseau du droit d'auteur.

---

<sup>62</sup> Hypothèse avancée par S. Yanisky & L. A. Velez-Hernandez, « *Copyrightability of Artworks produced by creative robots and originality: the Formality-Objective Model* », Minnesota Journal of Law, Science & Technology, Vol. 19, p. 18

<sup>63</sup> M. Vivant et J.-M. Bruguère, *Droit d'auteur et droits voisins*, Dalloz, Précis, 4<sup>e</sup> éd., 2019, n° 271. – Adde J. Larrieu, « Le robot et le droit d'auteur », *op. cit.* ; P.-Y. Gautier, *Propriété littéraire et artistique*, PUF, coll. Droit fondamental, 11<sup>e</sup> éd., 2019, n° 35.

<sup>64</sup> Comp. US Supreme Court, *Feist Publications, Inc. v. Rural Telephone Service*, 499 U.S. 340, 345 (1991) « *Original, as the term is used in copyright, means only that the work was independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity* ».

<sup>65</sup> A. Turing, « *Computing machinery and intelligence* », *Mind*, Oxford University Press, Volume LIX, Issue 36, Oct. 1950, p. 433, <https://academic.oup.com/mind/article/LIX/236/433/986238>

<sup>66</sup> V. les points suivants.

Doit-on alors considérer que l'IA peut être auteur ? L'idée d'une **personnalité juridique de l'intelligence artificielle** a déjà été avancée dans une Résolution du Parlement européen en 2017<sup>67</sup> et certains auteurs s'en sont faits les chantres<sup>68</sup>. Dans sa Résolution, le Parlement évoquait ainsi « la création, à terme, d'une personnalité juridique spécifique aux robots, pour qu'au moins les robots autonomes les plus sophistiqués puissent être considérés comme des personnes électroniques responsables, tenus de réparer tout dommage causé à un tiers »<sup>69</sup>.

Le présent rapport ne propose pas de s'attarder sur cette option, qui apparaît largement **impraticable**<sup>70</sup> : outre les difficultés éthiques et philosophiques qu'elle ne manquerait pas de soulever, l'octroi d'une personnalité juridique à l'intelligence artificielle ne résoudrait aucun des problèmes concrets posés par l'IA. D'une part, un tel bouleversement supposerait de définir les IA éligibles à la personnalité juridique des autres IA qui n'y sont pas, opération de qualification extrêmement complexe ; d'autre part, cette option se heurterait à l'identification d'un patrimoine autonome dans lequel seraient versés les droits d'auteur – et qui serait susceptible d'être gagé<sup>71</sup> ; encore, l'attribution d'une personnalité juridique ne résoudrait pas la question de l'autorité liée, en France notamment, à l'existence d'une personne physique – et non seulement d'une personnalité juridique ; enfin, elle risquerait de défausser les producteurs de l'IA de leurs propres responsabilités dans les résultats auxquels le fonctionnement de l'IA pourrait aboutir<sup>72</sup>. Une IA ne saurait donc être qualifiée d'auteur, ni d'ailleurs d'inventeur, comme l'a indiqué l'Office européen des brevets<sup>73</sup>.

Aussi, si l'identification d'un humain-créateur est nécessaire en l'état du droit, elle est peut-être à rechercher ailleurs, dans une relation créative inédite, une **relation intermédiée**. L'auteur serait donc plus éloigné que dans une analyse traditionnelle, mais il serait toujours à la base de choix qui influencent la création. Il convient alors d'évaluer les options de titularité.

### **2.1.3. La faisabilité de l'application du droit d'auteur aux produits générés par l'IA (questions de titularité)**

Rattacher les créations générées par l'IA au droit d'auteur implique, une fois que l'on en admet le principe, d'en vérifier la faisabilité juridique. Ainsi, **retenant une approche objectivée des critères d'accès à la protection, la voie du droit d'auteur semble ouverte, sauf à identifier un auteur personne physique**. La difficulté tient alors au fait que le processus créatif est obscurci par l'utilisation de l'IA, distendant le lien entre l'auteur et la

---

<sup>67</sup> Résolution du Parlement européen du 16 févr. 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique, 2015/2013(INL).

<sup>68</sup> A. Bensoussan, *Droit des robots*, Larcier, 2015, p. 41 s. et « La personne robot », D. 2017, p. 2044 ; M. Willick, « *Artificial Intelligence: Some legal approaches and applications* », (1983) 4 :2 AI Mag 5.

<sup>69</sup> Résolution précit., 2015/2013(INL), § 59 f).

<sup>70</sup> V. G. Loiseau, « La personnalité juridique des robots : une monstruosité juridique », JCP G 2018, 597 ; A. Bensamoun et G. Loiseau, « L'intégration de l'intelligence artificielle dans l'ordre juridique en droit commun : questions de temps », Dalloz IP/IT, avr. 2017, p. 239.

<sup>71</sup> R. Pearlman, « *Recognizing Artificial Intelligence (AI) as Authors and Inventors under US Intellectual Property Law* », 24 Rich. J. L. & Tech, no. 2, 2018 ; P. Samuelson, « *Allocating Ownership Rights in Computer-Generated Works* », 47 U. Pitt. L. Rev. 1185 (1985).

<sup>72</sup> G. Courtois, « Robots intelligents et responsabilité : quels régimes, quelles perspectives ? », Dalloz IP/IT 2016, p. 287 ; M. Bacache, « Intelligence artificielle et droits de la responsabilité et des assurances », in *Droit de l'intelligence artificielle*, dir. A. Bensamoun et G. Loiseau, LGDJ-Lextenso, coll. Les Intégrales, 2019, p. 69.

<sup>73</sup> Refus de deux demandes de brevets européens dans lesquelles une IA, « DABUS », était désignée comme inventeur : [epo.org/news-issues/news/2019/20191220.html](https://epo.org/news-issues/news/2019/20191220.html) (décision motivée annoncée pour courant janvier 2020).

création, là où ce lien – parfois vu comme un « cordon ombilical » – est une des caractéristiques du droit d’auteur « à la française ».

Plusieurs personnes pourraient être concernées, de l’utilisateur (2.1.3.2) au concepteur de l’IA (2.1.3.1). Il pourrait encore être proposé une identification légale, sur le modèle du droit anglais notamment (2.1.3.3).

### 2.1.3.1. Première hypothèse de titularité : cibler le concepteur de l’IA

Le droit d’auteur pourrait, dans une première hypothèse, remonter au concepteur de l’IA<sup>74</sup> (précisément celui qui réalise le modèle d’inférence qui fixe le cadre de la création). Ce concepteur définit le champ des possibles dans lequel l’IA va ensuite se mouvoir. Dans ce cas, l’IA ne fait en définitive qu’exécuter le cadre créatif fixé par le concepteur. Cette hypothèse a l’avantage d’identifier une personne humaine qui aura entretenu des liens très proches avec la machine, qui a été capable d’y **inoculer ses propres biais** et donc d’y laisser une certaine empreinte. Comme le relève P. Samuelson, dans de nombreux cas, l’IA aurait généré le même produit quelle que soit la personnalité de son utilisateur humain<sup>75</sup>. C’est alors l’auteur de l’IA qui semble le mieux placé pour recevoir les droits au titre des créations générées par la machine qu’il aura conçue. **Créateur « indirect »**<sup>76</sup>, le concepteur de l’IA guide en quelque sorte la réalisation algorithmique.

Désigner le « concepteur » de l’IA comme auteur des produits qu’elle génère ne lève cependant pas toutes les incertitudes.

L’une d’entre elles tient à l’**identification de ce personnage** : est-il le programmeur de l’algorithme, qui pose l’espace propre de la machine par des règles prédéfinies ? Est-il son « entraîneur », c’est-à-dire celui qui corrige l’IA en fonction de ses réactions spontanées à l’aléatoire, qui ajuste son comportement par des choix critiques dans la phase d’apprentissage sans avoir à expliciter ses choix propres par des règles prédéfinies<sup>77</sup> ? Est-il celui qui sélectionne les données, le « *training set* » ou « corpus d’entraînement » et qui inocule ses propres biais de sélection dans la masse de données qu’il met à la disposition de l’IA<sup>78</sup> ? Est-il, de manière plus générale et plus systématique, celui qui dispose du droit d’auteur sur le logiciel de l’IA lui-même ? Ou faut-il, dans cette création plurielle, s’intéresser davantage à celui qui a dirigé la création ? Par référence au régime de l’œuvre collective, prévu à l’article L. 113-5 du CPI, il serait alors sans doute pertinent de désigner la personne qui est à l’initiative du projet et l’a dirigé.

Une deuxième difficulté réside dans l’identification du **mécanisme juridique** par lequel les droits d’auteur sur la création générée par ordinateur se transmettraient au « concepteur » de

---

<sup>74</sup> V. en ce sens, not. Y. Michaux, « Singularité technologique, singularité humaine et droit d’auteur », *op. cit.* ; T. Lebrun, « L’apprentissage machine est une appropriation », *Les Cahiers de PI* 2018, vol. 30, n° 3, p. 895 ; Y. Gaubiac, *JCl. Propriété Littéraire et Artistique*, fasc. 1164, qui rappelle l’importance du rôle du concepteur de l’IA dans les résultats finaux qu’elle produit.

<sup>75</sup> P. Samuelson, « *Allocating Ownership Rights in Computer-Generated Works* », *op. cit.*

<sup>76</sup> B. Michaux, « Singularité technologique, singularité humaine et droit d’auteur », *op. cit.*, p. 413-414 ; T. Dreier, « *Creation and investment: artistic legal implications of computer-generated works* », *International Computer Law Adviser*, n° 3-4, vol. 5, 1991, spec. p. 16.

<sup>77</sup> C’est l’une des hypothèses défendue par J.-M. Deltorn, « Droit d’auteur et créations des algorithmes d’apprentissage », *PI* janvier 2016, n° 58, p. 4.

<sup>78</sup> Hypothèse aussi évoquée par J.-M. Deltorn, *ibid.* V. également F. Macrez et J.-M. Deltorn, « *Authorship in the age of machine learning and artificial intelligence* », *op. cit.*, qui prennent l’exemple d’*Emily Howell*, l’IA créée par David Cope basée sur un corpus de données choisies à partir de compositions classiques.

l'infrastructure génératrice de ces créations. Des auteurs de doctrine, notamment américains, marquent leur prédilection pour une analyse en termes de droit dérivé : le code de l'IA se retrouverait ainsi en filigrane, quoiqu'invisiblement, dans les créations générées par l'IA<sup>79</sup> ; dans cette hypothèse, la personne titulaire des droits sur l'IA serait nécessairement auteur de l'œuvre générée par l'IA. Les tenants de cette analyse admettent toutefois que celle-ci est subordonnée à ce que le juge soit capable d'identifier, dans l'œuvre générée, une expression reconnue du code lui-même<sup>80</sup>, ce qui peut se révéler compliqué, voire impossible (à tout le moins inédit ; la méthode serait à inventer). De ce point de vue, le produit généré par l'IA ne pourra pas prendre appui sur la qualification d'œuvre composite ou dérivée<sup>81</sup>. En effet, l'hypothèse se heurte à l'impossibilité de retrouver dans l'œuvre seconde (la création générée par l'IA) les éléments de l'œuvre première (l'IA elle-même)<sup>82</sup>.

Une autre assise pourrait être tirée du droit des biens, précisément du mécanisme de l'**accession par production**<sup>83</sup>, qui permet au propriétaire d'acquérir les accessoires que produit sa chose, augmentant l'assiette de la propriété principale. Ainsi, la chose frugifère, l'IA, génère des œuvres qui sont ses fruits et pourraient alors, par accession, devenir la propriété du ou des titulaires de droits sur l'IA<sup>84</sup>. La jurisprudence a déjà fait application de ces règles pour organiser le sort des redevances tirées de l'exploitation d'un brevet<sup>85</sup> ou de marques protégées<sup>86</sup>. Cette ouverture serait d'autant plus justifiée que le CPI admet lui-même, dans des cas particuliers, que le droit d'auteur soit attribué à d'autres personnes que la personne physique qui assure la création. Ainsi en est-il par exemple des œuvres posthumes ou collectives. En conséquence, on peut considérer que les modes d'appropriation du CPI ne sont pas fermés et laissent place à l'attribution des droits sur la création intelligente par voie d'accession.

Outre cette question pratique et théorique de mise en œuvre, la désignation de l'auteur de l'IA (son concepteur) comme auteur des créations générées par l'IA peut susciter des remarques.

Tout d'abord, la distinction entre la machine elle-même et le produit créé par la machine se fait de mieux en mieux. En effet, ainsi que le rappelle Anne-Marie Bridy, l'un des arguments commerciaux des concepteurs de l'IA a justement trait aux **interactions** promises entre la machine et ses utilisateurs, interactions visant à apporter à ces derniers le produit qui correspond le mieux à leur personnalité. Le dispositif est susceptible d'émanciper l'outil de sa programmation initiale. Il éloigne donc le concepteur du fonctionnement quotidien de son algorithme. C'est d'autant plus vrai que le concepteur de l'IA dispose toujours du choix de ne pas commercialiser son programme et de faire passer les œuvres générées comme les siennes. En outre, ayant déjà reçu une rémunération en échange de la vente ou de l'octroi de licences d'utilisation de son programme, et disposant déjà, dans la majorité des cas, d'un droit d'auteur

---

<sup>79</sup> Hypothèse envisagée par A. Bridy, « *Coding Creativity : Copyright and the Artificially Intelligent Author* », précit., p. 25.

<sup>80</sup> V. P. Samuelson, *Allocating Ownership Rights in Computer-Generated Works*, op. cit.

<sup>81</sup> Art. L. 113-2 et art. L. 113-4 du CPI.

<sup>82</sup> OMPI, « Recommandations en vue du règlement des problèmes découlant, sur le plan du droit d'auteur, de l'utilisation de systèmes informatiques pour l'accès aux œuvres ou pour la création d'œuvres – Utilisation de systèmes informatiques pour la création d'œuvres protégées », Droit Auteur 1982, 115, 9.

<sup>83</sup> C. civ., art. 546 : « La propriété d'une chose soit mobilière, soit immobilière, donne droit sur tout ce qu'elle produit, et sur ce qui s'y unit accessoirement soit naturellement, soit artificiellement. Ce droit s'appelle droit d'accession ». – En ce sens, v. P.-Y. Gautier, « De la propriété des créations issues de l'intelligence artificielle », JCP G 2018, 913.

<sup>84</sup> C'est la position de la SACEM.

<sup>85</sup> Cass. com., 28 avr. 2004, n° 02-21585, Bull. civ. IV, n° 80.

<sup>86</sup> CA Orléans, 10 juill. 2003, n° RG 02/01964, PIBD 2003, n° 773.

sur l'IA (à la considérer comme une œuvre complexe), il n'est pas, selon certains, la personne la plus légitime à réclamer des droits sur les créations qu'elle a générées.

De plus, le caractère **automatique** de la titularité a été critiqué comme permettant de faire l'économie d'une analyse portant sur l'originalité de l'œuvre générée, celle-ci ne pouvant pas par définition être caractérisée au cas par cas<sup>87</sup>. Néanmoins, si l'on accepte d'évaluer l'originalité de manière intrinsèque, c'est-à-dire dans l'œuvre, en relevant si elle est ou non créative, l'opposition pourrait être levée (par le juge en l'occurrence, comme c'est d'ailleurs le cas pour toutes les œuvres). Il y aurait donc bien une analyse, au cas par cas.

Pour autant, l'automaticité est aussi en elle-même critiquable, en ce qu'elle implique un raisonnement fictif qui consiste à faire comme si l'IA contenait déjà potentiellement toutes les œuvres à créer et ce, sans même que le concepteur n'ait une idée de l'immensité des débouchés de sa propre création<sup>88</sup>. Cette approche est, à tout le moins, inédite en droit d'auteur.

L'objection présente d'ailleurs des traductions concrètes : notamment, comment le concepteur de l'IA pourrait-il faire assurer le respect de ses droits alors que l'instrument est entre les mains d'un tiers ? En pratique, le contrat règlera sans doute les choses. Par ailleurs, comment concilier la volonté légitime de l'utilisateur de « diffuser » les réalisations qu'il a participé à générer avec le monopole du concepteur ? Est-il pertinent d'accorder un droit moral plein et entier, qui pourrait d'ailleurs indirectement venir augmenter le monopole d'exploitation du concepteur ?<sup>89</sup> Comment permettre à l'utilisateur d'user comme il le souhaite de l'instrument, de modifier l'*output*, sans porter atteinte à l'intégrité ? Et devrait-on considérer que le droit de divulgation est présumé ? Par ailleurs, comment gérer les productions *post-mortem*, c'est-à-dire les créations générées par l'IA alors que le concepteur de l'IA est mort ? Enfin, quel serait le point de départ de la protection ? La « publication » de l'œuvre, suivant en cela la logique des œuvres spéciales que sont les œuvres anonymes et pseudonymes ou les œuvres collectives ?<sup>90</sup>

Mais ce sont là des questions qui pourraient se poser en pratique et que la jurisprudence pourrait appréhender en partie. Dans tous les cas, il apparaît que **la désignation du concepteur de l'IA apparaît comme la solution la plus respectueuse du droit d'auteur**. Il convient cependant, à des fins de complétude, d'envisager les autres propositions.

