

# SYNTHÈSE

## DU GUIDE JURIDIQUE SUR L'UTILISATION DES DONNÉES PATRIMONIALES DANS LE CADRE DE L'INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE

---

Ce guide a été réalisé par l'INA et la BnF dans le cadre du consortium ArGiMi, lauréat de l'appel « Communs numériques pour l'intelligence artificielle générative » et financé par l'État dans le cadre de France 2030 (n° BPI DOS0238736).

---

Ont contribué à la rédaction du guide :

BnF : Service juridique : Harold Codant, Victor Musitelli / INA : Direction juridique : Jean-François Debarnot, Barbara Mutz, Sandrine Lardeux, Aude Formagne, Johanna Dong / Cabinet Valther : Xavier Près, avocat à la Cour

Mars 2026

<b>INTRODUCTION .....</b>	<b>3</b>
<b>1. RAPPEL SYNTHÉTIQUE DES RÈGLES DE PROTECTION DES DONNÉES PATRIMONIALES .....</b>	<b>4</b>
<b>1.1. Données patrimoniales protégées par des droits de propriété intellectuelle.....</b>	<b>4</b>
<b>1.1.1. Propriété littéraire et artistique .....</b>	<b>4</b>
a) Droit d’auteur.....	4
b) Droits voisins du droit d’auteur.....	5
c) Droits <i>sui generis</i> du producteur de bases de données.....	6
<b>1.1.2. Propriété industrielle .....</b>	<b>7</b>
<b>1.2. Autres protections possibles .....</b>	<b>7</b>
<b>1.2.1. Données personnelles .....</b>	<b>8</b>
<b>1.2.2. Droits de la personnalité .....</b>	<b>8</b>
<b>1.2.3. Droit de la concurrence .....</b>	<b>8</b>
<b>1.2.4. Secret des affaires .....</b>	<b>9</b>
<b>1.2.5. Droit des documents administratifs, des archives publiques et des informations publiques.....</b>	<b>9</b>
<b>1.3. Données patrimoniales protégées par contrats .....</b>	<b>9</b>
<b>2. ENTRAÎNEMENT DES MODÈLES D’INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE .....</b>	<b>10</b>
<b>2.1. Données d’entraînement .....</b>	<b>10</b>
<b>2.1.1. Diversité des données d’entraînement.....</b>	<b>10</b>
<b>2.1.2. Obligations de transparence résultant du Règlement européen sur l’intelligence artificielle .....</b>	<b>11</b>
<b>2.2. Conditions d’accès aux données patrimoniales.....</b>	<b>12</b>
<b>2.3. Conditions d’utilisation des données patrimoniales dont l’accès est licite .....</b>	<b>12</b>
<b>2.3.1. Conditions d’utilisation des données patrimoniales protégées par un droit de propriété intellectuelle .....</b>	<b>12</b>
a) Droit d’auteur (et droits voisins).....	12
b) Droit <i>sui generis</i> du producteur de bases de données .....	13
c) Droit des marques .....	13
d) Droit des brevets et des dessins et modèles .....	13
<b>2.3.2. Conditions d’utilisation des données patrimoniales protégées par d’autres droits .....</b>	<b>14</b>
a) Données à caractère personnel.....	14
b) Données protégées par des droits de la personnalité.....	15
c) Données protégées par le secret des affaires.....	15
d) Données protégées par le droit de la concurrence .....	15
e) Droit des informations publiques.....	15
<b>2.3.3. Conditions d’utilisation des données patrimoniales protégées par contrats .....</b>	<b>15</b>
<b>3. UTILISATION DES SYSTÈMES D’INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE.....</b>	<b>16</b>
<b>3.1. Respect des conditions d’utilisation des systèmes d’intelligence artificielle et des droits sur les données au stade de prompts ou en phase de RAG .....</b>	<b>16</b>
<b>3.2. Respect des droits sur les résultats générés par un système d’intelligence artificielle générative .....</b>	<b>16</b>
<b>3.2.1. Protection des résultats générés par un système d’intelligence artificielle générative .....</b>	<b>16</b>
a) Protection par le droit d’auteur .....	16
b) Protection par un droit de propriété industrielle .....	17
<b>3.2.2. Violation de droits de tiers.....</b>	<b>17</b>

## INTRODUCTION

---

Le guide propose un état des lieux du droit applicable en 2025 aux systèmes d'intelligence artificielle générative réutilisant des données patrimoniales détenues par les institutions culturelles. La présente synthèse en constitue un résumé permettant d'en faciliter la lecture et d'en relever les points saillants. Elle est complétée, pour aller plus loin, d'un renvoi aux paragraphes du guide sur chacun des points ici résumés.

### CADRE JURIDIQUE

Si le cadre juridique spécifique à l'intelligence artificielle (ci-après « **IA** ») est encore en construction, deux textes majeurs marquent une première étape : (i) un règlement européen du 13 juin 2024, qui encadre les systèmes d'intelligence artificielle, et (ii) une convention-cadre du Conseil de l'Europe sur l'IA et les droits fondamentaux adoptée le 17 mai 2024 par le Conseil de l'Europe. En France, une proposition de loi visant à adapter le droit d'auteur à l'IA a été déposée mais n'a pas encore été adoptée. Pour autant, l'absence de règles spécifiques ne signifie pas qu'il y aurait un vide juridique : les règles en vigueur s'appliquent, par analogie, en particulier en matière de propriété intellectuelle, de données personnelles, de contrats, de concurrence ou encore de secret des affaires.