### 2.1.3.2. Deuxième hypothèse de titularité : cibler l'utilisateur de l'IA

L'utilisateur de l'IA est celui qui a la maîtrise physique de l'outil et de l'élection du résultat en tant qu'œuvre. La désignation de l'utilisateur de l'IA comme auteur des créations générées par l'IA a pour elle l'avantage de la **simplicité** dès lors qu'elle abolit la distinction entre création assistée et création générée par ordinateur. Certains y voient également une certaine **légitimité financière**, dans la mesure où l'utilisateur aura souvent acheté une licence d'exploitation<sup>91</sup>. Cette solution est encore en accord avec la faisabilité technique : c'est

---

<sup>87</sup> Problèmes notamment pointés par P. Samuelson, *op. cit.*

<sup>88</sup> Arguments soulevés par Y. Gaubiac, G. Azzaria et A. Bridy dans leurs articles précit.

<sup>89</sup> V. les interrogations de V.-L. Benabou, « *AI, moral right and adaptation right* », *EU copyright, quo vadis ? From the EU copyright package to the challenges of Artificial intelligence*, colloque European Copyright Society, Bruxelles, mai 2018 : <https://europeancopyrightsocietydotorg.files.wordpress.com/2018/06/benabou-presentation-ai-moral-rights1.pdf>.

<sup>90</sup> CPI, art. L. 123-3.

<sup>91</sup> Arguments repris par P. Samuelson, « *Allocating Ownership Rights in Computer-Generated Works* », *op. cit.*

l'utilisateur qui dispose de la **garde** concrète du programme et qui divulguera l'œuvre ; il a ainsi un rôle incontournable<sup>92</sup>. Ces avantages ont justifié qu'une telle hypothèse ait reçu l'agrément d'une partie de la doctrine.

Se pose alors la question du critère de rattachement de l'auteur à l'œuvre. Le premier critère possible est l'existence d'un **choix créatif minime** : l'utilisateur arrête le processus et déclare le produit généré par l'IA comme étant une œuvre. Mais cela suffit-il à faire auteur ? D'autres auteurs, notamment américains, évoquent un deuxième critère pertinent en convoquant la doctrine du « *work made for hire* ». Cette théorie américaine<sup>93</sup> prévoit que dans certains cas, la titularité des droits sur une œuvre protégée créée par un agent commandité, auteur réel, remonte au commanditaire, auteur fictif, dès lors que c'est ce dernier qui l'a commandée et que c'est en son nom qu'elle aura été réalisée. Appliquée à l'IA, cette doctrine voudrait ainsi que l'IA soit « l'auteur en fait » de l'œuvre, mais que seul l'utilisateur soit son « auteur en droit » dès lors que c'est à sa demande que l'IA a produit l'œuvre pour laquelle la protection du droit d'auteur est recherchée<sup>94</sup>.

Cette hypothèse présente cependant des **oppositions de principe**. D'une part, elle pourrait être, dans de nombreux cas, **fictive**. Si l'utilisateur réalise, en aval, des choix créatifs, nul doute qu'il pourrait être qualifié d'auteur (voire de coauteur de l'œuvre de collaboration générée, si on retient aussi que le concepteur, en amont, a imprimé sa marque). Mais s'il ne fait aucun choix créatif, se contentant d'appuyer sur un simple bouton, la solution ne devrait pas être retenue. Le droit d'auteur perdrait de sa force s'il devait « récompenser » le seul acte de commandement d'une création ou tout acte annexe à la création elle-même. Il ferait dès lors fi de tous les principes. Ainsi que la Cour de cassation l'a rappelé, « la qualité d'auteur ne peut être reconnue à la personne qui s'est limitée à fournir une idée ou un simple thème »<sup>95</sup>. D'autre part, l'attribution d'un quelconque mérite à l'acte déclaratif par lequel un consommateur d'IA juge le produit spontanément généré par la machine comme étant « une œuvre » est, là aussi, fictive : la qualification serait quasiment vidée de son sens et les tribunaux se contenteraient de faire office de chambre d'enregistrement. Une telle solution contreviendrait d'ailleurs à la loi, dans la mesure où elle aboutirait à une protection automatique, en considération du genre de l'œuvre<sup>96</sup>.

### 2.1.3.3. Imposer légalement un régime de titularité

Pour lever les doutes existants, il est aussi envisagé de fixer légalement l'autorat et la titularité.

Aussi, certains auteurs de doctrine proposent de s'inspirer du **droit anglais**, qui met en place un régime dérogatoire pour les « œuvres générées par ordinateur », définies comme celles

---

<sup>92</sup> R. C. Denicola, « *Ex Machina: Copyright Protection for Computer Generated Works* », 69 Rutgers UL Rev., 251 (2016).

<sup>93</sup> V. art. 201 du *US Copyright Act*.

<sup>94</sup> V. A. Bridy, « *The Evolution of Authorship: Work Made by Code* », 39 COLUM. J.L. & ARTS 395 (2016) ; S. Yanisky-Ravid et L. Antonio Velez-Hernandez, « *Copyrightability of Artworks produced by creative robots and originality : the Formality-Objective Model* », Minnesota Journal of Law, Science & Technology, Vol. 19, p. 18 ; M. E. Kaminsky, « *Authorship, Disrupted: AI Authors in Copyright and First Amendment Law* », University of Colorado Law Legal Studies Research Paper No. 17-26, vol. 51, 2017.

<sup>95</sup> Cass. Civ. 1<sup>re</sup>, 8 nov. 1983, Bull. Civ. I., n° 260.

<sup>96</sup> L'article L. 112-1 du CPI interdit de prendre en compte le genre, la forme d'expression, le mérite ou la destination pour l'accès à la protection.

réalisées « dans des circonstances telles qu'il n'y a pas d'auteur humain »<sup>97</sup>. Le *Copyright, Designs and Patents Act* pose alors une fiction en vertu de laquelle l'auteur est celui qui a pris les « dispositions nécessaires pour la création de l'œuvre »<sup>98</sup>. Beaucoup considèrent que le bénéficiaire visé est alors l'utilisateur, encore qu'on puisse légitimement en douter<sup>99</sup>. D'ailleurs, la doctrine reste relativement divisée sur l'identification de cet auteur réputé. Seul un arrêt en fait une brève mention, semblant identifier le programmeur d'un algorithme de génération d'images dans un jeu vidéo comme la personne ayant « pris les dispositions nécessaires » à la création de l'œuvre, et non l'utilisateur. Toutefois, cet arrêt fut rendu au titre de motifs différents, qui ne permettent pas d'en tirer une interprétation satisfaisante<sup>100</sup>.

La formulation légale fait donc ici naître une solution imparfaitement résolue, juste déplacée en réalité. Au lieu de chercher à qualifier l'auteur, on cherche à qualifier la personne qui prend les dispositions nécessaires en le désignant comme auteur...

En outre, une telle solution prend, là encore, des libertés évidentes avec les principes, puisqu'elle se propose d'investir de la qualité d'auteur une personne qui se contente de « commander » la création, d'en organiser la génération, loin de toute idée de choix créatifs. En quoi le commanditaire d'une œuvre pourrait-il en être l'auteur ? Les fresques de la Chapelle Sixtine sont bien de Michel-Ange et non du pape Jules II, comme l'écrit, non sans provocation, Jane Ginsburg<sup>101</sup>.

Enfin, le droit anglais assume un **régime spécial de protection**. Ainsi, la durée de la protection patrimoniale est réduite à cinquante ans, au lieu des soixante-dix ans de principe, et le point de départ est la réalisation de l'œuvre et non la mort de l'auteur<sup>102</sup>. S'agissant du droit moral, il est fait exception au droit à la paternité<sup>103</sup> et au droit au respect de l'intégrité de l'œuvre<sup>104</sup>.

En conclusion, même incertaine sur certains points, la piste du droit d'auteur est séduisante. Elle impliquerait, pour une grande part, l'interprétation du juge, dans un rôle cependant traditionnel en droit d'auteur d'adaptation du droit. Une autre solution pourrait être de fixer dans la loi une solution *ad hoc*.

## 2.2. Les pistes alternatives au droit d'auteur classique

Certains régimes existants pourraient utilement servir de modèles à la détermination d'un régime de protection pour les produits de l'IA. On peut ainsi mobiliser les régimes spéciaux de droit d'auteur (2.2.1 et 2.2.2) ou encore le droit *sui generis* (2.2.3).

---

<sup>97</sup> *Copyright, Designs and Patents Act* (CDPA), 1988, art. 178 (b).

<sup>98</sup> CDPA, art. 9 (3).

<sup>99</sup> A. Guadamuz, « *Do Androids Dream of Electric Copyright? Comparative analysis of originality in artificial intelligence generated works* », *Intellectual Property Quarterly*, 2 (2017) : « [i]f the artificial agent is directly started by the programmer, and it creates a work of art, then the programmer is clearly the author in accordance to s. 9 (3) CDPA. However, if a user acquires a program capable of producing computer-generated works, and uses it to generate a new work, then ownership would go to the user. »

<sup>100</sup> *Nova Productions Limited v. Mazooma Games Limited & Others*, [2006] EWHC 24 (Ch) §105, 20 janv. 2006, IIC04C02882.

<sup>101</sup> J. C. Ginsburg, « *People Not Machines: Authorship and What It Means in the Berne Convention* », 49 *International Review of Intellectual Property and Competition Law*, no. 2, 131 (2018).

<sup>102</sup> CDPA, art. 12 (7).

<sup>103</sup> CDPA, art. 79 (2) (a).

<sup>104</sup> CDPA, art. 81 (2).

### 2.2.1. La création d'un droit d'auteur spécial

Restant à la lisière du droit d'auteur, on pourrait d'abord s'inspirer de figures spécifiques. Le premier modèle qui vient à l'esprit est sans doute celui du **logiciel** (ou programme d'ordinateur), d'une part, pour sa proximité de nature – une création informatique –, d'autre part, pour le modèle d'aménagement du régime. En effet, la particularité du logiciel a imposé non seulement une nouvelle définition jurisprudentielle de l'originalité (dans l'arrêt *Pachot* précité), mais encore un régime légal complètement dérogatoire. Ainsi, le contenu du monopole d'exploitation diffère (CPI, art. L. 122-6) ; les exceptions sont spécifiques (CPI, art. L. 122-6-1) ; le droit moral est réduit à peau de chagrin (et à l'expression minimale de la Convention de Berne – « atteinte à l'honneur et à la réputation » ; CPI, art. L. 121-7) ; les règles de titularité organisent une « dévolution » légale à l'employeur (CPI, art. L. 113-9), faisant ainsi exception au principe de l'indifférence du cadre de création (CPI, art. L. 111-1, al. 2). En définitive, le droit spécial du logiciel est un droit d'auteur qui s'en éloigne par de nombreux points.

En matière de création générée par une IA, le choix est clairement **politique**. Si la volonté politique s'exprime (comme cela a été le cas le logiciel), elle pourrait alors se traduire par la création d'un nouveau droit, rattaché au droit d'auteur (du fait de la proximité apparente du résultat créé), mais disposant de règles spécifiques. Le modèle de réflexion (celui du logiciel) pourrait dès lors inspirer ce droit spécial, notamment s'agissant des règles de titularité ou encore du droit moral (en effet, dans la mesure où le lien à l'auteur est distendu, on pourrait justifier un régime aménagé de droit moral).

On pourrait ensuite utiliser le modèle de l'**œuvre collective**<sup>105</sup>, précisément parce que celle-ci représente une vision plus économique du droit d'auteur, moins centrée sur le créateur. A ce titre, elle est dérogatoire des règles classiques de titularité, puisque les droits naissent, indépendamment de tout transfert, sur la tête de l'instigateur qui, dans un processus créatif vertical, dirige la création et la publie sous son nom. Il serait alors loisible d'ajouter un quatrième alinéa à l'article L. 113-2 du CPI pour définir l'œuvre en question, comme « la création générée par une intelligence artificielle et à la réalisation de laquelle n'a concouru aucune personne physique ». Quant au régime associé, on pourrait l'intégrer à l'article L. 113-5 du Code de la propriété intellectuelle, avec celui de l'œuvre collective. Le nouvel article L. 113-5 énoncerait ainsi : « L'œuvre collective et l'œuvre générée par une IA, sont, sauf preuve contraire, la propriété de la personne physique ou morale sous le nom de laquelle elles sont divulguées. Cette personne est investie des droits de l'auteur ». Cela étant, cette solution, si elle définit un premier titulaire, ne permet pas d'identifier d'auteur, ce qui contrarie la rigueur du raisonnement. Mais une relecture des critères d'accès à la protection le permettrait (v. *supra*, 2.1.2). En outre, dans les décisions de jurisprudence les plus récentes en droit d'auteur, on sent bien que l'identification des auteurs d'une œuvre collective n'est pas une priorité, au point que les juges semblent parfois caractériser pareille œuvre, avec un titulaire, mais sans auteur<sup>106</sup>...

### 2.2.2. La création d'un droit d'auteur à la manière d'un « droit voisin »

Une autre option pourrait être de créer un droit spécifique, en droit d'auteur, mais aussi à la lisière d'un droit voisin.

---

<sup>105</sup> Sur cette proposition, v. N. Enser, *op. cit.*, n° 612.

<sup>106</sup> V. A. Bensamoun, « La personne morale en droit d'auteur : auteur contre-nature ou titulaire naturel ? », D. 2013, p. 376.

On pourrait ici s'inspirer du régime de l'**œuvre posthume**<sup>107</sup>, lequel accorde un droit d'exploitation au propriétaire du support matériel de l'œuvre qui assure sa publication, alors même que la durée de protection patrimoniale est expirée en vertu du droit commun<sup>108</sup>. Ainsi, comme le propriétaire de l'œuvre posthume, celui qui divulgue une création générée par une IA bénéficierait de prérogatives de nature patrimoniale. Le fait générateur de cette protection reposerait dans la publication, dite aussi divulgation<sup>109</sup> ou, dans le texte de la directive « durée », communication au public (faite licitement)<sup>110</sup>. Même intégré au droit d'auteur – parce que portant sur une œuvre –, le monopole accordé au publicateur s'analyse, selon la doctrine, comme un « droit voisin »<sup>111</sup> spécifique, assis sur le support de l'œuvre. Un tel monopole spécial pourrait ainsi être accordé à celui qui endosse la responsabilité de la communication au public d'une création générée par une IA.

Il ne s'agirait cependant pas ici de faire naître un monopole sur de l'information brute ou des idées, sans autre condition. On parle bien ici d'un droit d'auteur, portant sur une œuvre. Aussi, la forme générée devrait être **intrinsèquement originale** ; elle devrait, en quelque sorte, passer le test de Turing<sup>112</sup> et pouvoir tromper sur son origine. Il s'agirait d'une création *assimilable* à une œuvre. Il est renvoyé ici à la démonstration opérée sur l'adaptation des critères du droit d'auteur.

Par ailleurs, il serait aussi possible, si la volonté politique était en ce sens, d'ajouter à ce cadre une condition d'**implication** ou d'**investissement** (matériel, humain ou financier). Plus largement, on pourrait ici s'inspirer du droit anglais pour désigner le titulaire.

La disposition pourrait alors être intégrée à l'article L. 123-4<sup>113</sup> du CPI, *in fine*, en ces termes : « Celui qui prend les dispositions nécessaires en termes d'investissement pour communiquer au public une création de forme générée par une intelligence artificielle et assimilable à une œuvre de l'esprit jouit d'un droit d'exploitation d'une durée de X<sup>114</sup> années à compter de la communication ». Pour permettre plus de souplesse, on pourrait ici privilégier les solutions contractuelles négociées en ajoutant en début de disposition « sauf stipulations contraires »<sup>115</sup>.

---

<sup>107</sup> Faisant cette proposition, N. Enser, *op. cit.*, n° 613.

<sup>108</sup> CPI, art. L. 123-4, al. 3 : « Si la divulgation est effectuée à l'expiration de cette période, il appartient aux propriétaires, par succession ou à d'autres titres, de l'œuvre, qui effectuent ou font effectuer la publication. »

<sup>109</sup> Le terme divulgation comme celui de publication devront ici être entendus au sens large de communication au public.

<sup>110</sup> Dir. 2006/116/CE du 12 déc. 2006 relative à la durée de protection du droit d'auteur et de certains droits voisins (version codifiée), art. 4.

<sup>111</sup> V. not. A. Lucas, A. Lucas-Schloetter et C. Bernault, *Traité de la propriété littéraire et artistique*, *op. cit.*, n° 673 ; P.-Y. Gautier, *Propriété Littéraire et Artistique*, PUF, coll. Droit fondamental, 11<sup>e</sup> éd., 2019, n° 401 ; F. Pollaud-Dulian, *Le droit d'auteur*, Economica, 2<sup>e</sup> éd., 2014, n° 2508 ; C. Caron, *Droit d'auteur et droits voisins*, LexisNexis, 5<sup>e</sup> éd., 2017, n° 351 ; M. Vivant et J.-M. Bruguière, *op. cit.*, n° 978.

<sup>112</sup> A. Turing, « *Computing machinery and intelligence* », *Mind*, Oxford University Press, Volume LIX, Issue 36, Oct. 1950, p. 433, <https://academic.oup.com/mind/article/LIX/236/433/986238>.

<sup>113</sup> Ce rattachement à la partie droit d'auteur permettrait de filer la logique du droit sur les œuvres posthumes. Cette option permet de rappeler que ces créations sont assimilables à des œuvres.

<sup>114</sup> Le choix d'une durée est un choix politique qu'il appartiendrait alors au législateur de trancher. Par comparaison, le nouveau droit voisin des éditeurs de presse a une durée de 2 ans à compter de la publication. La version initiale de 2016 proposait une durée de 20 ans.

<sup>115</sup> V. A. Bensamoun, « Intelligence artificielle et propriété intellectuelle », *op. cit.*

### 2.2.3. La création d'un droit *sui generis*

Le droit privatif sur les productions créatives d'une IA pourrait ensuite consister en un **droit *sui generis***, sur le modèle du droit accordé au producteur de bases de données<sup>116</sup>, avec un objectif affiché, celui de soutenir et de protéger l'investissement<sup>117</sup>.

La protection de la base de données par ce droit spécial de propriété littéraire et artistique est conditionnée à la réalisation d'un investissement<sup>118</sup>, financier, matériel ou humain, pour l'obtention, la constitution, la vérification ou la présentation du contenu de la base, à l'exclusion des investissements pour la création des données<sup>119</sup>. Le monopole permet alors au producteur d'interdire l'extraction et/ou la réutilisation de la totalité ou d'une partie substantielle de la base de données. La notion de producteur est connue du droit de la propriété littéraire et artistique<sup>120</sup>. Suivant une logique similaire, l'article L. 341-1 du Code de la propriété intellectuelle vise « la personne qui prend l'initiative et le risque des investissements correspondants ».

Appliqué à l'IA, ce droit *sui generis* permettrait à la personne qui prend le risque un retour sur investissement, déjouant les tentatives d'appropriation de valeur et encourageant du même coup l'investissement dans le domaine. La proposition pourrait être intégrée dans la première partie du Code de la propriété intellectuelle, consacrée à la propriété littéraire et artistique, à la suite du droit des bases de données précédemment décrit<sup>121</sup>, et imposer la protection suivante<sup>122</sup> : « Le producteur d'une intelligence artificielle permettant la génération de créations assimilables à des œuvres<sup>123</sup> de l'esprit bénéficie d'une protection sur ces créations lorsque celles-ci résultent d'un investissement financier, matériel ou humain substantiel ».

On pourrait encore décider de limiter le droit d'exploitation attribué au producteur – droit de reproduction et droit de communication au public – aux utilisations à titre lucratif<sup>124</sup>, afin de réduire le champ de la réservation aux comportements purement parasites. Certes, la prise en compte de l'usage commercial dans la délimitation du champ du monopole est ignorée du droit français. Mais le droit européen n'y est pas insensible<sup>125</sup>. Certains auteurs insistent

---

<sup>116</sup> V. dir. 96/9/CE, du 11 mars 1996, concernant la protection juridique des bases de données ; CPI, art. L. 341-1 et s.

<sup>117</sup> A. Bensamoun et G. Loiseau, « L'intégration de l'intelligence artificielle dans certains droits spéciaux », *Dalloz IP/IT*, mai 2017, p. 294, spéc. p. 297.

<sup>118</sup> Sur la notion, v. P. Gaudrat et F. Sardain, *Traité de droit civil du numérique*, t. 1, *Droit des biens*, Larcier, 2015, n° 1085 et s.

<sup>119</sup> CJCE, 9 nov. 2004, 4 arrêts, aff. C-203/02, *The British Horseracing Board Ltd et a. c/ William Hill Organization Ltd*, Rec. CJCE 2004, I, p. 10461 ; Comm. comm. électr. 2005, comm. 2, note C. Caron ; Propr. industr. 2005, comm. 7, 1<sup>er</sup> arrêt ; Propr. industr. 2005, comm. 22, 2<sup>e</sup> arrêt. CJCE, 9 nov. 2004, aff. C-46/02, *Fixtures Marketing Ltd c/ Oy Veikkaus AB*, Rec. CJCE 2004, I, p. 10396. CJCE, 9 nov. 2004, aff. C-338/02, *Fixtures Marketing Ltd c/ Svenska Spel AB*, Rec. CJCE 2004, I, p. 10532, pt 29. CJCE, 9 nov. 2004, aff. C-444/02, *Fixtures Marketing Ltd c/ Organismos prognostikon agonon podofairou AE (OPAP)*, Rec. CJCE 2004, I, p. 10590. Sur ces arrêts, v. not. S. Lemarchand et S. Rambaud, *Propr. intell.* 2005, p. 99 ; F. Pollaud-Dulian, *RTD com.* 2005, p. 90 ; M. Vivant, *RLDI*, mars 2005, n° 104.

<sup>120</sup> V. les producteurs de phonogrammes (CPI, art. L. 213-1) et de vidéogrammes (CPI, art. L. 215-1).

<sup>121</sup> Peut-être dans un nouveau titre V du livre III ?

<sup>122</sup> V. A. Bensamoun, « Intelligence artificielle et propriété intellectuelle », *op. cit.*

<sup>123</sup> Là encore, la référence à l'œuvre devra se faire de manière objective.

<sup>124</sup> Sur une problématique proche, v. J. Lapousterle et A. Latil (dir.), *L'usage commercial des biens intellectuels*, Mare et Martin, PUS, 2020, à paraître.

<sup>125</sup> V. par ex. la jurisprudence sur les liens hypertextes. – V. généralement J. Lapousterle, « L'usage commercial en droit d'auteur : indifférence véritable ou déni ? », in *L'usage commercial des biens intellectuels*, *op. cit.*

d'ailleurs sur la nécessité de prendre en compte la fonction du droit de propriété intellectuelle pour en déterminer le périmètre<sup>126</sup>.

Une telle option aurait, d'une part, le mérite de la simplicité et, d'autre part, celui de la cohérence. Elle ne risquerait pas de fragiliser le droit d'auteur en cherchant à en étendre les catégories juridiques. Elle permettrait l'adaptation d'un régime aux spécificités de la création générée par ordinateur et aux objectifs d'intérêt général. Elle n'empêcherait pas de qualifier telle ou telle création d'œuvre de l'esprit, lorsque les critères traditionnels sont remplis – notamment dans des cas s'apparentant surtout à de la création *assistée*. Elle permettrait, enfin, de préserver les incitations à l'investissement voulues dans le domaine de l'IA en octroyant une protection limitée aux droits patrimoniaux, et pour une durée plus courte que celle que le droit d'auteur prévoit. La solution pourrait donc être une voie intéressante, mais seulement si le besoin était avéré et que l'avenir révélait que le droit d'auteur ne peut s'appliquer. D'autant que la voie d'un droit *sui generis* présente l'inconvénient de nier la parenté évidente entre les créations traditionnelles et les créations intelligentes.

De manière générale, si un nouveau droit devait être créé, il faudrait s'interroger sur son périmètre exact. En particulier, il serait peut-être pertinent, pour préserver la valeur de la créativité humaine, d'accorder des droits moins étendus que ceux qu'offre le droit d'auteur classique. Par ailleurs, une attention devrait être portée aux droits des artistes-interprètes car leur monopole est assis sur l'exécution d'une œuvre de l'esprit.

### 2.3. *L'absence de droit privatif*

Avant toute chose, il convient ici de rappeler que dès lors qu'un humain intervient de manière suffisamment créative, l'IA doit être appréhendée comme un **outil**<sup>127</sup>, peu important la sophistication ou la technicité de ce dernier.

Cette hypothèse mise à part, des auteurs de doctrine doutent de la nécessité d'une protection des créations générées par une IA, constatant notamment l'**existence de protections en amont** de la chaîne (v. *supra*). Certains appuient en outre leur position par le fait que la valeur culturelle se trouve plus dans l'**algorithme « éduqué »** que dans la création qu'il génère. L'acte créatif est davantage dans la programmation de l'IA que dans son fonctionnement. Selon eux, le résultat de l'IA est le pur produit du hasard ou d'une planification mécanique : il n'exprime donc que très peu une quelconque créativité propre à la personne humaine qui mériterait la protection du droit d'auteur. Or, le but même de la protection accordée par le droit d'auteur est de favoriser la « communication humaine »<sup>128</sup>.

Par ailleurs, le programmeur de l'IA étant rémunéré par la vente de l'algorithme ou par une **licence** d'utilisation du logiciel, le droit d'auteur n'aurait donc pas à être sollicité pour protéger les créations de l'IA, au risque d'être dénaturé.

---

<sup>126</sup> M. Vivant, « *Intellectual property rights and their functions: determining their legitimate 'enclosure'* », in H. Ullrich, P. Drahos et G. Ghidini (dir.), *Kritika: Essays on intellectual property*, vol. 3, Edward Elgar Pub., 2018, p. 44.

<sup>127</sup> On pense par exemple à *Daddy's car*.

<sup>128</sup> V. S. Ricketson, « *People or Machines: The Bern Convention and the Changing Concept of Authorship* », 16 Colum.-VLA J.L & Arts 1 (1991-1992) ; C. J. Craig and I. R. Kerr, « *The Death of the AI Author* », mars 2019 : <https://ssrn.com/abstract=3374951> ou <http://dx.doi.org/10.2139/ssrn.3374951>.

Dès lors, la loi assumerait une absence de protection immatérielle. Seul resterait, le cas échéant, un simple droit de propriété corporelle sur le bien matériel produit par l'IA, qui serait probablement attribué à celui qui détient le titre d'utilisation de l'IA elle-même. Ce régime aurait également vocation à être complété par un volet concurrentiel permettant d'éviter tout détournement du domaine public. La théorie de l'enrichissement sans cause pourrait également être convoquée afin de lutter contre l'utilisation illicite de créations non protégées<sup>129</sup>, quoique la jurisprudence sur ce sujet ne soit pas très fournie.

Le groupe français de l'AIPPI<sup>130</sup> rejette pour sa part l'idée d'un monopole considérant que les conditions ne sont tout simplement pas remplies en l'**absence d'auteur humain**. Cette constatation interdit la mise en place de pareille protection qui contreviendrait alors aux fondements du droit d'auteur. En outre, le risque de **massification** de la création et de saturation du champ créatif fragiliserait la position de la personne physique créatrice.

D'autres commentateurs encore avancent des arguments éprouvés, comme l'**atteinte à la liberté d'expression** que constitue toute exclusivité<sup>131</sup>. Par ailleurs, à l'heure où toute réservation par la propriété intellectuelle est remise en cause dans l'opinion publique, ces auteurs considèrent qu'il est à craindre que la mise en place d'un nouveau droit soit analysée comme un excès de protection.

D'autres encore privilégient la souplesse et soutiennent la voie du **règlement contractuel** entre le concepteur de l'IA et l'utilisateur ou la protection par la technique, par l'apposition éventuelle de **mesures techniques**. Cette dernière voie impliquerait de modifier le droit puisque pour l'heure, le contournement est sanctionné pour autant que la mesure vise à protéger un objet éligible au droit d'auteur ou aux droits voisins (CPI, art. L. 331-5).

Par ailleurs, l'absence de droit privatif ne signifie pas nécessairement absence de protection. Le concepteur de l'IA peut ainsi protéger le processus créatif par le mécanisme du **secret des affaires** (C. comm., art. L. 151-1, transposant la directive 2016/943 du 8 juin 2016), ce qui lui permettra, en prenant des mesures matérielles et juridiques raisonnables, de s'opposer à la divulgation illicite de son savoir-faire et à l'appropriation de son investissement. Il pourra encore, sur le fondement du droit commun de la responsabilité civile extracontractuelle (C. civ., art. 1240 s.), agir contre les actes de **concurrence déloyale** ou les **agissements parasitaires**.

Enfin, selon une analyse doctrinale, ces œuvres sans auteur devraient être considérées comme des communs par nature – des **commons by design** –, intégrés dans un domaine public par défaut<sup>132</sup>. Cette qualification extensive du domaine public serait soumise à deux conditions : que l'IA joue un rôle prépondérant dans l'expression de l'œuvre et que l'intervention humaine

---

<sup>129</sup> A. Lebois, « Quelle protection juridique pour les créations des robots journalistes ? », CCE 2015, Etude 2.

<sup>130</sup> V. le rapport du groupe français AIPPI (Association Internationale pour la Protection de la Propriété Intellectuelle, 2019 – *Question Study – Copyright in artificially generated works* : [https://aippi.org/wp-content/uploads/2019/06/2019\\_FR\\_2019\\_Study\\_Question\\_Copyright\\_in\\_artificially\\_generated\\_works\\_2019-05-24.pdf](https://aippi.org/wp-content/uploads/2019/06/2019_FR_2019_Study_Question_Copyright_in_artificially_generated_works_2019-05-24.pdf).

<sup>131</sup> En ce sens, not. J. Cabay, « Droit d'auteur et intelligence artificielle : comparaison n'est pas raison », févr. 2019, <https://difusion.ulb.ac.be/vufind/Record/ULB-DIPOT:oai:dipot.ulb.ac.be:2013/285602/Holdings>.

<sup>132</sup> V. not. G. Azzaria, « IA et création non appropriable », <https://www.lesconferences.ca/videos/mise-en-commun-numerique/> et « Intelligence artificielle et droit d'auteur : l'hypothèse d'un domaine public par défaut », *Les Cahiers de PI*, 2018, vol. 30, n° 3, p. 925. Comp. D. Gervais, « *Can Machines be Authors* », <http://copyrightblog.kluweriplaw.com/2019/05/21/can-machines-be-authors/>, mai 2019 ; « La machine en tant qu'auteur », *Propri. intell.* 2019, n° 72, p. 7.

soit en conséquence résiduelle, ne permettant pas d'imprimer des choix créatifs, lesquels sont alors le résultat de l'IA.

Un autre auteur propose de fonder l'intégration au domaine public, non sur l'absence d'auteur, mais sur des **raisons écologiques** (coût écologique de ces productions de masse) et sur le risque d'**appauvrissement culturel** lié à ces productions. Ce qui justifierait de ne préserver, dans un choix de politique législative, qu'une valorisation contractuelle<sup>133</sup>.

Mais cette voie, qui a le mérite de la simplicité en choisissant de soustraire au droit privatif les créations purement générées par l'IA, pose d'autres problèmes. Ainsi qu'on l'avait signalé, elle **laisse sans protection les candidats à la titularité de droits voisins**, précisément les interprètes des créations artificielles. Pour certains, elle favoriserait également les effets de substitution et de parasitisme, multipliant les risques de contentieux, et risquerait de limiter l'incitation à l'investissement dans le domaine de l'IA<sup>134</sup>. Ces aspects mériteraient cependant une validation économique empirique pour être étayés. De plus, la voie ne résout pas la question de la distinction entre les créations « assistées » par l'ordinateur, qui seraient éligibles au droit d'auteur traditionnel, et les créations « générées », qui tomberaient dans le domaine public – quoique cette distinction puisse être laissée à la libre appréciation des juges du fond. Enfin, elle serait, dans les faits, assez impraticable et favoriserait les **comportements de détournement** : l'utilisateur de l'IA n'aurait aucune incitation à révéler que son « œuvre » est générée par ordinateur et qu'elle tombe dans le domaine public.

#### 2.4. *Quelle solution ?*

Le choix d'un statut des productions culturelles de l'IA apparaît comme un sujet presque **philosophique**. Il est aussi clairement question de **politique législative**. Le dernier mot reviendra en effet au législateur. Et avant tout, il devra se prononcer sur la pertinence ou non d'une protection par un droit privatif des réalisations de l'IA. Aujourd'hui, il semblerait que le règlement de la question s'opère par la **voie contractuelle**. Précisément, lorsqu'une IA est mise à disposition pour créer, deux formules sont proposées<sup>135</sup> : une utilisation gratuite pour une exploitation privée des résultats créatifs et une utilisation payante pour une exploitation commerciale de ceux-ci. Aussi une intervention législative nécessiterait-elle de mettre en exergue le besoin de créer un droit privatif, en évaluant ces solutions contractuelles et par une étude d'impact.

Il importe ensuite de rappeler que le droit d'auteur doit rester rattaché à un **humain**<sup>136</sup>. La mise en œuvre du droit d'auteur ne peut se passer de la présence humaine (même si le lien à la création pourrait être plus distendu). En outre, dès que l'IA est un outil entre les mains d'un créateur personne physique – qui choisit les intrants de la base d'entraînement en amont, réalise la curation en aval... –, le droit d'auteur classique a vocation à s'appliquer. C'est aujourd'hui la situation la plus courante (celle de *Daddy's car* ou d'*Edmond de Belamy*). Il s'agit alors juste d'évaluer l'implication de la personne physique qui crée, soit en amont en

---

<sup>133</sup> En ce sens, A. Billion, audition CSPLA.