**Pour en savoir plus** : Consultez le Guide de l'IA : §1 à §4

### DÉFINITION DE L'IA

L'IA n'a pas encore de définition unique figée, car il s'agit d'un domaine en évolution rapide. De manière générale, elle désigne la capacité de systèmes techniques à imiter certaines fonctions humaines (raisonnement, planification, créativité) de façon plus ou moins autonome, à partir de données existantes. Le règlement européen parle de « systèmes d'intelligence artificielle » (ci-après « **systèmes d'IA** »), entendus comme des systèmes automatisés, conçus pour fonctionner à différents niveaux d'autonomie et pouvant faire preuve d'une capacité d'adaptation après leur déploiement, et qui déduisent, à partir des entrées qu'ils reçoivent, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions. Ces systèmes d'IA reposent sur des « modèles d'IA », c'est-à-dire des constructions mathématiques qui apprennent à partir des données.

Parmi ces systèmes, l'IA générative occupe une place particulière. Elle désigne la capacité d'une IA à générer automatiquement des contenus (textes, images, vidéos, sons, codes...) en réponse à une ou plusieurs instructions. Ces systèmes d'IA sont dits « à usage général », car ils peuvent être utilisés pour une grande diversité d'usages. Ils reposent sur des modèles d'apprentissage profond, eux-mêmes fondés sur des réseaux de neurones artificiels imitant le fonctionnement du cerveau humain.

Leur fonctionnement est probabiliste : ils produisent un résultat en calculant la réponse la plus plausible, ce qui peut aussi entraîner des erreurs ou des réponses incohérentes (ou « hallucinations »).

Enfin, la donnée est un élément central du fonctionnement de ces systèmes d'IA : leur efficacité dépend fortement de la qualité et de la diversité des données utilisées.

**Pour en savoir plus** : Consultez le Guide de l'IA : §4 à §10

# 1. RAPPEL SYNTHÉTIQUE DES RÈGLES DE PROTECTION DES DONNÉES PATRIMONIALES

## DÉFINITION DES DONNÉES PATRIMONIALES

Les données patrimoniales s'entendent aux fins des présentes comme les documents numériques de toute nature (documents écrits, graphiques, photographiques, sonores, audiovisuels, multimédias, ainsi que les logiciels bases des données et métadonnées) conservés par les institutions du patrimoine culturel, que ces données aient été reçues dans le cadre des missions légales (y compris dans le cadre du dépôt légal) ou par tout procédé d'acquisition emportant transfert de propriété au profit de l'institution (tels que notamment dons, legs, achat, dation en paiement, etc.) ou qu'elles aient été produites par l'institution, directement ou indirectement.

## PROTECTION DES DONNÉES PATRIMONIALES

Les données patrimoniales (documents, images, vidéos, sons, logiciels, bases de données, etc.) peuvent être protégées par plusieurs dispositifs juridiques.



**Point de vigilance** : L'utilisation de données patrimoniales ne peut être opérée sans l'autorisation des titulaires de droits, étant précisé qu'une même donnée patrimoniale peut être protégée à plusieurs titres et par conséquent nécessiter l'autorisation de plusieurs titulaires de droits différents.

**Exemple** : l'utilisation d'une photographie peut, selon les cas, nécessiter l'autorisation du photographe (auteur de la photographie) et de la personne photographiée (droit à l'image). L'utilisation de la même photographie peut encore nécessiter de vérifier dans le contrat par lequel l'institution culturelle a acquis le cliché photographique (par donation, legs, contrat de commande, etc.) les éventuelles réserves liées à son utilisation.

### 1.1. DONNÉES PATRIMONIALES PROTÉGÉES PAR DES DROITS DE PROPRIÉTÉ INTELLECTUELLE

Au sein de la propriété intellectuelle, l'on distingue la propriété littéraire et artistique (1.1.1), qui concerne le droit d'auteur et les droits voisins du droit d'auteur, et la propriété industrielle (1.1.2) qui regroupe le droit des marques, le droit des dessins et modèles, et le droit des brevets (*cf. schéma de synthèse page 7*).

#### 1.1.1. PROPRIÉTÉ LITTÉRAIRE ET ARTISTIQUE

##### a. Droit d'auteur

- **Droit d'auteur** : Une œuvre (texte, image, musique, etc.) est protégée s'il s'agit d'une création de forme originale. Aucun dépôt n'est requis. L'auteur bénéficie :
  - D'un **droit moral** qui est perpétuel, imprescriptible, inaliénable, mais transmissible notamment aux héritiers de l'auteur après sa mort. Il recouvre 4 prérogatives :
    - Le **droit de divulgation**, par lequel l'auteur décide de communiquer ou non son œuvre au public
    - Le **droit à la paternité**, obligeant que le nom et la qualité de l'auteur soient mentionnés sur tous les supports d'exploitation de l'œuvre
    - Le **droit au respect**, permettant à l'auteur de s'opposer à toute modification ou atteinte à son œuvre.
    - Le **droit au retrait et de repentir**, permettant à l'auteur, sous certaines conditions, de revenir sur sa décision de divulguer l'œuvre.

- De **droits patrimoniaux** qui sont limités dans le temps, pour la durée de la vie de l’auteur, plus 70 ans après sa mort. À l’issue de cette période, l’œuvre tombe dans le domaine public et peut être utilisée librement, sous réserve de respecter le droit moral de l’auteur. Ces droits patrimoniaux peuvent être cédés contre rémunération. Ils recouvrent :
  - Le **droit de reproduction** : permettant à l’auteur d’autoriser ou d’interdire toute fixation de son œuvre sur un support quelconque (papier, vidéo, enregistrement, etc.).
  - Le **droit de représentation** : permettant à l’auteur d’autoriser ou d’interdire toute communication de son œuvre au public par tout procédé (télévision, cinéma, réseaux sociaux, etc.).
- Une œuvre peut être créée par un ou plusieurs auteurs. En cas de pluralité d’auteurs, l’œuvre sera qualifiée d’œuvre de collaboration, d’œuvre collective ou d’œuvre dérivée (ou composite), chaque qualification entraînant l’application de règles spécifiques.
- Certaines **exceptions légales** permettent d’utiliser une œuvre sans autorisation, dans des cas bien précis (copie privée, fouilles de textes et de données, parodie, courte citation, cercle de famille, dépôt légal, etc.).
- L’auteur d’une œuvre graphique ou plastique bénéficie également d’un **droit de suite** lui conférant, sous certaines conditions, une rémunération sur une partie du prix de vente en cas de revente du support de l’œuvre.