<sup>134</sup> Contestant cette idée d'effet incitatif : J. Cabay, « Contre la protection du droit d'auteur pour les productions de l'intelligence artificielle », nov. 2018, <https://orbi.uliege.be/handle/2268/232795> ; <http://hdl.handle.net/2268/232795>.

<sup>135</sup> C'est le cadre notamment choisi par Aiva Technologies.

<sup>136</sup> V. en ce sens la résolution 2019 – *Study Question – Copyright in artificially generated works*, sept. 2019 : [https://aippi.org/wp-content/uploads/2019/10/Resolution\\_Copyright\\_in\\_artificially\\_generated\\_works\\_English.pdf](https://aippi.org/wp-content/uploads/2019/10/Resolution_Copyright_in_artificially_generated_works_English.pdf).

programmant et en entraînant la machine, soit en aval en « contrariant » la production par des actions en aval<sup>137</sup>.

Cependant, on ne peut ignorer que la technique est en développement et que l'outil gagne en autonomie, en réduisant le rôle de l'humain. Cette évolution en cours pourrait d'ailleurs justifier que le législateur ne se précipite pas pour intervenir<sup>138</sup>. Il serait alors urgent d'attendre...

En ce sens, il importe de relever que **le droit d'auteur semble suffisamment souple pour recevoir ces créations** et que l'attribution des droits au concepteur de l'IA semble de nature à apporter des solutions pertinentes. C'est ce que semble avoir décidé une **cour chinoise**, au bénéfice de l'entreprise Tencent<sup>139</sup>. Aussi, sauf à ce que la jurisprudence mette en exergue un besoin nouveau ou refuse de recevoir ces réalisations, **le droit positif devrait pour l'heure pouvoir être appliqué, dans une lecture renouvelée des critères d'accès à la protection.**

Si, à l'avenir, une insuffisance du droit positif et un besoin de protection étaient rapportés, une intervention législative serait alors nécessaire. Sur ce point, il est probable que la première réglementation ou décision régionale pourrait influencer les autres cadres normatifs<sup>140</sup>. Pour quel contenu et à quelle échelle ?

Une voie intéressante pourrait alors être, au vu des différentes positions et analyses, celle de la création d'un **droit spécial du droit d'auteur** (pour manifester la parenté évidente de ces créations avec les œuvres classiques), assis sur les critères classiques dans une lecture renouvelée. En effet, prenant en compte le fait que le lien à l'auteur, au sens classique du terme, apparaît plus distendu et que le contenu de la protection peut s'avérer inadapté (sur le terrain du droit moral notamment), le droit spécial pourrait permettre une protection ajustée.

Dans tous les cas, un point d'attention devrait être porté au sujet du droit des artistes-interprètes, pour s'assurer que leur sort serait réglé lorsqu'ils interprètent pareilles réalisations.

Quoi qu'il en soit, même si une solution nationale est possible – à considérer que le champ est non-harmonisé –, il serait souhaitable que la solution retenue s'impose à une **échelle internationale**<sup>141</sup>, *a minima* **européenne**. Sur ce point, une protection par le droit d'auteur permettrait de profiter des outils normatifs déjà en vigueur, notamment de la Convention de Berne, ce qui aurait pour intérêt d'appliquer la règle du traitement national aux créations intelligentes.

Le législateur européen n'ignore pas que le sujet impose parfois son intervention. En ce sens, la question du régime des œuvres entrantes, qui permettent la fabrication du modèle d'inférence, a été (imparfaitement) traitée dans la directive « *Digital Single Market* », grâce à l'exception « fouille de données » dédiée aux usages de l'IA (v. *infra* 3.2.1.2). S'il est

---

<sup>137</sup> V. J. C. Ginsburg et L. A. Budiardjo, « Authors and machine », Columbia Public Law Research Paper No. 14-597 ; Berkeley Technology Law Journal, Vol. 34, No. 2, 2019.

Disponible à SSRN : <https://ssrn.com/abstract=3233885> ou <http://dx.doi.org/10.2139/ssrn.3233885>. Mais les auteurs considèrent que pour être auteur, la personne doit avoir la maîtrise de la conception et de l'exécution de l'*output*.

<sup>138</sup> Considérant que la technique est encore en développement, l'AIPPI considère dans sa résolution qu'il serait prématuré de se prononcer sur la pertinence d'une protection nouvelle.

<sup>139</sup> <https://venturebeat.com/2020/01/10/chinese-court-rules-ai-written-article-is-protected-by-copyright/>.

<sup>140</sup> Nous n'ignorons pas que le traitement national de la Convention de Berne implique que même si les créations produites par une IA étaient protégées aux EU, elles ne le seraient pas nécessairement en Europe.

<sup>141</sup> V. dans le même sens les conclusions du *Summary Report* de l'AIPPI, 2019 – *Question Study – Copyright in artificially generated works*, juill. 2019 : « A clear majority of the responding Groups considers that harmonization regarding the protection of artificially-generated works is desirable » : [https://aippi.org/wp-content/uploads/2019/07/SummaryReport\\_COPYRIGHT-DATA\\_London2019\\_final\\_160719.pdf](https://aippi.org/wp-content/uploads/2019/07/SummaryReport_COPYRIGHT-DATA_London2019_final_160719.pdf).

intervenu sur l'amont, le législateur devrait être à même, si le besoin était avéré, de réglementer l'aval.

D'évidence, le sujet intéresse et la question est en mouvement. Aussi, et comme il l'a été évoqué par le président du CSPLA en début de mission<sup>142</sup>, la réflexion devrait être suivie par l'institution. Il s'agit là d'un premier point d'étape qui pourrait mériter que le CSPLA réévalue la question – en constante évolution – dans les prochaines années.

---

<sup>142</sup> Le président était alors P.-F. Racine.

### Partie 3. – Le régime des œuvres permettant la production de créations culturelles par l'IA

Le principe même d'apprentissage machine repose, on l'a vu, sur une utilisation massive de quantités de données, préalablement sélectionnées ou non. L'entraînement de la machine porte sur sa façon d'appréhender ces données, de les calculer et d'en extraire un résultat. Cette extraction suppose une analyse des données présentes dans la machine, afin d'en identifier les caractéristiques, y compris esthétiques.

Doit-on considérer que la décomposition ou l'analyse des œuvres ingérées<sup>143</sup>, leur permettant d'être représentées dans un espace intelligible par la machine et de révéler les liens ou les similarités par recombinaison, est un **acte d'exploitation donnant prise au droit d'auteur** ? Comment comprendre la **nouvelle exception de fouille de données** permise par l'article 4 de la directive 2019/790 et ouverte largement aux acteurs de l'IA, mais qui préserve également une possibilité d'*opt-out* technique par le titulaire de droits ? D'évidence, cette limitation, adoptée sous la pression des acteurs du numérique pour permettre une utilisation facilitée de l'IA, aboutit à un compromis fort singulier. D'abord, ce texte, bien que posant une limitation, constitue indirectement la reconnaissance d'un monopole étendu. Ensuite, il permet de déroger à l'exclusivité de l'auteur. Enfin, il autorise à sortir de l'exception pour retourner à une réservation. Nul doute que l'articulation ne se fera pas sans heurts.

Aussi, le droit d'auteur s'applique-t-il au phénomène d'analyse, d'utilisation et de décomposition d'œuvres protégées (3.1) ? Comment trouver un équilibre entre les objectifs d'innovation liés à l'IA et le besoin de protection légitime des titulaires de droits (3.2) ?

#### *3.1. Le droit d'auteur s'applique-t-il aux œuvres dont se nourrit l'IA créative ?*

On le sait, le monopole d'exploitation appartenant à l'auteur est particulièrement large en droit français, où le législateur a adopté une approche synthétique et ouverte, en la forme de deux prérogatives catégorielles qui intègrent l'ensemble des actes d'exploitation de l'œuvre : le droit de représentation (ou droit de communication au public) et le droit de reproduction. Mais les œuvres « ingérées » par l'IA ne sont pas classiquement exploitées : elles sont déconstruites, décomposées, afin d'élaborer un modèle reposant sur les spécificités communes de celles-ci. Elles sont ainsi utilisées non pour elles-mêmes, mais pour leur **valeur informationnelle**. L'œuvre est, en quelque sorte, « **data-ifiée** ». Ce que l'IA utilise dans son processus créatif est moins l'œuvre primaire que les traits caractéristiques qu'elle a identifiés. Le processus ultérieur de recombinaison ne portera donc pas tant sur l'œuvre elle-même que sur le brassage et l'interaction des données que l'IA aura extraites. Doit-on voir ici une exploitation des œuvres entrantes ?

Selon certains, chaque ingestion par l'intelligence artificielle consisterait en une « **appropriation** » de l'œuvre<sup>144</sup>, susceptible de violer le monopole d'exploitation. Il se pourrait même que l'auteur des œuvres brassées par l'IA puisse influencer le sort de l'œuvre générée. A suivre ce raisonnement, il serait possible de retrouver, dans l'œuvre générée par la

<sup>143</sup> Il s'agit par ex. de passer d'une image à un ensemble de valeurs d'intensité ou d'une séquence sonore à un ensemble de fréquences.

<sup>144</sup> T. Lebrun, « L'apprentissage machine est une appropriation », Vol. 30, n° 3, Les Cahiers de propriété intellectuelle, 2018, p. 895.

machine, certains traits caractéristiques des œuvres primaires extraites au cours du processus de décomposition et de recombinaison<sup>145</sup>. Dans ce cas, la création générée par l'IA serait, pour partie au moins, analysée comme une œuvre dérivée des œuvres traitées par l'intelligence artificielle lors de son processus créatif.

Cependant, il sera en réalité rare que la création finale reflète des composantes essentielles et reconnaissables de l'œuvre initiale, ce critère étant déterminant pour qualifier une œuvre d'œuvre dérivée<sup>146</sup>. Certes, on pourrait se demander si l'appréciation des juges du fond ne pourrait pas s'adapter aux spécificités du processus de création par IA afin de fixer un seuil de reconnaissance suffisamment bas pour permettre cette qualification d'œuvre dérivée : plusieurs décisions, sans nécessairement concerner la création par IA, ont déjà marqué une tendance à l'appréciation laxiste du critère de reconnaissance de l'œuvre originale<sup>147</sup>. Au Canada, un artiste a récemment attaqué en justice un autre artiste ayant utilisé une IA traitant les œuvres issues du premier pour produire de nouvelles œuvres (non commercialisées). L'affaire est pendante<sup>148</sup> et il sera intéressant de voir comment les similitudes seront appréciées.

Mais la voie semble peu convaincante. Il sera délicat d'estimer, dans le cas des œuvres générées par IA, que le processus de décomposition puis recombinaison puisse être considéré comme une représentation de l'œuvre primaire au public. En effet, l'œuvre primaire sera la plupart du temps rendue méconnaissable par le processus créatif de l'IA (il est même parfois ajouté dans le dispositif intelligent un « logiciel anti-plagiat ») ; il est alors impossible de considérer qu'elle est communiquée au public. Aussi, le critère essentiel pour savoir si le droit d'auteur et ses exclusions ont vocation à s'appliquer pourrait être de déterminer si l'œuvre contient ou non des « **éléments matériels détectables** » d'une œuvre préalable, et non pas de savoir si cette œuvre préalable a constitué une source d'inspiration « immatérielle » pour produire la création subséquente<sup>149</sup>.

S'il apparaît douteux que le droit de communication au public est mis en œuvre, en revanche, il appert que l'œuvre digérée par l'IA est bien souvent, en amont, **reproduite**. Cette reproduction constitue-t-elle un acte d'exploitation soumise au consentement de l'auteur ?

Certains le réfutent, considérant que la reproduction est ici une étape **technique** et non un acte d'exploitation autonome. Or, copie technique et copie juridique doivent être distinguées, la première ne caractérisant pas nécessairement la seconde<sup>150</sup>, même si le législateur a plutôt

---

<sup>145</sup> V. notamment F. Macrez et J.-M. Deltorn, « *Authorship in the age of machine learning and artificial intelligence* », *op. cit.*

<sup>146</sup> L'accord de l'auteur de l'œuvre première n'est sollicité que si les éléments retrouvés sont bien empruntés à l'œuvre première. Tel n'est pas le cas lorsque les ressemblances entre deux chansons peuvent résulter de la simplicité de la mélodie et de leur inspiration commune (rythme antillais) : CA Paris, 25 avr. 1972, RIDA juill. 1972, p. 221.

<sup>147</sup> V. notamment *Bridgeport Music, Inc v. Dimension Films*, 410 F. 3d 792 (6th Cir. 2005) ; Bundesgerichtshof, *Kraftwerk et al. v. Moses Pelham, et al.*, No. I ZR 112/06, 20 nov. 2008 ; TGI de Paris, 5 juill. 2000, CCE mars 2001, comm. n° 23, obs. C. Caron.

<sup>148</sup> Affaire *Chamandy c./ Basanta* (2018) : <https://www.theglobeandmail.com/arts/art-and-architecture/article-artist-faces-lawsuit-over-computer-system-that-creates-randomly/>.

<sup>149</sup> J.-M. Deltorn, « "In the style of..." - deep learning, style transfer and the limits of copyright protection. A European perspective », in *Artificial intelligence and intellectual property, Annali italiani del diritto d'autore, della cultura e dello spettacolo (AIDA)*, vol. XXVII, mai 2019, p. 337.

<sup>150</sup> A. Lucas, A. Lucas-Schloetter et C. Bernault, *Traité de la propriété littéraire et artistique*, LexisNexis, 5<sup>e</sup> éd., 2017, n° 263, au sujet des reproductions techniques provisoires. V. aussi S. Dusollier, « L'exploitation des œuvres : une notion centrale en droit d'auteur », in *Mélanges A. Lucas*, LexisNexis, 2014, p. 263, spéc. p. 267.

tendance à consacrer des exceptions en ces endroits<sup>151</sup>. En matière de création par une IA, la reproduction ne vise qu'à extraire le contenu informationnel de l'œuvre « ingérée »<sup>152</sup>. Et il n'est pas assuré que le droit exclusif ait vocation à saisir cette valeur informationnelle de l'œuvre. Cependant, il faut convenir que la loi n'opère pas de telles distinctions. L'article 2 de la directive 2001/29 reconnaît à l'auteur une **prérogative particulièrement large**, formulée comme « le droit exclusif d'autoriser ou d'interdire la reproduction directe ou indirecte, provisoire ou permanente, par quelque moyen et sous quelque forme que ce soit, en tout ou en partie ». Dans ces conditions, l'approche fonctionnelle, qui a pu être retenue en droit des marques<sup>153</sup>, ne semble pas opérante en droit d'auteur<sup>154</sup>.

Néanmoins, une décision récente de la **CJUE** pourrait également influencer le débat s'agissant de la mise en œuvre du droit de reproduction. Dans un arrêt *Pelham* du 29 juillet 2019<sup>155</sup>, au sujet de l'échantillonnage musical (*sampling*), la grande chambre de la Cour a décidé qu'un échantillon non reconnaissable n'est pas une reproduction – en réalité n'est pas couvert par le droit exclusif. Précisément, après avoir reconnu, interprétant l'article 2.c) de la directive 2001/29, que « la reproduction par un utilisateur d'un échantillon sonore, même très bref, d'un phonogramme doit, en principe, être considérée comme une reproduction “en partie” de ce phonogramme » (pt 29), elle ajoute, modifiant la solution attendue, qu'il n'y a pas de reproduction si l'échantillon est utilisé, pour créer une nouvelle œuvre, « sous une forme modifiée et non reconnaissable à l'écoute » (pt 31). Certes, la décision porte non sur le droit d'auteur mais sur le droit voisin des producteurs. Et la solution mérite sans aucun doute d'être discutée. Cela étant, transposée aux créations générées (et même assistées) par une IA, elle pourrait conduire à considérer l'exclusion du champ de la protection pour les reproductions amont des œuvres (comme données d'entraînement) dans l'IA, dans la mesure où ces œuvres ne sont en principe **pas reconnaissables dans la création finale**.

Cette analyse peut sembler, à première vue, en contradiction avec la nouvelle exception *text and data mining* de la directive 2019/790. En effet, pourquoi faire une exception si l'acte ne donne pas prise au monopole d'exploitation ?

Mais il est possible que la contradiction ne soit qu'apparente, même s'il est difficile de tirer des conclusions certaines de la décision *Pelham*. En effet, si l'arrêt tranche la question du droit de reproduction, il donne l'impression que c'est par renvoi au droit de communication au public puisque l'œuvre n'est pas reconnaissable **à l'écoute** (ce qui constitue une communication). Or l'acte de reproduction est autonome de celui de communication au public. Dans ces conditions, l'échantillonnage devrait être considéré comme une hypothèse particulière, mêlant les deux actes d'exploitation et qui diffère donc, par nature, de la fouille de données, laquelle ne vise que l'acte de reproduction.

---

<sup>151</sup> V. par ex. l'article 5.1 de la directive 2001/29 sur la copie technique provisoire qui permet de fluidifier la circulation des réseaux.

<sup>152</sup> C. Bernault, *Open access et droit d'auteur*, Larcier, 2016, n° 185-187. – P. Kamocki, « “Laissez fouiller!” – L'argument pour les “utilisations orthogonales” des œuvres de l'esprit dans le contexte du débat sur l'exploitation des données », RIDA janv. 2016, p. 5-85, spéc. p.55.

<sup>153</sup> J. Azéma et J.-C. Galloux, *Droit de la propriété industrielle*, Dalloz, 7<sup>e</sup> éd., 2012, n° 1434-1435.

<sup>154</sup> A. Lucas et J. Ginsburg, « Droit d'auteur, liberté d'expression et libre accès à l'information (Etude comparée de droit américain et européen) », RIDA juill. 2016, p. 5-153, spéc. p. 107-109. – Sur les différentes fonctions qui fondent la propriété intellectuelle : W. Fisher, « *Theories of intellectual property* », <https://cyber.harvard.edu/people/ffisher/iptheory.pdf>.

<sup>155</sup> CJUE, gde ch., 29 juill. 2019, aff. C-476/17, *Pelham*.

### **3.2. Le difficile équilibre entre les objectifs d'innovation par l'IA et de protection légitime des auteurs**

A considérer que l'acte amont d'ingestion par l'IA donne prise au monopole, le consentement de chaque titulaire de droits sur les œuvres versées dans le processus d'apprentissage puis de création de l'IA devrait être récolté (moyennant rémunération, le cas échéant). La création par IA en serait très certainement freinée.

Il convient dès lors d'envisager si des exceptions peuvent jouer en pareille situation. On peut encore se demander si un système de licence permettant de mieux prendre en compte la spécificité de la création par l'IA ne pourrait pas être mis en place.

#### **3.2.1. Les possibles exceptions au monopole conféré par le droit d'auteur**

Les exceptions en vigueur (directive 2001/29, art. 5) sont impropres à appréhender la situation (3.2.1.1). Raison pour laquelle l'UE s'y est intéressée, en créant une nouvelle exception « fouille de données » qui vise clairement l'hypothèse de l'ingestion d'œuvres (3.2.1.2).

##### **3.2.1.1. L'inapplicabilité des exceptions actuellement en vigueur**

Trois exceptions ont pu être évoquées afin de limiter le monopole d'exploitation dans le cas de la création générée par IA, mais aucune ne semble pouvoir, en l'état, apporter une solution satisfaisante à l'hypothèse d'ingestion d'œuvres par une IA. Notons au préalable, pour mémoire, que si la Cour de cassation a pu, à plusieurs reprises, sembler consacrer une forme d'exception qui ne figurait pas dans les textes, en exonérant les reproductions « fortuites » ou « accessoires »<sup>156</sup>, ce courant jurisprudentiel est jugé par la majorité de la doctrine comme étant incertain. En tout état de cause, la reproduction des œuvres dans le processus de création par IA ne présente, en aucune manière, un aspect fortuit (critère d'« inclusion fortuite » exigé à l'article 5.3, i) de la directive 2001/29), dès lors qu'il constitue au contraire une démarche volontaire au cœur même du processus créatif envisagé.

La première exception au monopole du droit d'auteur concerne la **copie privée**, faisant obstacle à l'interdiction, par l'auteur, des « copies ou reproductions réalisées à partir d'une source licite et strictement réservées à l'usage privé du copiste », non destinées à une utilisation collective (2° de l'article L. 122-5 du CPI). Il est clair que la reproduction des œuvres destinées à être déversées dans l'algorithme de l'IA pour leur utilisation dans le processus créatif ne saurait se réclamer de cette exception, dès lors qu'elle ne s'effectuera pas, dans la majorité des hypothèses, pour un strict usage privé du copiste.

La seconde exception empêche l'interdiction, par l'auteur, de la « **reproduction provisoire présentant un caractère transitoire ou accessoire**, lorsqu'elle est une partie intégrante et essentielle d'un **procédé technique** et qu'elle a pour unique objet de permettre l'utilisation licite de l'œuvre ou sa transmission entre tiers par la voie d'un réseau faisant appel à un intermédiaire » (6° de l'article L. 122-5 du CPI). Pour l'essentiel, cette exception pour copie

---

<sup>156</sup> V. not. Cass. 1<sup>re</sup> civ., 12 mai 2011, D. 2011. AJ 1409, obs. J. Daleau ; D. 2011. 1875, note C. Castets-Renard ; RTD com. 2011. 553, obs. F. Pollaud-Dulian ; CCE 2011, n° 62, note C. Caron ; JCP 2011. 814, note M. Vivant ; Légipresse 2011. III. 627, note Renault ; RIDA juill. 2011, p. 341, note P. Sirinelli ; RLDI 2011/72, 2371, obs. A. Bensamoun. – Cass. 1<sup>re</sup> civ., 12 juill. 2012, RIDA 4/2012, p. 547 et p. 537, obs. P. Sirinelli ; CCE 2012, comm. 91, note C. Caron ; JCP E 2012, 1627, note J.-M. Bruguière ; Propr. intell. 2012, p. 405, obs. A. Lucas ; RTD Com. 2012, p. 775, obs. F. Pollaud-Dulian ; D. 2012, p. 2071, note C. Castets-Renard.

technique, instaurée par la directive 2001/29 du 22 mai 2001, exempte d'autorisation les copies rendues nécessaires par les transmissions numériques des œuvres.

On pourrait être tenté d'opérer un rapprochement avec la situation étudiée – ce que fait d'ailleurs le considérant 9 de la directive 2019/790 – dans la mesure où la reproduction pourrait être analysée comme un processus accessoire intégré à l'opération technique et qu'elle peut tout à fait être temporaire (elle n'est plus nécessaire une fois le modèle d'inférence élaboré). On peut cependant s'interroger sur le sens exact de la condition d'utilisation licite. Notamment, la reproduction réalisée en cas de fouille n'a pas pour unique objet l'utilisation licite de l'œuvre : c'est précisément le caractère licite de son utilisation qui est en question. Aussi, même si la directive semble indiquer que certaines hypothèses seront couvertes par cette exception, on peine à les envisager et à comprendre la frontière avec la nouvelle exception dédiée.

Enfin, l'**exception à des fins de recherche** pourrait servir de justification à l'acte amont d'ingestion par l'IA, mais outre qu'il s'agit d'une exception facultative (art. 5.3, a)), elle est peu adaptée à la technologie et les licences d'utilisation (notamment offertes aux chercheurs) peuvent exclure la fouille de données<sup>157</sup>.

### 3.2.1.2. La nouvelle exception de fouille de données

La directive 2001/29 n'avait pas anticipé les difficultés que les **pratiques d'exploration (ou « fouille ») de données (*text and data mining*, TDM)** allaient susciter au regard du droit d'auteur. La question s'est néanmoins posée lors des débats parlementaires ayant précédé l'adoption de la loi pour une République numérique et au cours du lent processus institutionnel ayant accouché de la réforme de la directive sur le droit d'auteur dite « DSM »<sup>158</sup>. Cette nouvelle pratique, dont le concept apparaît en 1989 sous l'acronyme KDD (*Knowledge Discovery in Databases*), pourrait entraîner « un changement révolutionnaire des méthodes de travail des chercheurs », le CNRS comparant la fouille de données à la « création du microscope »<sup>159</sup>. De la cartographie du génome humain à la recherche pharmaceutique en passant par les humanités numériques, ce sont tous les domaines de la recherche qui sont concernés<sup>160</sup>. Mais pas seulement. L'**apprentissage machine** implique le TDM et le besoin va donc bien au-delà de la recherche proprement dite.

Le Conseil Supérieur de la Propriété Littéraire et Artistique a défini le TDM comme un « processus de recherche automatisé qui porte sur un ensemble de données numériques dans le but de découvrir de nouvelles connaissances ou de nouvelles idées »<sup>161</sup>. Le rapport Villani sur l'IA parle, lui, d'un « ensemble de traitements informatiques consistant à extraire des

---

<sup>157</sup> En ce sens, v. cons. 10, dir. 2019/790.

<sup>158</sup> V. les rapports remis à la Commission européenne : J.-P. Triaille, J. de Meeûs d'Argenteuil et A. de Francquen, *Study on legal framework text and data mining*, mars 2014, et I. Hargreaves (dir.), *Standardisation in the area of innovation and technological development, notably in the field of text and data mining*, juill. 2014. – Adde rapport Reda (*Rapport sur la mise en œuvre de la directive 2001/29/CE du Parlement européen et du Conseil du 22 mai 2001 sur l'harmonisation de certains aspects du droit d'auteur et des droits voisins dans la société de l'information*), 2014/2256 (INI), 24 juin 2015, proposition n° 28.

<sup>159</sup> Audition du CNRS, rapportée par C. Mélot, avis n° 525 au nom de la commission de la culture, de l'éducation et de la communication du Sénat sur le projet de loi pour une République numérique, 5 avr. 2016, p. 45.

<sup>160</sup> P. Kamocki, « “Laissez fouiller !” – L'argument pour les “utilisations orthogonales” des œuvres de l'esprit dans le contexte du débat sur l'exploitation des données », RIDA janv. 2016, p. 5 s., spéc. p. 7.

<sup>161</sup> J. Martin et L. de Carvalho, Mission du CSPLA sur l'exploration des données, juillet 2014, p. 9.

connaissances selon un critère de nouveauté ou de similarité dans des textes ou des bases de données ».

La Loi pour une République numérique a, dès **2016**, prévu une exception au droit d'auteur pour favoriser la fouille de données<sup>162</sup>. Ainsi, les titulaires de droits (auteurs comme producteurs de base de données) ne peuvent interdire les « copies ou reproductions numériques », à partir d'une source licite, en vue de l'exploration « de textes et de données incluses ou associées aux **écrits scientifiques** » dans un cadre de **recherche publique**, à l'exclusion de toute finalité commerciale<sup>163</sup>.

L'exception française est doublement limitée : par son champ d'application, réservé à l'écrit, excluant ainsi la fouille sur des œuvres audiovisuelles, plastiques ou musicales ; par ses bénéficiaires, puisque seule une partie de la recherche publique est concernée, celle qui ne repose pas sur des partenariats publics-privés. La démarche peut étonner, alors que le droit européen était sur le point d'évoluer et dans la mesure où la directive 2001/29, qui contient à l'article 5 une liste exhaustive des exceptions possibles au sein de l'UE, ne prévoit pas expressément pareille limitation.

Estimant que le cadre législatif existant permettait déjà aux États membres d'autoriser les activités d'exploration de textes et de données dans le cadre de l'exception pour la recherche scientifique (article 5, paragraphe 3, point a), de la directive 2001/29), le Royaume-Uni connaît aussi une exception pour l'analyse des textes et des données (art. 29A du *Copyright, Designs and Patents Act 1988*<sup>164</sup>), introduite en 2014. L'exception est limitée aux utilisations non commerciales. Cependant, les bénéficiaires de l'exception britannique ne sont pas seulement des organismes de recherche, mais aussi des chercheurs individuels (toute personne qui a un accès légal à une œuvre)<sup>165</sup>.

---

<sup>162</sup> Loi n° 2016-1321 du 7 oct. 2016, art. 38. – V. art. L. 122-5, 10°, et L. 342-3, 5°, du CPI.

<sup>163</sup> Le 10° de l'art. L. 122-5 vise « Les copies ou reproductions numériques réalisées à partir d'une source licite, en vue de l'exploration de textes et de données incluses ou associées aux écrits scientifiques pour les besoins de la recherche publique, à l'exclusion de toute finalité commerciale. Un décret fixe les conditions dans lesquelles l'exploration des textes et des données est mise en œuvre, ainsi que les modalités de conservation et de communication des fichiers produits au terme des activités de recherche pour lesquelles elles ont été produites ; ces fichiers constituent des données de la recherche ».

<sup>164</sup> « *Copies for text and data analysis for non-commercial research* :

(1) *The making of a copy of a work by a person who has lawful access to the work does not infringe copyright in the work provided that :*

(a) *the copy is made in order that a person who has lawful access to the work may carry out a computational analysis of anything recorded in the work for the sole purpose of research for a non-commercial purpose, and*  
(b) *the copy is accompanied by a sufficient acknowledgement (unless this would be impossible for reasons of practicality or otherwise).*

(2) *Where a copy of a work has been made under this section, copyright in the work is infringed if :*

(a) *the copy is transferred to any other person, except where the transfer is authorised by the copyright owner, or*  
(b) *the copy is used for any purpose other than that mentioned in subsection (1)(a), except where the use is authorised by the copyright owner.*

(3) *If a copy made under this section is subsequently dealt with :*

(a) *it is to be treated as an infringing copy for the purposes of that dealing, and*  
(b) *if that dealing infringes copyright, it is to be treated as an infringing copy for all subsequent purposes.*

(4) *In subsection (3) "dealt with" means sold or let for hire, or offered or exposed for sale or hire.*

(5) *To the extent that a term of a contract purports to prevent or restrict the making of a copy which, by virtue of this section, would not infringe copyright, that term is unenforceable. »*

<sup>165</sup> N. Jondet, « L'exception pour le *data mining* dans le projet de directive sur le droit d'auteur. Pourquoi l'Union européenne doit aller plus loin que les législations des États membres », *Propriété Intellectuelle* 2018. 25. – *Comp.* : il existe aussi une exception TDM en Allemagne dans le cadre de la recherche scientifique et pour une utilisation non commerciale (UrhG, § 60d).

Le Japon<sup>166</sup> dispose également d'une exception de TDM large et flexible, qui permet toute exploitation, avec ou sans but commercial, dès lors que l'« analyse de l'information (c'est-à-dire l'extraction, la comparaison, la classification, ou autre analyse du langage, d'un son, d'une image, ou de tout autre élément dont un grand nombre d'œuvres ou une grande quantité de données sont composées) » n'a pas pour objectif de « jouir » de l'œuvre<sup>167</sup>.

L'Union européenne s'est saisie de la question puisque, désormais, l'article 2 de la directive 2019/790 du 17 avril 2019<sup>168</sup> définit le TDM comme une « technique d'analyse automatisée visant à analyser des textes et des données sous une forme numérique afin d'en dégager des informations, ce qui comprend, à titre non exhaustif, des constantes, des tendances et des corrélations ».

Les opérations de fouille supposent souvent en amont des reproductions intermédiaires des œuvres ou des bases de données concernées qu'il faut stocker, au moins temporairement, et parfois modifier (format, coupes, fusions, compilations...) pour les rendre analysables<sup>169</sup> : chacune de ces opérations est alors susceptible d'être en contradiction avec les règles de réservation de la propriété intellectuelle, nécessitant un accord exprès de leurs titulaires. Dès lors, la mise en place d'une exception implique indirectement que le monopole avait vocation à s'appliquer en pareille situation. En conséquence, une exception était nécessaire. Mais elle s'applique, selon la directive, dès lors qu'une reproduction est réalisée et que l'acte met en œuvre le droit exclusif sans pouvoir être excusé par l'exception obligatoire pour copie technique<sup>170</sup>. L'approche suggère donc une certaine casuistique dans la mise en œuvre du régime pertinent.

Le texte nouveau procède en deux temps.

La directive « DSM » impose d'abord à l'**article 3**<sup>171</sup> une exception obligatoire et à laquelle on ne peut déroger par le contrat (art. 7.1<sup>172</sup>), au bénéfice « des organismes de recherche et

---

<sup>166</sup> Art. 30-4, point *ii* de la loi japonaise sur le droit d'auteur (loi n° 48 du 6 mai 1970, mod. 2018 ; <http://www.japaneselawtranslation.go.jp/law/detail/?id=3379>) :

« *It is permissible to exploit a work, in any way and to the extent considered necessary, in any of the following cases, or in any other case in which it is not a person's purpose to personally enjoy or cause another person to enjoy the thoughts or sentiments expressed in that work; provided, however, that this does not apply if the action would unreasonably prejudice the interests of the copyright owner in light of the nature or purpose of the work or the circumstances of its exploitation: (...)*

*(ii) if it is done for use in data analysis (meaning the extraction, comparison, classification, or other statistical analysis of the constituent language, sounds, images, or other elemental data from a large number of works or a large volume of other such data; the same applies in Article 47-5, paragraph (1), item (ii) (...)* ».

<sup>167</sup> « Jouir de l'œuvre » signifie « satisfaire le désir de savoir ou s'amuser intellectuellement en écoutant ou en voyant ces œuvres » : v. M. Nagatsuka, « Lettre du Japon. La réforme de la loi sur le droit d'auteur en 2018 et à venir », *Propr. intell.* oct. 2019, n° 73, p. 121.

<sup>168</sup> Dir. 2019/790 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les directives 96/9/CE et 2001/29/CE, 17 avr. 2019.

<sup>169</sup> E. Rosati, « *Copyright as an Obstacle or an Enabler? A European Perspective on Text and Data Mining and Its Role in the Development of AI Creativity* », *Asia Pacific Law Review*, sept. 2019, <https://ssrn.com/abstract=3452376>.