**Point de vigilance** : Toute utilisation non autorisée d’une œuvre protégée par un droit d’auteur est une **contrefaçon**, sanctionnée **civilement** (dommages et intérêts notamment) et/ou **pénalement** (amende, emprisonnement, etc.).

**Pour en savoir plus** : Consultez le Guide de l’IA : §15 à §22

## b. Droits voisins du droit d’auteur

- **Droits voisins** : À côté des droits d’auteur, il existe des droits dits « voisins » du droit d’auteur, qui protègent les contributions de certaines personnes, que sont les artistes interprètes, les producteurs de phonogrammes ou de vidéogrammes, les chaînes de télévision ou encore les éditeurs de presse.
  - Chacun bénéficie d’un **monopole d’exploitation** pour une durée limitée.
  - Seuls les artistes-interprètes bénéficient d’un droit moral.
  - Voici un tableau récapitulatif :

Qualité	Définition	Droit moral	Droits patrimoniaux	Durée de protection des droits patrimoniaux
<b>Artiste-interprète</b>	La personne qui représente, chante, récite, déclame, joue ou exécute une œuvre de l’esprit	<b>Oui</b> , pour le respect de son nom, de sa qualité et de son interprétation.	Le droit d’autoriser ou d’interdire toute exploitation de l’interprétation, sauf exceptions	<b>50 ans</b> (à compter de son interprétation – 70 ans pour la musique).
<b>Producteur de phonogramme</b>	La personne, physique ou morale, qui a l’initiative et la responsabilité de la première fixation d’une séquence de sons.	<b>Non</b>	Le droit d’autoriser ou d’interdire toute exploitation du phonogramme, sauf exceptions.	<b>70 ans</b> (à compter de la mise à disposition du public du phonogramme).

<b>Producteur de vidéogrammes</b>	La personne, physique ou morale, qui a l'initiative et la responsabilité de la première fixation d'une séquence d'images, sonorisée ou non.	<b>Non</b>	Le droit d'autoriser ou d'interdire toute exploitation du vidéogramme, sauf exceptions.	<b>50 ans</b> (à compter de la mise à disposition du public du vidéogramme).
<b>Entreprises de communication audiovisuelle</b>	Essentiellement les chaînes de TV et radios.	<b>Non</b>	Le droit d'autoriser ou d'interdire toute exploitation de leurs programmes, sauf exceptions.	<b>50 ans</b> (à compter de la première communication au public des programmes).
<b>Éditeurs et agences de presse</b>	Les personnes qui éditent et publient des publications de presse (journaux en ligne, etc.).	<b>Non</b>	Le droit d'autoriser ou d'interdire les exploitations en ligne de leurs publications, sauf exceptions.	<b>2 ans</b> (à compter de la publication de presse concernée).



**Point de vigilance** : Toute utilisation d'une interprétation, d'un phonogramme, d'un vidéogramme ou d'un programme audiovisuel, sans autorisation du titulaire de droits, constitue une **contrefaçon**, sanctionnée **civilement** (dommages et intérêts notamment) et/ou **pénalement** (amende, emprisonnement, etc.).

**Pour en savoir plus** : Consultez le Guide de l'IA §23 à §30

### c. Droits *sui generis* du producteur de bases de données

- **Définition** : Le producteur d'une base de données est la personne qui prend l'initiative et le risque des investissements (financiers, matériels ou humains) pour la constitution de ladite base de données.
- Le producteur de la base de données bénéficie d'un droit spécial, complémentaire du droit d'auteur, portant non pas sur l'architecture de la base et/ou les données qu'elle contient, mais sur les investissements (financiers, matériels ou humains) substantiels liés à l'obtention, la constitution, la vérification ou la présentation du contenu de la base, les données qui y sont intégrées étant protégées en tant qu'ensemble et non de manière individuelle.
- **Monopole** : Ce droit permet au producteur d'autoriser ou d'interdire toute extraction ou réutilisation non autorisée de la base de données, ou de ses parties importantes.
- **Durée** : Le droit *sui generis* de la base de données a une durée de 15 ans à compter de la date d'achèvement de la base de données, renouvelable pour la même durée en cas de mise à jour significative.



**Point de vigilance** : En cas d'atteinte aux investissements substantiels du producteur de base de données, celui-ci peut agir en justice pour obtenir des **dommages et intérêts**. Dans certains cas, des sanctions pénales sont également applicables.

**Pour en savoir plus** : Consultez le Guide de l'IA §31 à §33

## 1.1.2. PROPRIÉTÉ INDUSTRIELLE

	Définition	Droits	Durée
<b>Marques</b>	Le droit des marques protège, sous certaines conditions les signes distinctifs (nom, logo, couleur, son, etc.) identifiant des produits et/ou services d'un opérateur économique.	Droit exclusif d'exploitation, permettant, sous certaines conditions et sauf exceptions, d'autoriser ou d'interdire toute exploitation, dans la vie des affaires, d'un signe identique ou similaire pour des produits et/ou des services identiques ou similaires, dans le territoire dans lequel la marque a été déposée.	<b>10 ans</b> (à compter de la date de dépôt de la marque), <b>renouvelable indéfiniment.</b>
<b>Dessins et modèles</b>	Le droit des dessins et modèles protège, sous certaines conditions, l'apparence d'un produit, ou d'une partie de ce produit, caractérisée en particulier par ses lignes, ses contours, ses couleurs, sa forme, sa texture ou ses matériaux.	Droit exclusif d'exploitation, permettant, sous certaines conditions et sauf exceptions, d'autoriser ou d'interdire l'exploitation du dessin ou modèle, dans le territoire dans lequel le dessin ou modèle a été déposé.	<b>5 ans</b> (à compter de la date de dépôt du dessin ou modèle), <b>renouvelable 5 fois.</b>
<b>Brevets</b>	Le droit des brevets protège, sous certaines conditions, les inventions techniques nouvelles résultant d'une activité inventive et qui sont susceptibles d'application industrielle.	Droit exclusif d'exploitation, permettant, sous certaines conditions et sauf exceptions, d'autoriser ou d'interdire l'exploitation du produit ou du procédé breveté, sur le territoire dans lequel le brevet a été déposé.	<b>20 ans</b> (à compter de la date de dépôt).