<sup>170</sup> Dir. 2019/790, cons. 9.

<sup>171</sup> « Article 3 - Fouille de textes et de données à des fins de recherche scientifique

1. Les États membres prévoient une exception aux droits prévus à l'article 5, point a), et à l'article 7, paragraphe 1, de la directive 96/9/CE, à l'article 2 de la directive 2001/29/CE et à l'article 15, paragraphe 1, de la présente directive pour les reproductions et les extractions effectuées par des organismes de recherche et des institutions du patrimoine culturel, en vue de procéder, à des fins de recherche scientifique, à une fouille de textes et de données sur des œuvres ou autres objets protégés auxquels ils ont accès de manière licite.

des institutions du patrimoine culturel, en vue de procéder, à des fins de **recherche scientifique**, à une fouille de textes et de données sur des œuvres ou autres objets protégés auxquels ils ont accès de manière licite. »

La directive prévoit encore, à l'**article 4**<sup>173</sup>, une autre exception TDM, plus large, pour **tous les usages de l'IA, y compris commerciaux** (avec dans tous les cas la limite du triple test ; art. 5.5 de la dir. 2001/29), qui peut apparemment être limitée par le contrat (pas de mention à l'article 7). Cette nouvelle exception offre en outre une possibilité de sortie – d'*opt-out* – pour les titulaires de droits<sup>174</sup>. Il s'agit donc d'une sorte de consentement présumée, exigeant de l'auteur qu'il manifeste explicitement son refus de l'accorder, ce qui, dans le cas d'un contenu laissé à la disposition du public en ligne, nécessitera une réservation « au moyen de procédés lisibles par machine, y compris des métadonnées et les conditions générales d'utilisation d'un site internet ou d'un service »<sup>175</sup>. Cette faculté de sortie a été introduite comme contrepartie au caractère obligatoire et au risque de parasitisme du fait de l'ouverture large de l'exception.

Ces deux dispositions (articles 3 et 4 précités) concernent **tous les types de données** : textes, sons, images, etc., et elles exigent l'**accès légitime** à l'œuvre, ce qui garantit la rémunération des auteurs. Elles ne font exception que pour le droit de reproduction, et non le droit de représentation : le TDM ne peut être utilisé comme cheval de Troie pour la diffusion d'objets protégés sans autorisation des ayants droit. Enfin, les copies des contenus protégés pourront être conservées, notamment pour la vérification des résultats de la recherche, mais des obligations de sécurité devront garantir leurs conditions de stockage<sup>176</sup>. Ne figure cependant aucune disposition spécifique sur l'effacement automatique des données. Le point 2 de l'article 4 mentionne néanmoins que les reproductions peuvent être conservées « aussi longtemps que nécessaire aux fins de la fouille de textes et de données ». Ce renvoi à la *nécessité* doit être interprété comme une exigence d'effacement dès que la copie n'a plus d'utilité pour l'opération.

---

2. Les copies des œuvres ou autres objets protégés effectuées dans le respect du paragraphe 1 sont stockées avec un niveau de sécurité approprié et peuvent être conservées à des fins de recherche scientifique, y compris pour la vérification des résultats de la recherche.

3. Les titulaires de droits sont autorisés à appliquer des mesures destinées à assurer la sécurité et l'intégrité des réseaux et des bases de données où les œuvres ou autres objets protégés sont hébergés. Ces mesures n'excèdent pas ce qui est nécessaire pour atteindre cet objectif. »

4. Les États membres encouragent les titulaires de droits, les organismes de recherche et les institutions du patrimoine culturel à définir d'un commun accord des bonnes pratiques concernant l'application de l'obligation et des mesures visées aux paragraphes 2 et 3, respectivement. »

<sup>172</sup> « Toute disposition contractuelle contraire aux exceptions prévues aux articles 3, 5 et 6 est non exécutoire. »

<sup>173</sup> « Article 4 - Exception ou limitation pour la fouille de textes et de données

1. Les États membres prévoient une exception ou une limitation aux droits prévus à l'article 5, point a), et à l'article 7, paragraphe 1, de la directive 96/9/CE, à l'article 2 de la directive 2001/29/CE, à l'article 4, paragraphe 1, points a) et b), de la directive 2009/24/CE et à l'article 15, paragraphe 1, de la présente directive pour les reproductions et les extractions d'œuvres et d'autres objets protégés accessibles de manière licite aux fins de la fouille de textes et de données.

2. Les reproductions et extractions effectuées en vertu du paragraphe 1 peuvent être conservées aussi longtemps que nécessaire aux fins de la fouille de textes et de données.

3. L'exception ou la limitation prévue au paragraphe 1 s'applique à condition que l'utilisation des œuvres et autres objets protégés visés audit paragraphe n'ait pas été expressément réservée par leurs titulaires de droits de manière appropriée, notamment par des procédés lisibles par machine pour les contenus mis à la disposition du public en ligne.

4. Le présent article n'affecte pas l'application de l'article 3 de la présente directive. »

<sup>174</sup> V. *supra* l'article 4.3 de la directive 2019/790.

<sup>175</sup> Cons. 18.

<sup>176</sup> Cons. 15.

En consacrant ces exceptions, et précisément celle de l'article 4 qui intéresse particulièrement les usages de l'IA, le législateur européen a **renforcé** le champ du monopole à un endroit où on pouvait douter de son applicabilité.

Des interrogations demeurent néanmoins quant à la **faisabilité** même du mécanisme d'*opt-out* prévu. Comment pourra-t-on considérer qu'un auteur aura « réservé de manière appropriée les droits de reproduction et d'extraction »<sup>177</sup> ? Faudra-t-il qu'il se déclare expressément, par exemple sur un fichier centralisé<sup>178</sup> ? Par quels moyens techniques s'assurer de l'efficacité de cette réserve ? Sur ce point, la mise en place d'un **standard technique** serait d'un intérêt évident, pour peu qu'il soit correctement pris en compte. *Quid* par ailleurs en cas de contradiction sur des fichiers différents et avec quelle responsabilité ? On peut aussi se demander si l'exception ne pourrait pas devenir un outil de négociation entre les mains des acteurs de l'IA. Peut-on imaginer qu'un moteur de recherche refuse de référencer les contenus protégés pour lesquels le titulaire aurait expressément interdit le TDM, en ayant exercé sa faculté d'*opt-out*, ou pour lesquels il y aurait un doute ? A moins que cette exception, qui suscite la réprobation chez certains ayants droit, ne les conduise tout simplement à exercer l'*opt-out*, par précaution/par défaut, pour éviter un pillage massif.

Ces incertitudes au sujet de l'article 4 de la directive 2019/790 laissent présager des difficultés de transposition et, par voie de conséquence, une harmonisation relative au sein de l'UE.

Au-delà de ces interrogations, ce texte nécessitera très certainement le concours des grandes plateformes et outils de recherche afin de garantir, par des moyens techniques, sa bonne application et le suivi des différentes réserves émises par les auteurs dont les œuvres sont mises à la disposition du public en ligne.

### 3.2.2. Le mécanisme des licences comme solution adaptée ?

Le mécanisme des exceptions au monopole d'exploitation est, on l'a vu, trop incertain pour pouvoir tout à la fois assurer la protection légitime des ayants droit et le développement du secteur de l'IA, dont l'essor doit être favorisé.

Plutôt qu'un dispositif réglementaire d'exception « molle » (et incertaine), il pourrait être envisagé d'avoir recours, sur la base du **volontariat**, à un mécanisme de **licence « générale »**, sur le modèle du contrat général de représentation, « par lequel un organisme professionnel d'auteurs confère à un entrepreneur de spectacles la faculté de représenter, pendant la durée du contrat, les œuvres actuelles ou futures, constituant le répertoire dudit organisme aux conditions déterminées par l'auteur ou ses ayants droit » (CPI, art. L. 132-18). En ce sens, les organismes de gestion collective notamment pourraient accorder des licences sur leurs répertoires (pour leurs adhérents) dans le but de permettre la fouille. On peut même se demander si ne pourraient pas être mis en place des systèmes permettant non seulement la

---

<sup>177</sup> Le cons. 18 de la directive indique : « En ce qui concerne les contenus mis à la disposition du public en ligne, la réservation de ces droits ne devrait être jugée appropriée que si elle est effectuée au moyen de procédés lisibles par machine, y compris des métadonnées et les conditions générales d'utilisation d'un site internet ou d'un service. La réservation de droits aux fins de la fouille de textes et de données ne devrait pas affecter d'autres utilisations. Dans d'autres cas, il peut être approprié de réserver les droits par d'autres moyens, comme des accords contractuels ou une déclaration unilatérale. Les titulaires de droits devraient pouvoir appliquer des mesures destinées à garantir le respect de leurs droits réservés. »

<sup>178</sup> Mais cela ne contreviendrait-il pas à l'esprit du principe même du droit d'auteur, qui protège en dehors de toute démarche préalable ?

licitation des usages mais également l'accès, dans un format déjà structuré pour le TDM, aux « données » utiles<sup>179</sup>.

Enfin, l'article 12 de la directive 2019/790<sup>180</sup> offre la possibilité aux Etats membres d'adopter des mécanismes de « licences collectives ayant un effet étendu », en considération de « la nature de certaines utilisations ainsi que le nombre habituellement élevé d'œuvres ou autres objets protégés concernés » (cons. 45 ; v. aussi cons. 47), permettant « aux organismes de gestion collective de conclure des licences sur une base volontaire, indépendamment du fait que l'ensemble des titulaires de droits aient autorisé l'organisme concerné à le faire » (cons. 46). Le paragraphe 3 de l'article 12 liste les garanties appropriées propres à protéger les intérêts légitimes des titulaires de droits y compris non-membres de l'organisme de gestion collective. D'abord, cet organisme doit être « suffisamment représentatif » ; il doit exister ensuite une « égalité de traitement » de tous les titulaires de droits ; ceux-ci doivent encore conserver la possibilité de reprendre le contrôle de leurs œuvres en s'opposant à l'application de ce mécanisme « à tout moment, facilement et de manière effective », cette possibilité d'*opt-out* (de la licence) devant donner lieu à publicité.

Un tel dispositif pourrait, dans certains secteurs, utilement accompagner le développement de la création par l'IA, même si la territorialité en limite sans doute l'intérêt.

Pour autant, une telle option obligerait sans doute, en définitive, pour les ayants droit définitivement opposés au TDM à réaliser deux fois l'*opt-out* (pour échapper au TDM et pour s'extraire de la gestion collective).

---

<sup>179</sup> V. aux USA le modèle *RightFind XML for Mining* : <https://www.youtube.com/watch?v=-gUhAkwZbVQ> ; <http://www.copyright.com/business/xmlformining/>

<sup>180</sup> V. également l'article 7 qui prévoit les licences collectives étendues pour les œuvres indisponibles dans le commerce. – V. en doctrine, F. Siiriainen, « Les licences collectives à effet étendu : entre gestion collective consentie, facilitation des autorisations et sécurité juridique (Commentaire de l'article 12 de la directive) », CCE oct. 2019, dossier 5 ; C. Bernault, « Gestion collective – Licence collective étendue », PI juill. 2019, n° 72, p. 59.

## **Partie 4. – Améliorer le partage de données pour répondre aux enjeux du développement de l’IA**

Dès lors que l’intelligence artificielle est présentée comme une des clés du pouvoir de demain dans une économie numérique et que celle-ci se nourrit de données, la question, en amont, de l’accès aux données et des asymétries critiques entre les acteurs en matière d’accès se révèle cruciale. Pour les données « œuvres », la compatibilité entre l’essor des modèles économiques reposant sur l’IA et le respect de la protection par la propriété intellectuelle suscite des interrogations, comme vient de le montrer la partie précédente. Pour d’autres catégories que sont les données d’usage et les métadonnées définies dans la première partie de ce rapport, les enjeux de la circulation et du partage de données, sans toujours être directement reliés à des questions de propriété intellectuelle, n’en sont pas moins considérables. La réflexion s’inscrit dans un contexte plus général de libre circulation des données, avec l’adoption du Règlement 2018/1807, du 18 novembre 2018, « établissant un cadre applicable au libre flux des données à caractère non personnel dans l’Union européenne »<sup>181</sup>, ou les différentes stratégies européennes dévoilées visant à favoriser le partage des données<sup>182</sup>.

Dans le secteur culturel, après avoir précisé l’importance des enjeux de maîtrise de ces données dans la perspective d’un essor des utilisations de l’IA (4.1), les possibles améliorations de l’accès et du partage de données seront exposées (4.2).

### ***4.1. Les enjeux de l’accès et du partage des données pour le secteur culturel***

Les enjeux du partage de ces données ne sont pas les mêmes dans tous les cas. Pour les données d’usage, les enjeux sont ceux non seulement de la question classique de la transparence, mais également de la maîtrise de la relation-client (4.1.1). Pour les métadonnées, il s’agit essentiellement de qualifier les contenus afin d’accompagner des applications développées à tous les stades de la chaîne de valeur (4.1.2).

#### **4.1.1. Les données d’usage : des enjeux de transparence et de maîtrise de la relation-client**

La consommation de contenus en ligne s’accompagne d’une production sans précédent de données d’usage. Mais, paradoxalement, alors que le monde numérique est marqué par une traçabilité potentielle presque totale et bien plus riche de l’ensemble des interactions avec les utilisateurs finaux, les données d’usage sont parfois moins accessibles que dans le monde physique. Dans ce dernier, la question de la transparence de la rémunération des ayants droit est un des enjeux centraux qui a motivé l’intervention législative pour réduire les asymétries d’informations au profit de l’aval.

C’est pourquoi il existe, en France depuis 1948, un contrôle strict, réalisé par le CNC, des entrées réalisées dans les salles, afin de procéder de manière fiable aux opérations de remontées de recettes au profit des ayants droit. Il existe, de même, des prémices d’exigences

---

<sup>181</sup> Principalement, ce texte prohibe les restrictions à la localisation de données au sein de l’UE, impose la libre accessibilité des données par les autorités nationales et encourage, par l’élaboration de codes de conduite, le partage des données au bénéfice des utilisateurs professionnels.

<sup>182</sup> Communication de la Commission, *Créer une économie européenne fondée sur les données*, COM(2017) 09 final, 10 janv. 2017 ; Communication de la Commission, *L’intelligence artificielle pour l’Europe*, COM(2018) 237 final, 25 avr. 2018.

de transparence dans le Code de la propriété intellectuelle au travers de l'obligation de rendre compte des éditeurs, étendue à l'édition numérique, en matière de représentation de spectacles, de production audiovisuelle ou de cession des droits de l'artiste-interprète à un producteur de phonogrammes (voir respectivement les articles L. 132-21, L. 132-25 et L. 212-15 du CPI).

La transparence des comptes de production et d'exploitation des œuvres cinématographiques et audiovisuelles est également organisée par la loi n° 2016-925 du 7 juillet 2016 relative à la liberté de la création, à l'architecture et au patrimoine (LCAP) qui introduit, dans le Code du cinéma, un nouveau chapitre sur ce sujet. Ce texte systématise la transmission des comptes de production et d'exploitation et apporte une nouveauté, celle d'étendre les obligations de restitution régulière des recettes et des coûts, non seulement aux ayants droit mais également aux coproducteurs et à toute personne intéressée aux recettes : ces restitutions sont également transmises aux partenaires financiers du film, ainsi qu'aux artistes-interprètes et aux techniciens.

Les obligations légales de transparence des données sur les rémunérations ne concernent cependant pas les services de vidéo à la demande (VAD) par abonnement alors même que dans ces modèles, où les revenus des éditeurs dépendent du nombre de pages lues ou du temps passé, l'accès aux données est déterminant. Certains acteurs comme Netflix, dont le modèle d'affaire n'oblige pas à rémunérer leurs fournisseurs de contenus à l'acte, ont, après avoir fourni très peu de données aux ayants droit, évolué : la SACD a ainsi conclu des accords avec Netflix, tandis qu'une association européenne de producteurs audiovisuels a ouvert des discussions avec la plateforme. Mais il s'agit toujours d'engagements purement contractuels. De plus, toutes ces dispositions ne traitent que de la cession de droits et ne s'appliquent pas aux sortes de plateformes qui assurent la diffusion des contenus numériques sans s'inscrire dans ce schéma contractuel.

Aucun instrument de droit français ne permet donc aux ayants droit de participer à une économie pilotée par la donnée, faute d'avoir accès aux données d'usage détenues par les acteurs de l'aval : celles qui permettent de savoir, par exemple dans l'édition, comment les lecteurs lisent, quel passage ils ont surligné, à quel rythme ils lisent, quand ils abandonnent, les pages qu'ils survolent...