**Point de vigilance** : Marques, dessins et modèles et brevets sont des **titres de propriété industrielle** délivrés par un office compétent (INPI, EUIPO, etc.), valables pour la zone géographique concernée (France, UE, etc.) et pour la durée de protection susvisée.

À l'expiration de la durée de protection, le monopole cesse de produire ses effets : il n'est alors plus nécessaire de demander une autorisation préalable auprès des titulaires de droits.

Tandis que, pendant la durée de protection de chacun de ces titres, toute atteinte à l'un de ces droits de propriété intellectuelle constitue une **contrefaçon**.

**Pour en savoir plus** : Consultez le Guide de l'IA :

Pour les marques : §35 à §37

Pour les dessins et modèles : §38 à §40

Pour les brevets : §41 à §44

## 1.2. AUTRES PROTECTIONS POSSIBLES

Les données patrimoniales peuvent être protégées par d'autres droits, dont le droit des données personnelles (1.2.1), les droits de la personnalité (1.2.2), le droit de la concurrence (1.2.3) ou encore le secret des affaires (1.2.4).

### 1.2.1. DONNÉES PERSONNELLES

- Lorsque les données patrimoniales contiennent des informations relatives à une personne physique identifiée ou identifiable, elles sont protégées par le droit des données personnelles.
- La personne concernée bénéficie notamment du droit d'accéder à ses données, de les rectifier, de s'opposer à certains traitements ou encore de demander leur suppression.
- Certaines données dites « sensibles » (origine, santé, etc.) sont très encadrées. Leur traitement est par principe interdit, sauf exceptions.
- Lorsqu'une personne souhaite collecter et traiter certaines données personnelles dont elle détermine les finalités du traitement, celle-ci doit respecter un certain nombre d'obligations imposées par la Loi informatique et libertés et le règlement européen sur la protection des données personnelles (RGPD).



**Point de vigilance** : Le traitement des données personnelles suppose qu'elles soient collectées de manière licite (consentement de la personne concernée, intérêt légitime du responsable de traitement, etc.). Par exemple, les traitements de données personnelles fondés sur le consentement explicite de la personne concernée, ou qui concernent des données personnelles ayant été manifestement rendues publiques par la personne concernée, sont licites.

**Pour en savoir plus** : Consultez le Guide de l'IA : §46 à §48

### 1.2.2. DROITS DE LA PERSONNALITÉ

- Les droits de la personnalité sont inhérents à la personne humaine. Ils protègent la vie privée de la personne, son image, sa voix, son nom... Ainsi, ils interdisent toute atteinte injustifiée aux droits les plus fondamentaux des personnes : dignité, corps humain, vie privée, etc.
- La violation du droit au respect à la vie privée peut entraîner une condamnation au paiement de dommages et intérêts.



**Point de vigilance** : Les droits de la personnalité s'éteignent à la mort de la personne concernée et ne sont pas transmissibles aux héritiers. Les proches du défunt ne peuvent agir que s'ils prouvent l'existence d'un préjudice personnel, spécialement en cas d'atteinte à leurs propres droits, ou à la mémoire du défunt.

**Pour en savoir plus** : Consultez le Guide de l'IA : §49 à §50

### 1.2.3. DROIT DE LA CONCURRENCE

- Les données ont une valeur stratégique pour les systèmes d'IA. Leur accès peut être un enjeu concurrentiel dès lors que la qualité des données d'entrée déterminera la performance du système d'IA.
- Ces enjeux créent, *de facto*, une compétition entre les différents fournisseurs de systèmes d'IA.
- Le droit interdit cependant les abus de position dominante, les ententes, et les comportements déloyaux.
- L'Autorité de la concurrence surveille de près ces pratiques.



**Le saviez-vous ?** Dans un avis public du 28 juin 2024, l'Autorité de la concurrence recommande de vérifier que le marché de la donnée se construit en assurant un « *équilibre entre juste rémunération des ayants droit et accès des développeurs de modèles aux données nécessaires pour innover en prenant en compte la diversité des cas d'usage* »

**Pour en savoir plus** : Consultez le Guide de l'IA : §51 à §52

#### 1.2.4. SECRET DES AFFAIRES

- Certaines données patrimoniales peuvent contenir des informations protégées par le secret, et notamment par le secret des affaires.
- Le détenteur légitime de cette information peut agir contre l'obtention ou la divulgation illicite de ce secret.
- Certaines exceptions existent (obligations légales, procédures judiciaires...).



**Le saviez-vous ?** Le secret des affaires peut ainsi être opposé en cas d'obtention, d'utilisation ou de divulgation illicites d'informations protégées.



**Point de vigilance** : Toute atteinte au secret des affaires engage la responsabilité civile de son auteur. Dans certains cas, des sanctions pénales sont également applicables.

**Pour en savoir plus** : Consultez le Guide de l'IA : §53 à §56

#### 1.2.5. DROITS DES DOCUMENTS ADMINISTRATIFS, DES ARCHIVES PUBLIQUES ET DES INFORMATIONS PUBLIQUES

- Les données patrimoniales (telles que les copies numériques des collections, les métadonnées associées aux documents sources ou numérisés) sont susceptibles d'être qualifiées cumulativement de documents administratifs, d'archives publiques et/ou d'informations publiques.
- Leur réutilisation peut donner lieu à l'établissement d'une licence, sous réserve du respect des droits des tiers (droits d'auteur, données à caractère personnel, etc.)