En Europe, la directive « DSM » 2019/790 tend à garantir le droit à l'information des auteurs et interprètes en introduisant, dans son article 19, des obligations de transparence afin que les ayants droit reçoivent des informations sur l'exploitation de leurs œuvres notamment sur l'ensemble des revenus générés et la rémunération due. Mais « ces dispositions portent la marque de l'ère analogique, où tout ce qu'un auteur avait besoin et pouvait envisager de savoir était le nombre d'exemplaires de ses ouvrages vendus et les revenus que cela lui apportait »<sup>183</sup>. Une interprétation ambitieuse de la directive 2019/790, pourrait cependant plaider pour un partage des données d'usage, dans l'intérêt conjugué des titulaires et des utilisateurs, dès lors qu'ils y consentent, afin de respecter leur vie privée.

Au-delà de l'enjeu de transparence pour une juste répartition des revenus, au profit des ayants droit, ce sont, en réalité, **tous les équilibres au sein de la filière qui s'avèrent susceptibles d'être remis en cause par l'accès aux données d'usage ou au contraire par la perte de la maîtrise de la relation-client.** Dans un univers numérique d'individualisation et de

---

<sup>183</sup> V.-L. Benabou *et alii*, *Droit de la propriété littéraire et artistique, données et contenus numériques*, rapport du CSPLA, sept. 2018, p. 124.

connaissance de plus en plus fine des pratiques des consommateurs, le contrôle de cette relation est devenu crucial. Dans de nombreux cas, dans le monde numérique, la seule transparence, c'est-à-dire la publication de données de consommation globale sur les performances d'un contenu ne suffit plus. Pour des usages mobilisant de l'IA, ce type d'accès à des données de faible granularité se révèle, le plus souvent, décevant. D'autres données plus qualifiées deviennent indispensables : performances des autres œuvres proposées par les fournisseurs concurrents à des fins de comparaison/mise en perspective, profils des utilisateurs, données contextuelles de présentation des œuvres, provenance du lien menant au contenu, nombre de fois où il est proposé par un algorithme de recommandation...

Pour les auteurs, l'accès aux données d'usage peut jouer un rôle potentiellement important dans le processus de création, dans une logique de « *data-driven creation* » (création guidée par la donnée), afin notamment de promouvoir une meilleure interaction entre l'auteur et son public. Pour les éditeurs, les producteurs, les distributeurs, l'accès aux données d'usage est également devenu indispensable. Ce sont les données d'usage qui permettent de connaître la façon dont l'œuvre est reçue, de séduire un public dans une perspective de distribution plus fine, plus personnalisée, d'adapter la stratégie de distribution aux plateformes sur lesquelles le contenu se révèle plus performant. Ce sont elles qui rendent possibles d'enrichir l'offre de services en proposant des recommandations personnalisées, en appréhendant les nouvelles tendances, de renouveler les services proposés, d'accompagner les prises de décision (motiver la production ou l'achat d'un contenu, ou argumenter dans le cadre de négociations commerciales). Ce sont elles qui autorisent la publicité ciblée facilitée par le numérique.

Or, la constitution de bases de données d'usage par les fournisseurs de terminaux, de systèmes d'exploitation, les réseaux sociaux ou les agrégateurs de contenus pourrait conduire à la perte du contrôle, par les éditeurs et distributeurs traditionnels de contenus, de la relation avec le client final. Ainsi, des éditeurs de presse ou de services audiovisuels n'ont pas accès aux données sur leurs propres contenus circulant sur les réseaux sociaux. Les opérateurs de télécommunications qui proposent des box et des services par contournement (*over the top service* ou service OTT) récoltent de multiples données d'usage en les restituant de manière très limitée, selon des accords contractuels, avec les chaînes de télévision qui éditent les programmes. Afin de reprendre le pouvoir sur les données utilisateurs, les chaînes de télévision ont ainsi pris l'habitude de demander une inscription à leur service numérique, comme M6 qui la première, dès décembre 2015, a rendu obligatoire l'inscription à son service de vidéo en *replay* : elle demande à l'internaute de renseigner son sexe et son âge, ce qui peut faire l'objet de réticences de l'utilisateur alors même que les opérateurs de télécommunication ou de réseau social, qui fonctionnent dans des univers déjà « *loggés* » (identifiés), n'ont pas besoin de l'exiger.

Collectées par les acteurs économiques qui exercent leurs activités au plus près des consommateurs finaux des œuvres, les données d'usage sont donc cruciales pour l'ensemble de la filière. La **concentration des données chez ces acteurs « aval »**, notamment les plateformes de mise à disposition de contenus, et l'éventuelle **absence de partage** de ces données avec l'amont, pourraient être regardées comme de véritables **freins à l'innovation permise par l'IA**. Selon son intensité, l'asymétrie d'information peut créer de nouveaux rapports de force au sein des filières, une forte dépendance de l'amont sur l'aval et des comportements opportunistes de l'aval. Tandis que des opérateurs de l'amont travaillent « à l'aveugle » et ne sont pas en mesure de développer des services innovants à partir de l'activité informationnelle générée par l'utilisation des œuvres, de nouveaux métiers apparaissent précisément centrés sur l'analyse de plus en plus sophistiquée des comportements d'usagers.

#### 4.1.2 Les métadonnées : des enjeux de qualification des contenus

Les métadonnées visant à qualifier précisément les contenus se multiplient et jouent un rôle majeur dans les services ; elles accompagnent les processus de création artificielle et renouvellent considérablement les formes de prévision ou de recommandation, proposées dans les industries culturelles.

Or, parmi l'ensemble des métadonnées pertinentes, celles, pourtant basiques, concernant l'identification de chaque œuvre ne sont pas toujours harmonisées. Au-delà de la bataille pour l'identification, la manière dont les multiples métadonnées descriptives ou enrichies sont produites et rendues disponibles pour chaque œuvre n'est pas davantage mutualisée. Les informations peuvent être incomplètes, inexactes ou simplement non accessibles. Les métadonnées peuvent en effet être générées ou enrichies à différentes étapes de la chaîne de valeur et les acteurs de chaque secteur ont tendance à créer eux-mêmes des bases de métadonnées correspondant à leurs objectifs propres, sans se soucier d'une vision globale et de leurs interdépendances.

Parmi les professionnels, le regard porté sur les métadonnées est très différent selon les secteurs. Dans l'industrie musicale, la question fait l'objet d'une réelle prise de conscience et les collaborations se multiplient. Dans l'audiovisuel, malgré quelques tentatives, les standards et les normes de métadonnées sont soumises à une très forte concurrence les unes avec les autres, sans qu'aucune, publique ou privée, ne réussisse à s'imposer et sans que le sujet se soit véritablement invité dans les agendas professionnels. Sur le marché de l'art, la commercialisation en ligne tardive n'a pas conduit à une réflexion sur la mise en place de standards de métadonnées. Des bases de reproductions d'œuvres, notamment de photos, circulent largement sur le web, mais selon des règles d'association avec des métadonnées éparses.

Sur ce marché de la métadonnée, l'IA est largement sollicitée car elle transforme les possibilités d'analyse des œuvres, d'extraction et de production de métadonnées associées. Si l'IA prolonge et automatise les possibilités de production et d'extraction de métadonnées, cette automatisation reste cependant coûteuse et source d'erreurs, ce qui nécessite encore, à ce stade, de nombreuses interventions manuelles. Le chantier de l'amélioration qualitative des métadonnées produites reste donc, encore, largement inachevé.

#### 4.2. Les voies futures du partage

Pour répondre aux enjeux, dont on a vu l'importance, de l'accès et du partage des données dans le secteur culturel, plusieurs voies peuvent être explorées.

La première est, dans un cadre coopératif, la constitution d'outils de gestion à gouvernance partagée à l'instar de ceux fonctionnant, par exemple, dans les *smart cities*, sur la logique des **communs**, théorisée par l'économiste Ostrom. De multiples exemples de partage coopératif existent dans le secteur culturel.

La seconde option est de laisser les acteurs économiques convenir, entre eux, dans le cadre du **marché**, de la meilleure façon d'organiser le partage de données. Dans les relations commerciales, la réutilisation des données s'inscrit déjà largement dans une logique d'API selon laquelle un acteur privé propose à d'autres de tirer profit de la valeur d'usage de

données qu'il détient. La limite de cette option est que ce partage reste dans ce cas « à la main » du premier opérateur qui peut toujours en limiter ou en refuser l'accès.

C'est pourquoi une troisième option, celle des **modalités réglementaires d'ouverture et de partage**, peut s'avérer complémentaire. C'est à cette dernière option que nous nous intéressons dans cette partie, en explorant tout d'abord la piste, fournie par le droit de la concurrence, des facilités essentielles (4.2.1). Nous discutons ensuite des modalités de partage des deux catégories de données identifiées dans le secteur culturel, les données d'usage (4.2.2) et les métadonnées (4.2.3).

#### **4.2.1. Faire des données culturelles des infrastructures essentielles : une proposition séduisante mais inadaptée**

La théorie des infrastructures essentielles trouve son origine en droit américain de l'antitrust au début du XX<sup>e</sup> siècle. Elle n'est pas législative, et sa formalisation dépend de la jurisprudence. Lorsque le titulaire d'une position dominante maîtrise l'accès à une ressource et en refuse la fourniture, alors qu'elle conditionne l'entrée sur le marché, cette situation peut constituer un abus de position dominante. La ressource est alors appelée « infrastructure essentielle », notion qui donne son nom à une théorie, laquelle peut conduire à contraindre une entreprise, publique ou privée, investie d'une mission de service public ou non, qui possède ou contrôle une facilité essentielle, à coopérer : « quand l'accès à une ressource est indispensable pour opérer sur un marché, le propriétaire de cette facilité peut, dans certaines circonstances, être obligé de le garantir aux autres opérateurs même s'il s'agit de ses concurrents »<sup>184</sup>, contre rémunération. Si cette théorie trouve à s'appliquer, le régulateur sera alors en mesure d'imposer au détenteur de l'infrastructure concernée de mettre son installation à la disposition d'autres entreprises afin qu'elles opèrent sur son marché ou sur un marché amont ou aval, à des conditions de prix raisonnables (par référence aux coûts d'exploitation) et non discriminatoires.

La Commission européenne a reçu cette théorie dans le droit européen pour la première fois en 1992, à propos d'installations portuaires<sup>185</sup>. Elle a édicté les conditions de son application. Pour qu'une infrastructure soit jugée essentielle, il faut que l'entreprise concernée soit en situation de position dominante sur le marché où se situe l'infrastructure ; que l'infrastructure soit effectivement indispensable pour l'activité de l'entreprise en réclamant l'accès ; et qu'il soit impossible à un opérateur, d'un point de vue pratique ou raisonnable, de reproduire l'installation essentielle : le fait que la reproduction soit difficile ou simplement coûteuse n'est pas suffisant.

C'est en 1995 que la CJUE, dans son arrêt *Magill*<sup>186</sup>, fait entrer cette théorie dans le droit de la concurrence européen sur une question concernant spécifiquement la propriété intellectuelle, celle de l'accès à une licence. Dans cet arrêt, la Cour a estimé que si le refus des chaînes de télévision de fournir des licences d'exploitation à un journal présentant l'ensemble de leurs programmes télévisés ne constituait pas en soi un abus de position dominante, dans la mesure où le droit d'auteur leur confère un droit exclusif de reproduction, tel pouvait être le cas dans certaines « circonstances exceptionnelles » dans la mesure où le refus de fournir

---

<sup>184</sup> C. Castets-Renard, F. Marty, v° « *Facilités essentielles* », in *Dictionnaire des communs*, dir. M. Cornu, J. Rochfeld, F. Orsi, PUF, 2017.

<sup>185</sup> Décis. n° 94/19 de la Commission, 21 décembre 1993, *Sea Containers c./ Stean Sealink*, Europe, mars 1994, n° 115, obs. L. Idot.

<sup>186</sup> CJCE, 6 avril 1995, *RTE et ITP c./ Commission*, aff. jointes C-241/91 P et C-242/91 P, Rec. I. 743.

accès aux données brutes aurait pour effet de faire obstacle à l'apparition d'un produit nouveau pour lequel il existait une demande potentielle des consommateurs, et sans que le journal ne puisse se tourner vers une infrastructure qui soit substituable.

En France, la théorie des infrastructures essentielles fut reçue par l'Autorité de la concurrence (avis n° 02-A-08 du 22 mai 2002 relatif à la saisine de l'Association pour la promotion de la distribution de la presse) et la Cour de cassation, qui en précisèrent les conditions d'application.

Dans l'ensemble de la jurisprudence, l'infrastructure essentielle peut tout à la fois être constituée par une infrastructure ou un équipement, par un produit ou un service. Rien n'empêche de considérer que des biens immatériels constituent des ressources pouvant être qualifiées d'infrastructures essentielles. Il est cependant difficile d'obliger une entreprise à ouvrir l'accès à une infrastructure alors même qu'il ne lui est reconnu aucun droit dessus (dans la mesure où les données, en tant que telles, ne sont pas l'objet de droits de propriété). Il faudrait alors réserver la théorie des infrastructures essentielles aux seuls accès aux bases de données appropriables au sens du droit de la propriété intellectuelle, c'est-à-dire protégées par le droit *sui generis* de l'article L. 341-1 du Code de la propriété intellectuelle ou par le droit d'auteur. Une base de données appropriable peut bien *a priori* être jugée comme une infrastructure essentielle à laquelle l'entreprise qui détient ces droits, si elle est située en position dominante, devrait pouvoir être obligée de livrer l'accès afin de permettre à d'autres acteurs, d'exercer leurs propres activités économiques.

Une autre solution serait de considérer que les données n'étant pas appropriables, hormis dans les hypothèses où elles sont structurées en bases de données protégées, ce n'est guère que l'accès à ces données qui doit être autorisé – c'est-à-dire l'accès aux fichiers qui les « abritent ». Cela équivaldrait, en quelque sorte, à un « **droit de passage** » **immatériel** dans les serveurs de l'entreprise en situation de position dominante afin d'accéder aux données qu'elle ne possède pas mais dont elle a l'accès, par exemple par le moyen d'API.

La détention de données peut donc accroître le pouvoir de marché de certains acteurs et créer les conditions d'un abus de position dominante. Encore faut-il que ces données soient considérées comme des infrastructures essentielles, c'est-à-dire comme n'étant pas reproductibles ou substituables par des moyens économiques raisonnables, ce qui sera rarement le cas. La plupart des données nécessaires à une activité peuvent en effet être obtenues par des sources différentes ; les données sur les goûts musicaux par exemple sont susceptibles d'être connues par des achats sur des sites marchands, des sites de *streaming*, des moteurs de recherche généralistes ou les pages personnelles des réseaux sociaux et ne pourront donc pas relever de facilités essentielles. Ce n'est pas la donnée qui, en tant que telle, n'est pas reproductible mais son croisement avec d'autres et sa contextualisation.

Une étude allemande mandatée par le ministère fédéral pour les affaires économiques et l'énergie et menée par l'Institut de Düsseldorf d'économie de la concurrence s'est elle aussi posée la question des « **données essentielles** ». Elle a recommandé de faire explicitement mention de la possibilité d'accès raisonnable à des données essentielles comme critère pertinent à prendre en compte lors de la détermination d'un pouvoir de marché d'un opérateur. Elle proposait ainsi de rendre possible l'évaluation des « intérêts légitimes » en cause sur le marché de la donnée à la lumière de la finalité concurrentielle lorsqu'une entreprise refuse l'accès d'une autre entreprise à ses propres données. Toutefois, elle note également que le seuil à franchir pour qualifier le refus d'accès d'abusif doit nécessairement

être abaissé par rapport au seuil correspondant aux cas plus traditionnels de refus d'accès aux infrastructures essentielles. Le rapport précise ainsi que tel serait plus facilement le cas lorsque le refus porte sur des données d'usage brutes, générées automatiquement, que l'entreprise concernée ne produit qu'incidemment – c'est-à-dire sans investissement particulier de sa part.

La théorie des infrastructures essentielles, parce qu'elle conduit à imposer une véritable obligation de contracter aux acteurs économiques opérant sur un marché, reste une **théorie prudente**. Très strictement entendue par les autorités de la concurrence nationales ou européennes, elle ne repose sur aucun texte en particulier, mais bien sur une pratique. A la seule main des régulateurs et, en dernier ressort, des juges, elle ne s'applique qu'en fonction des circonstances de l'espèce et d'une casuistique variable. Elle n'offre donc aucun levier direct d'action pour permettre de fiabiliser les pratiques dans ce domaine, ni de déployer une politique publique d'ensemble. La doctrine considère en général qu'il n'y a pas de « droit à l'accès », mais uniquement un droit à l'accès en cas d'abus dans le refus d'accès.

La question de savoir si la qualification, exigeante, d'infrastructures essentielles, pourrait être appliquée au *big data*, donc aux masses de données détenues par les puissantes entreprises du numérique, continue cependant à susciter d'intenses débats parmi les spécialistes du droit de la concurrence. Selon la professeure Béhar-Touchais<sup>187</sup>, cette qualification ne s'impose pas sans discussion mais, **dès lors que l'IA se développe, se posera la question de savoir s'il faut créer des licences, permettant à des opérateurs d'avoir un accès payant aux grandes masses de données détenues par d'autres**. On peut penser qu'une régulation sectorielle serait utile pour permettre cet accès.

**Plus qu'une théorie générale aux contours difficilement cernables, c'est donc une approche différente qu'il conviendrait de suivre** afin de répondre, dans la culture, aux enjeux de l'accès et du partage de données permettant d'alimenter les algorithmes d'apprentissage automatique. Il faut en revenir aux fondements-mêmes des interrogations qui ont précédé et prendre le chemin d'une approche segmentée correspondant aux deux catégories de données identifiées dans le secteur culturel, les données d'usage et les métadonnées.

#### 4.2.2. La piste de la portabilité des données d'usage

**Afin d'encourager l'accès et le partage des données d'usage, une option consisterait à créer un nouveau droit à la « portabilité » de ces données pour rectifier l'asymétrie informationnelle entre les acteurs des filières culturelles**. Ce droit à la portabilité des données d'usage dont les débiteurs seraient non seulement les exploitants au sens du Code de la propriété intellectuelle, mais aussi tous ceux qui fournissent au public des services d'intermédiation donnant accès aux œuvres et objets protégés, est une des propositions formulées dans un précédent rapport du CSPLA<sup>188</sup>. Une clause contractuelle impérative pourrait ainsi être imposée dans les contrats liant les titulaires de droits sur les contenus protégés et l'opérateur ayant obtenu l'autorisation de les exploiter.

---

<sup>187</sup> M. Béhar-Touchais, « Intelligence artificielle et droit de la concurrence », in *Droit de l'intelligence artificielle*, dir. A. Bensamoun et G. Loiseau, LGDJ-Lextenso, coll. Les Intégrales, 2019, p. 171.

<sup>188</sup> V.-L. Benabou *et alii*, *Droit de la propriété littéraire et artistique, données et contenus numériques*, rapport du CSPLA, sept. 2018.

L'objectif serait d'assurer la connaissance des données d'usage portant sur les œuvres concernées, dont disposent les opérateurs de l'aval en relation directe avec le consommateur final. **La portabilité ne viserait pas le simple accès et la transparence mais bien, dans un monde « data driven » où les métiers évoluent, la réutilisation, la valorisation d'un spectre plus large de données d'usage afin de mettre en œuvre les nouveaux services rendus possibles par l'IA.**

La revendication des titulaires de disposer d'une connaissance fine des utilisateurs supposerait d'inclure des données à caractère personnel, dans le respect des dispositions de protection de la vie privée. Dans les secteurs culturels, sont en effet considérées comme des données personnelles au sens Règlement Général pour la Protection des Données (RGPD) les données d'usage telles que nous les avons définies dans le cadre de ce rapport, dès lors qu'elles sont reliées directement ou indirectement à un individu et concernent les caractéristiques, comportements, habitudes, goûts ou jugements d'une personne précise. La transparence et la maîtrise des données personnelles par les individus sont au cœur du Règlement européen précité qui en prévoit cependant une forme de « partage » par le biais d'un droit à la portabilité ; cette forme de portabilité citoyenne qui a pour objectif de permettre aux internautes de maîtriser leurs données personnelles reste donc à la seule « main » de ces derniers.

La difficulté de l'instauration d'une nouvelle forme de portabilité des données d'usage tient également à la manière de cerner son champ d'application et les éléments constitutifs de la clause impérative. Il s'agira, pour le législateur comme pour les cocontractants, d'identifier les données d'usage concernées, leur niveau de granularité et de traitement ainsi que leur format, la fréquence d'actualisation ou encore la période pendant laquelle les données doivent être disponibles. La clause impérative devrait ainsi probablement laisser aux cocontractants une certaine marge de manœuvre dans la fixation des conditions dans lesquelles les données d'usage seraient remontées (rémunération, format des données, périodicité...). Cette souplesse, qui semble indispensable tant il paraît impossible de prévoir tous les cas de figure, diminuerait cependant d'autant l'effectivité de la loi. Cette piste, prometteuse, reste encore, nécessairement, à affiner.

#### **4.2.3. La piste des métadonnées d'intérêt général**

En France, le mouvement d'ouverture de l'*open data* est organisé pour les données publiques, depuis 1978 par la loi portant diverses mesures d'amélioration des relations entre l'administration et le public, dont le titre I est relatif à « la liberté d'accès aux documents administratifs » (dite loi CADA). La loi pour une République numérique en 2016 a conforté et amplifié ce mouvement, au-delà des données publiques et d'une seule logique d'accès. Les données publiques sont les informations publiques telles que définies par l'article L. 321-1 du Code des relations entre le public et l'administration (CRPA), c'est-à-dire les informations contenues dans les documents administratifs, diffusées ou communicables à tous, non grevées de droit de propriété intellectuelle de tiers. Depuis plusieurs années, le cadre réglementaire qui définit les modalités de mise à disposition et de réutilisation des informations publiques a beaucoup évolué. La loi pour une République numérique, promulguée le 7 octobre 2016, introduit ainsi deux nouveautés en matière de partage.

La première concerne les données publiques. La loi pose le principe de l'*open data* par défaut : l'ouverture des données publiques, déjà pratiquée, devient la règle et non plus l'exception. Dans le secteur culturel, sont concernées au titre des données publiques, celles

détenues par des institutions comme les musées et les bibliothèques. Cependant, le cadre juridique définissant les conditions de mise à disposition et de réutilisation des données des institutions culturelles publiques reste d'une grande complexité et toutes les questions ne sont pas, à ce jour, réglées<sup>189</sup>.

Seconde innovation, la loi pour une République numérique (RN) propose une nouvelle notion, celle de données d'intérêt général (DIG) dont l'accès doit être permis à tous. Leur statut est différent puisqu'il ne s'agit pas de données publiques ; ce sont des données de nature privée mais dont l'ouverture peut se justifier pour répondre à un motif d'intérêt général. A la différence des données personnelles et des données publiques, les DIG n'ont pas de définition stabilisée dans le droit positif ; la loi RN a regroupé sous cet intitulé des dispositions relatives à l'ouverture des données des concessionnaires de service public et des bénéficiaires de subventions ainsi qu'à l'accès de la statistique publique aux données des entreprises privées.

Le rapport Cytermann<sup>190</sup>, qui a introduit la notion de DIG et dont une partie des travaux a été reprise dans la loi RN, met l'accent sur quatre finalités principales justifiant le caractère « d'intérêt général » de ces données et leur ouverture : l'information des citoyens, la conduite de politiques publiques, la recherche et l'innovation économique. La grande nouveauté introduite par la notion de DIG est donc d'organiser l'ouverture de certaines données détenues par des entités privées. C'est dans cette perspective que se situe le rapport Villani<sup>191</sup> qui prône une politique offensive de la donnée visant à favoriser, selon les secteurs, l'accès, le partage et la circulation de certaines données privées des entreprises. Cependant, s'il est légitime de promouvoir l'ouverture des données privées d'intérêt général, il existe également une tension entre la préservation des incitations à constituer de nouvelles bases de données et leur ouverture, la plus large possible, notamment pour stimuler l'innovation économique. Le thème de la circulation des données, de leur partage à des fins d'intérêt général, dans le respect de la liberté d'entreprendre est donc central.

**La question posée dans le secteur culturel est celle de savoir si des motifs d'intérêt général pourraient conduire à ce que certaines métadonnées fassent l'objet de formes de mutualisation et de partage.** Pour répondre aux enjeux de la production qualitative de métadonnées, celles-ci doivent-elles circuler, être accessibles et largement partagées ? Des professionnels soulignent que la qualification des contenus peut être faite de multiples manières et que la qualité de ce travail, parce qu'elle constitue un facteur de différenciation, un marché porteur pour de nombreuses *start-ups*, est un élément d'émulation concurrentielle ; les bases de métadonnées devraient donc rester privatives.

Divers rapports (Lescure, 2013<sup>192</sup> ; *Bearing Point*, 2015<sup>193</sup> ; *Berklee Institute for Creative Entrepreneurship*<sup>194</sup>) ont cependant préconisé, sans être, jusqu'à présent, suivis d'effets, la création de bases de métadonnées consolidées et ouvertes. On peut en effet considérer que la mutualisation de la production de métadonnées et leur partage se révéleraient bénéfiques pour un grand nombre d'acteurs, qui séparément, comme le montre la situation actuelle, ne sont pas incités à investir dans cette production. **L'objectif n'est pas de renforcer la production déjà**

---

<sup>189</sup> Voir notamment J. Farchy, M. de La Taille, *Economie des licences libres dans le secteur culturel*, 2018.

<sup>190</sup> L. Cytermann *et alii*, *Rapport relative aux données d'intérêt général*, sept. 2015.

<sup>191</sup> C. Villani, *Donner un sens à l'intelligence artificielle*, mars 2018.

<sup>192</sup> P. Lescure, *Contribution aux politiques culturelles à l'ère numérique*, Mission « Acte II de l'exception culturelle », La Documentation française, mai 2013.

<sup>193</sup> Etude de faisabilité relative à la mise en place d'un registre ouvert de métadonnées étudiant la proposition n° 79 du rapport de P. Lescure.

<sup>194</sup> Berklee ICE, *Fair Music: Transparency and Money Flows in the Music Industry*, 2015.

**pléthorique de métadonnées mais d'améliorer leur qualité par divers croisements et mutualisations.** Toutes les métadonnées ne seraient pas concernées mais, en priorité, celles descriptives et juridiques.

De plus, comme l'efficacité des algorithmes d'analyse et de reconnaissance des contenus qui sont utilisés pour la génération automatique de nouvelles métadonnées dépend du volume des bases de données utilisées pour l'étape d'entraînement de ces algorithmes, ces outils bénéficient surtout aux acteurs installés qui ont déjà de grands volumes de métadonnées. Ainsi, l'écart se creuse entre les possibilités offertes à la recherche publique ou aux opérateurs de taille modeste et celles mises en œuvre dans un cadre industriel par des entreprises qui disposent, grâce aux métadonnées, d'un avantage concurrentiel important. L'absence de bases de métadonnées complètes, de qualité et accessibles constitue un obstacle pour la communauté des chercheurs qui, avec les progrès de l'apprentissage profond, doit s'entraîner sur des masses considérables de métadonnées pour espérer bénéficier des effets de l'IA. **La constitution de bases de métadonnées partagées entre les acteurs serait susceptible de permettre d'amoinrir ces barrières à l'entrée et, dans une perspective de politique culturelle, de favoriser la diversité des acteurs mettant en œuvre, dans ce secteur, les services innovants permis par l'IA.**

Pour financer les investissements dans la « mise en données », sources de gains d'efficience pour tous mais qu'aucun acteur n'est incité à prendre en charge seul, et faciliter le partage de métadonnées, il serait possible de favoriser des « boucles de rétrocession vertueuses » des données et le développement d'« infrastructures de données », à l'instar de ce qui est promu par l'administrateur général des données s'agissant des données publiques. Il conviendrait dès lors de penser la construction d'un environnement favorable à la réciprocité des gains escomptés par l'exploitation des données en fonction de l'investissement réalisé afin de permettre à chacun d'améliorer son propre service, dans une logique de « coopération »<sup>195</sup>.

### **Vers une acculturation du secteur culturel à une économie « *data driven* »...**

Même lorsque certains contrats ou textes législatifs prévoient des modalités d'accès aux données, leur analyse se heurte, le plus souvent, à la faible acculturation des acteurs. Comme nous l'avons déjà noté, la valeur de la donnée n'existe pas en soi et des investissements nombreux sont nécessaires pour lui donner une valeur économique, mobilisant non seulement des compétences techniques mais également organisationnelles ou managériales. Loin d'être une simple préoccupation technique, une véritable gouvernance de la donnée s'impose pour analyser et intégrer les effets d'une économie « *data driven* ».

Le ministère de la Culture pourrait accompagner cette acculturation, notamment pour des acteurs de taille modeste, afin que, dans ce secteur comme dans d'autres, les potentialités d'une économie de la donnée et de l'IA s'expriment dans l'intérêt collectif.

---

<sup>195</sup> H. Verdier, *La donnée comme infrastructure essentielle*. Rapport de l'administrateur des données sur la donnée dans les administrations, 2016-2017, La documentation française, 2018. Il est à noter que le rapport Verdier utilise la notion d'infrastructure essentielle pour les données, au sens politique du terme, et non au sens du droit de la concurrence examiné précédemment.

## Liste des personnes ayant contribué à la mission

### ➤ Autorités et représentants officiels :

**B. Pailhès**, coordonnateur national de la stratégie d'IA

#### **HADOPI :**

D. Rapone, président

P. Blassel, secrétaire générale

#### **Ministère de la Culture :**

A. de Nervaux, chef du service des affaires juridiques et internationales

S. Jacquier, chargée de mission auprès du chef du service des affaires juridiques et internationales

M.-V. Leroi, DIN

N. Orsini, DIN

A. Therwath, DGCA, délégation à la photographie

C. Devillers-Sigaud, DGCA, BAJ

### ➤ Organisations intéressées :

#### **ADAGP :**

T. Maillard, directeur juridique

#### **Editeurs/filière livre :**

V. Barthez, Les Editeurs d'éducation, directrice

F. Piacentino, SNE, Commission numérique

M. Trauman, directrice adjointe de Hachette Innovation Program

A. Robert, directeur juridique Hachette

J. Chouraqui, directeur juridique SNE

L. Bernard, chargée de mission juridique SNE

L. de Carvalho, responsable juridique Madrigal

M. Benesty, ingénieur *machine learning*, ancien avocat, Dalloz

C. Blache, responsable relations institutionnelles internationales SNE

L. Bernard, chargée de mission commissions Juridique et Illustration, Déléguée à la protection des données

#### **Editeurs/presse :**

*L'Alliance de la presse d'information générale*

S. Ouachtati, responsable du pôle juridique et social

F. Rimbart, Responsable développement numérique

#### **GFII :**

D. Berthault, société LexisNexis, président du GFII

M. Dessertenne, déléguée générale du GFII

A. Raulin, directeur d'études, Bureau van Dijk, administrateur du GFII

F. Campana, responsable juridique, DILA, administrateur du GFII

T. Parisot, responsable des relations institutionnelles CAIRN, membre du CA

**SACD :**

H. Tilliet, directeur des affaires juridiques et des contrats audiovisuels

**SACEM :**

D. El Sayegh, secrétaire général

C. Bonin, directrice juridique

**SAIF :**

A. Defaux, directrice juridique

**SGDL :**

G. Pelletier, directeur

M. Bensimon, responsable juridique

A. Morvan, juriste

D. Couet-Lannes, juriste

**SNEP**

A. Lasch, directeur général

➤ **Entreprises :**

**Aiva Technologies :**

Pierre Barreau, CEO

**IBM :**

J.-M. Leclerc, Government and Regulatory Affairs Executive

**Groupe M6 :**

A. Cathala, juriste contentieux droit d'auteur

**Netflix :**

Marie-Laure Daridan, director Public Policy France

Josh Korn, Innovation and Technology Public Policy

**Rakuten France :**

B. Moutte-Caruel, directeur des affaires juridiques et réglementaires

L. Ach, senior manager Europe, Rakuten Institute of Technology

**Spotify :**

F. Pachet, directeur du Spotify Creator Technology Research Lab

O. Régnier, director of European Policy

**TF1 :**

A. Level, directeur des affaires règlementaires numériques

A. Briffod, responsable juridique

➤ **Experts et personnalités qualifiées**

**G. Azzaria**, professeur, Université Laval-Québec

**V.-L. Benabou**, professeure, Université d'Aix-Marseille, personnalité qualifiée CSPLA

**J.-A. Benazeraf**, avocate, Artlaw, personnalité qualifiée CSPLA

**A. Billion**, doctorant, AI ethics and legal, IBM

**J. Cabay**, professeur, Université libre de Bruxelles

**Raja Chatila**, professeur, Université Pierre et Marie Curie, directeur de l'ISIR (Institut des Systèmes Intelligents et de Robotique)

**De Gaulle Fleurance et associés :**

G. Courtois, avocat associé

J.-S. Mariez, avocat associé

**J.-M. Deltorn**, docteur en physique, doctorant en droit, examinateur à l'OEB

**S. Dormont**, maître de conférences, Université Paris-Est Créteil

**N. Enser**, docteur en droit, avocate, cabinet Vercken-Gaullier

**IRCAM :**

G. Beller, Directeur département Interface recherche-crédation

J.-L. Giavitto, directeur de recherche au CNRS et directeur-adjoint du laboratoire Sciences et technologies de la musique et du son (STMS)

P. Esling, Responsable de l'axe de recherche Orchestration

J. Nika, Human-machine musical interactions: Research (PhD) & Computer Music Design for performance & composition

**T. Lebrun**, juriste en droit du numérique, doctorant en littérature, Université Laval-Québec

**T. E. Synodinou**, professeure, Université de Chypre

**E. Treppoz**, professeur, Université Jean Moulin Lyon 3, président du groupe français de l'AIPPI

**C. Zolynski**, professeure, Université Paris 1, personnalité qualifiée CSPLA

## **Contributions écrites**

Hadopi

SACEM

SGDL

De Gaulle Fleurance et associés