**Point de vigilance** : Toute personne réutilisant des informations publiques en violation des prescriptions prévues au code des relations entre le public et l'administration (CRPA) est susceptible d'être sanctionnée par une amende ou des dommages-intérêts.

**Pour en savoir plus** : Consultez le Guide de l'IA : §57 à §62

### 1.3. DONNÉES PATRIMONIALES PROTÉGÉES PAR CONTRATS

- Les données patrimoniales peuvent avoir été produites par une institution culturelle ou obtenues de différentes manières, telles qu'une acquisition à titre onéreux ou à titre gratuit (don ou legs), ou encore dans le cadre d'une commande ou d'un contrat de prestation de services.

- Les contrats conclus par l'institution culturelle pour l'exploitation ou l'acquisition des données patrimoniales peuvent comporter des clauses limitant, voire s'opposant à l'usage de ces données patrimoniales.

**À noter que** le transfert de propriété d'un support matériel (livre, tableau, cliché photographique, fichier sonore ou audiovisuel, etc.) n'emporte pas transfert des droits de propriété intellectuelle qui protègent l'œuvre qui y est incorporée, sauf si cela est expressément prévu dans le contrat.

- Dans ces cas, l'autorisation des titulaires de droits de propriété intellectuelle (auteurs, artistes-interprètes, producteurs, éditeurs...) doit être obtenue pour pouvoir utiliser ou exploiter leurs œuvres, objets ou bases de données.



**Le saviez-vous ?** Les autorisations des titulaires de droits de propriété intellectuelle peuvent être obtenues soit **en gestion individuelle**, c'est-à-dire en négociant de gré à gré avec chacun des titulaires de droits, soit **en gestion collective**, c'est-à-dire en négociant avec les organismes de gestion collective (OGC) qui représentent l'ensemble des titulaires de droits qui leur ont confié leurs droits.

**Pour en savoir plus :** Consultez le Guide de l'IA : §63 à §69

## 2. ENTRAÎNEMENT DES MODÈLES D'IA GÉNÉRATIVE

Les IA génératives nécessitent pour la plupart l'utilisation de données pour leur entraînement et leur développement. Ces données d'entraînement peuvent être constituées de données patrimoniales qui ne sont pas libres de droit car soumises à autorisation d'un ou plusieurs titulaires des droits précédemment identifiés.

### 2.1. DONNÉES D'ENTRAÎNEMENT

Les IA génératives sont entraînées grâce à des jeux de données ou « données d'entraînement ». Les données d'entraînement sont diverses (2.1.1) et doivent par ailleurs être identifiables pour assurer l'effectivité des réglementations applicables selon la nature des données d'entraînement utilisées (2.1.2).

#### 2.1.1. DIVERSITÉ DES DONNÉES D'ENTRAÎNEMENT

- **Pourquoi les données sont-elles nécessaires à l'entraînement des IA ?** Les systèmes d'IA générative ont besoin de très grandes quantités de données (textes, images, sons, vidéos, etc.) pour fonctionner. Ces données permettent « d'entraîner » les modèles d'IA, c'est-à-dire d'ajuster les paramètres des algorithmes pour qu'ils soient capables de produire des résultats pertinents.

- Trois types de données à distinguer :

Données utilisées pour l'entraînement des IA		
Données d'entraînement	Données de validation	Données de test
Données utilisées pour former le modèle en ajustant ses paramètres.	Données utilisées pour évaluer le modèle en cours de formation et ajuster les réglages secondaires.	Données utilisées à la fin pour vérifier que le modèle fonctionne bien avant d'être mis en service ou commercialisé.

- **Origine des données** : Les données peuvent venir de nombreuses sources, en particulier de contenus accessibles en ligne (sites internet, bases de données, réseaux sociaux, etc.). Les **données patrimoniales** (ex. : archives, contenus culturels) peuvent faire partie de ces sources.
- **Données d'entrée et de sortie** : Traditionnellement, on distingue les **données d'entrée** (utilisées pour entraîner les IA) et les données de sortie (contenus générés par l'IA). Mais en pratique, cette frontière est poreuse : certaines données générées, généralement appelées données de résultat (outputs) peuvent être réutilisées comme données d'entrée (inputs) pour continuer à entraîner les IA.



**Le saviez-vous ?** Les données fournies par l'utilisateur (textes, sons, images, etc.) aux IA génératives, ainsi que les réponses générées par l'IA, sont le plus souvent réutilisées comme données d'entraînement des IA.

**Pour en savoir plus** : Consultez le Guide de l'IA : §70 à §73

### 2.1.2. OBLIGATIONS DE TRANSPARENCE RÉSULTANT DU RÈGLEMENT EUROPÉEN SUR L'INTELLIGENCE ARTIFICIELLE

- Il est difficile de comprendre comment les systèmes d'IA fonctionnent et quelles données (le cas échéant, patrimoniales) ont été utilisées pour les entraîner. Des outils techniques permettent toutefois de mieux tracer ces données. Le Règlement européen sur l'IA cherche à répondre à cette opacité en imposant des **obligations de transparence**.
- L'objectif du Règlement est d'encourager l'innovation tout en protégeant les droits fondamentaux au sein de l'Union européenne. Le Règlement prévoit une **régulation graduée selon le niveau de risques** : plus un système d'IA présente de risques pour les personnes ou la société, plus les exigences juridiques sont élevées. À cet égard, les systèmes d'IA générative ou à « usage général » peuvent appartenir à différentes catégories de risques, et sont donc soumis à des obligations spécifiques en fonction de leur qualification.
- Le Règlement porte sur les systèmes d'IA (et non sur les données elles-mêmes), il impose donc certaines obligations de transparence les concernant, notamment pour les IA génératives :
  - Les utilisateurs doivent être informés qu'ils interagissent avec une IA ;
  - Les contenus générés (texte, image, audio, vidéo) doivent être marqués comme produits ou modifiés par une IA, de manière lisible par des machines ;
  - En cas d'hypertrucage (ou « deepfake »), cela doit être clairement indiqué ;

Les fournisseurs de modèles d'IA à usage général doivent :

- Tenir une documentation technique à jour sur leur modèle,
- Respecter le secret des affaires et les droits de propriété intellectuelle,
- Mettre en place une politique conforme au droit d'auteur, notamment pour respecter le droit d'opposition des titulaires de droits (possibilité de s'opposer à l'utilisation de leurs œuvres pour l'entraînement des modèles).



**Point de vigilance** : Une obligation spécifique vise les fournisseurs de modèles d'IA à usage général : ils doivent publier un **résumé suffisamment détaillé des contenus utilisés pour entraîner leurs modèles**.

Ce résumé doit :

- Être complet sur le contenu (mais pas nécessairement technique),
- Indiquer les principaux jeux de données utilisés,
- Mentionner les autres sources de données (avec explication), pour permettre aux titulaires de droits d'identifier les usages et faire valoir leurs droits,
- Être rendu public.

Un modèle de résumé a été publié par la Commission européenne le 24 juillet 2025.

**Pour en savoir plus** : Consultez le Guide de l'IA : §74 à §82

## 2.2. CONDITIONS D'ACCÈS AUX DONNÉES PATRIMONIALES

Concernant les données patrimoniales relevant du dépôt légal (dont l'INA, la BnF et le CNC sont les organismes dépositaires) et protégées par un droit de propriété intellectuelle, celles-ci ne sont consultables que dans les emprises de ces organismes par des chercheurs dûment habilités. Tout accès en dehors des emprises ou par des personnes non accréditées, nécessite l'accord préalable et écrit de l'ensemble des titulaires de droits.

Concernant les contenus audiovisuels et radiophoniques produits ou coproduits par les sociétés nationales de programmes (France Télévisions, Radio France et France Médias Monde) dont les droits de propriété et/ou d'exploitation sont légalement dévolus à l'INA, ceux-ci sont accessibles aux professionnels en exécution de contrats de licence conclus avec l'INA et, pour une partie seulement, en ligne sur les services de l'INA, sous réserve du respect des conditions générales d'utilisation et de vente de ces services.

Concernant les données patrimoniales ne relevant pas du dépôt légal, conservées par les bibliothèques ou archives, leur accès est conditionné par le respect des droits des tiers, selon qu'elles sont protégées ou non par la propriété intellectuelle, par des secrets, des données à caractère personnel ou autres droits privatifs.

## 2.3. CONDITIONS D'UTILISATION DES DONNÉES PATRIMONIALES DONT L'ACCÈS EST LICITE

L'utilisation de données patrimoniales pour entraîner une IA générative doit respecter les règles en vigueur, notamment celles relatives à la propriété intellectuelle (2.3.1) et celles relatives aux autres mécanismes de protection (2.3.2).

### 2.3.1. CONDITIONS D'UTILISATION DES DONNÉES PATRIMONIALES PROTÉGÉES PAR UN DROIT DE PROPRIÉTÉ INTELLECTUELLE

#### a. Droit d'auteur (et droits voisins)

- L'utilisation de contenus protégés par des droits d'auteur (ou des droits voisins) est, sauf exception, soumise à l'autorisation préalable du titulaire des droits.
- L'utilisation d'œuvres ou d'autres objets protégés pour entraîner une IA suppose souvent de les reproduire (par copie notamment), ce qui relève du droit exclusif reconnu aux auteurs et titulaires de droits voisins.

- Dans le cas où l'autorisation préalable des auteurs serait nécessaire, l'on peut se demander si des exceptions légales pourraient néanmoins s'appliquer, dont notamment l'exception de fouille de textes et de données :
  - Les institutions publiques culturelles, qui ont un accès licite à leurs collections, devraient pouvoir bénéficier de l'exception dès lors que la fouille de textes et de données est mise en œuvre **à des fins de recherches scientifiques**.
  - **Pour une utilisation à d'autres fins que de recherches scientifiques** l'exception ne peut trouver à s'appliquer que si les titulaires de droits n'ont pas expressément refusé que leurs œuvres soient utilisées à des fins de fouille de textes et de données (appelé « opt-out »).



**Point de vigilance** : compte tenu des nombreuses réserves sur l'application de l'exception de fouille de textes et de données à l'entraînement de l'IA (failles du mécanisme de l'opt out, invalidité au regard du « test des trois étapes »...), l'utilisation d'œuvres (ou interprétations) pour entraîner des modèles d'IA devrait nécessiter l'autorisation préalable des titulaires de droits.

Une exception existe toutefois en faveur des institutions culturelles : l'exception de fouilles de textes et de données qui leur permet de réaliser des copies ou reproductions numériques d'œuvres (ou interprétations) à des fins de la recherche scientifique sans autorisation des titulaires de droits.

**Pour en savoir plus** : Consultez le Guide de l'IA : §83 à §102

#### b. Droit *sui generis* du producteur de bases de données

- Le producteur d'une base de données peut interdire l'extraction de sa base, même si elle est consultable.
- L'extraction et l'utilisation du contenu d'une base de données protégée au titre du droit *sui generis* pour entraîner un système d'IA générative devrait donc être considérée comme un acte relevant du monopole du producteur, et donc soumis à son autorisation préalable.
- L'exception de fouille de textes et de données devrait permettre aux institutions publiques culturelles d'extraire ou de faire des copies numériques de bases de données **à des fins de recherche scientifique**.

**Pour en savoir plus** : Consultez le Guide de l'IA : §103

#### c. Droit des marques

- La protection d'un signe par le droit des marques vise à éviter la confusion sur l'origine commerciale des produits et des services protégés par cette marque.
- L'utilisation d'une marque au sein d'un jeu de données d'entraînement d'un système d'IA, sans lien avec un produit ou un service, ne semble pas enfreindre le droit des marques dès lors que l'acte en cause ne devrait pas s'analyser comme un usage à titre de marque.
- Il en résulte que l'autorisation du titulaire d'une marque ne semble pas nécessaire si une marque est utilisée dans le cadre de l'entraînement d'un système d'IA.

**Pour en savoir plus** : Consultez le Guide de l'IA : §104 à §106

#### d. Droit des brevets et des dessins et modèles

- La protection d'un dessin ou modèle s'applique dans certaines situations précises, comme la fabrication ou la vente d'un produit incorporant un dessin ou un modèle.

- Si un dessin ou modèle est simplement utilisé pour entraîner une IA, cet usage ne devrait pas être considéré comme une atteinte aux droits du titulaire.



**Point de vigilance** : Certains juges ont déjà estimé qu'une simple reproduction visuelle, par exemple sur une affiche publicitaire, n'ayant pas pour objet de commercialiser le produit incorporant le modèle mais un produit différent, pouvait constituer un acte de contrefaçon (CA Paris, 27 novembre 2015, 13/21612). Mais même dans ce cas-là, ils ont rappelé que la protection s'appliquait surtout **au produit qui incorpore le dessin ou modèle**.

Dans tous les cas, il faut être vigilant au regard du droit d'auteur car un dessin ou modèle peut également être protégé par un droit d'auteur s'il est original.

- La reproduction d'un brevet enregistré au sein de données d'entraînement d'une IA ne devrait pas constituer, en soi, une atteinte aux droits du titulaire du brevet.
- En effet, les actes, dont l'utilisation est subordonnée à autorisation préalable, ont trait, comme en matière de dessins et modèles, au produit ou au procédé breveté lui-même.

### 2.3.2. CONDITIONS D'UTILISATION DES DONNÉES PATRIMONIALES PROTÉGÉES PAR D'AUTRES DROITS

#### a. Données à caractère personnel

- Entraîner une IA avec des données permettant d'identifier des personnes physiques est un traitement de données personnelles soumis à la réglementation en matière de données personnelles.
- Dans ce cas, le fournisseur du système d'IA est responsable de traitement au sens de la réglementation et doit notamment :
  - Justifier d'une base légale pour effectuer ce traitement, c'est-à-dire qu'il doit être fondé juridiquement (par exemple, sur le consentement explicite de la personne concernée, l'intérêt légitime du responsable de traitement, ou encore pour justifier d'une mission d'intérêt public),
  - Informer les personnes concernées que leurs données personnelles sont traitées dans le cadre d'un entraînement d'un système d'IA,
  - Assurer la sécurité des données personnelles,
  - Respecter les principes de minimisation et de transparence imposés par la réglementation.
- Pour entraîner une IA à partir de données contenant notamment des données à caractère personnel, il est possible d'échapper à l'application de la réglementation en matière de protection des données en procédant à une **anonymisation complète et irréversible** de ces données.



**Le saviez-vous ?** La CNIL publie des recommandations sur l'application du Règlement général sur la protection des données (RGPD) au développement des systèmes d'IA pour aider les professionnels à concilier innovation et respect des droits des personnes. Elles sont consultables sur le site internet de la CNIL.



**Point de vigilance** : Les personnes physiques ont également le droit de ne pas être soumises à une décision prise uniquement par une machine, sans qu'un humain n'intervienne (appelée « *décision fondée exclusivement sur un traitement automatisé, y compris le profilage* »).

Ce droit s'applique notamment si cette décision a des conséquences juridiques (par exemple : refus d'octroyer un prêt, licenciement) ou impacte fortement la vie de la personne.

Autrement dit, une IA ne peut pas, à elle seule, décider ; un humain doit toujours valider ou intervenir.

**Pour en savoir plus** : Consultez le Guide de l'IA : §112 à §119

## **b. Données protégées par des droits de la personnalité**

- Les droits de la personnalité, qui protègent les droits les plus fondamentaux de la personne, telle que le droit au respect à la vie privée sont protégés au même titre que les données personnelles. En effet, les données inhérentes à la personne humaine (sa vie, son corps, son image, sa voix, son nom, sa santé, ses croyances, etc.), sont par définition, des données à caractère personnel.
- Leur utilisation nécessite donc l'accord préalable de la personne concernée.

**Pour en savoir plus** : Consultez le Guide de l'IA : § 120 à § 121

## **c. Données protégées par le secret des affaires**

- Certaines données utilisées pour entraîner une IA peuvent être confidentielles, notamment parce qu'elles sont protégées par le secret des affaires.
- L'utilisation des données, y compris patrimoniales, couvertes par le secret des affaires est susceptible de constituer une violation des règles de confidentialité dès lors qu'elles sont communiquées à un tiers, sans autorisation.
- Ce risque n'existe pas seulement au stade des données d'entraînement, de test ou de validation, mais également au stade des prompts, ces derniers étant susceptibles de contenir des données protégées par la confidentialité ou le secret des affaires.

**Pour en savoir plus** : Consultez le Guide de l'IA : § 122 à § 123

## **d. Données protégées par le droit de la concurrence**

- Lorsqu'un opérateur économique impose des conditions injustes d'accès ou d'exploitation des données, il pourrait être sanctionné par les autorités pour **abus de position dominante**.

**Pour en savoir plus** : Consultez le Guide de l'IA : § 124 à § 127

## **e. Droit des informations publiques**

La réutilisation d'informations publiques peut donner lieu à l'établissement d'une licence, non exclusive. Elle est conditionnée au respect des droits des tiers (propriété intellectuelle, données à caractère personnel, secrets, etc.) Cette licence est en principe gratuite, sauf exceptions.

**Pour en savoir plus** : Consultez le Guide de l'IA : § 128 à § 136

### **2.3.3. CONDITIONS D'UTILISATION DES DONNÉES PATRIMONIALES PROTÉGÉES PAR CONTRATS**

- Certaines clauses d'un contrat peuvent limiter ou empêcher l'usage de données, y compris patrimoniales, à des fins d'entraînement d'une IA.
- Ces clauses peuvent figurer dans tout type de contrat, qu'il s'agisse :
  - d'un transfert gratuit (ex. : don ou legs),
  - ou d'un transfert onéreux (ex. : vente, licence, cession de droits, etc.).
- L'usage de données, y compris patrimoniales, nécessite de vérifier s'il existe ou non un obstacle de nature contractuel en restreignant les conditions d'usage.

Les autorisations nécessaires pour permettre l'entraînement des IA peuvent faire l'objet de contrats, afin de lever certains obstacles. Les contrats peuvent être conclus de gré à gré avec les titulaires de droits ou leurs représentants (en gestion individuelle) ou avec les organismes de gestion collective (en gestion collective) ayant reçu mandat (légal ou contractuel) de leurs membres pour délivrer ces autorisations.

**Pour en savoir plus** : Consultez le Guide de l'IA : §137 à §140

### 3. UTILISATION DES SYSTÈMES D'INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE

---

L'utilisation d'un système d'IA est soumise à des conditions générales d'utilisation proposées par le fournisseur de ce système d'IA auxquelles les utilisateurs sont soumis et au respect des droits des tiers sur les données intégrées dans les requêtes (prompts) ou en phase de RAG (3.1). À cet encadrement contractuel, s'ajoutent les règles susceptibles de s'appliquer aux contenus générés par les IA génératives (3.2).

#### 3.1. RESPECT DES CONDITIONS D'UTILISATION DES SYSTÈMES D'IA ET DES DROITS SUR LES DONNÉES AU STADE DE PROMPTS OU EN PHASE DE RAG

- Les IA génératives sont accessibles via des plateformes qui imposent des conditions générales d'utilisation (ci-après désignées « **CGU** ») à leurs utilisateurs. Ces CGU prennent la forme de contrats d'adhésion, non négociables, que l'utilisateur doit préalablement accepter pour pouvoir utiliser le service.
- Même si ces contrats sont imposés, ils doivent néanmoins respecter la réglementation en vigueur. En particulier, toute clause déséquilibrée est invalide et peut être réputée « non écrite » par les juges.
- Les CGU prévoient en général que l'utilisateur reste propriétaire des contenus qu'il fournit ou obtient (données d'entrée et de sortie), l'utilisateur devant en tout état de cause respecter les lois et règlements en vigueur (droits d'auteur, vie privée, etc.), notamment au stade des prompts.



**Point de vigilance** : les droits des tiers (propriété intellectuelle, confidentialité ou secrets, protection des données à caractère personnel etc.) ne doivent pas seulement être respectés lors de l'entraînement sur les données d'entrée mais également au stade des **prompts** (ces derniers pouvant contenir des données protégées) ou en phase de **RAG**.

**Pour en savoir plus** : Consultez le Guide de l'IA : §142 à §149

#### 3.2. RESPECT DES DROITS SUR LES RÉSULTATS GÉNÉRÉS PAR UN SYSTÈME D'IA GÉNÉRATIVE

Les IA génératives sont capables de produire des contenus relevant de genres différents (images, vidéos, textes, etc.). Ces contenus sont susceptibles, à certaines conditions, d'être eux-mêmes protégés par le droit, et notamment par le droit de la propriété intellectuelle.

##### 3.2.1. PROTECTION DES RÉSULTATS GÉNÉRÉS PAR UN SYSTÈME D'IA GÉNÉRATIVE

###### a. Protection par le droit d'auteur

- S'agissant du **droit d'auteur**, tout dépend de la manière dont l'homme interagit avec la machine. Un contenu généré par une IA ne devrait pas pouvoir bénéficier de la protection par le droit d'auteur en France, sauf à démontrer l'existence d'une **création intellectuelle humaine**.

- La protection d'un contenu généré par un système d'IA par le droit d'auteur est difficile à considérer en droit français, lorsque, par le prompt, l'IA générative fait partie intégrante et exclusive du processus créatif car elle conduit à la génération de contenus par nature aléatoires, échappant à la création humaine.
- En revanche, si l'IA générative est utilisée comme un simple outil technique pour l'une des étapes, initiale et/ou intermédiaire, du processus créatif, et qu'une création humaine a été réalisée en amont ou en aval du prompt, l'IA n'intervient alors que comme simple instrument d'assistance technique.



**Point de vigilance** : La question de la protection des contenus générés par une IA par le droit d'auteur n'a pas encore été tranchée en France. La loi ne dit mot et aucune décision judiciaire française n'est encore intervenue.

Dans la conception juridique française du droit d'auteur, celui-ci est attaché à la personne humaine. Il paraît donc difficile, en l'état du droit positif, de protéger un contenu exclusivement généré par un système d'IA par le droit d'auteur.



**Le saviez-vous ?** En **Chine**, les tribunaux reconnaissent la possibilité de protéger par le droit d'auteur des œuvres générées par IA lorsqu'une intervention humaine suffisamment créative est démontrée.

Par exemple, dans une affaire concernant un contenu généré par Midjourney puis modifié sur Photoshop, les juges ont retenu la protection de ce contenu par le droit d'auteur en considérant qu'« *à travers le choix libre et l'agencement uniques des scènes, de l'environnement, des couleurs, des jeux de lumières et d'ombre, ainsi que des angles et de la composition* », cette œuvre manifestait « *l'originalité de l'auteur* » (Changshu Judicial Court, Decision No. Su 0581 Min Chu 6697, 28 octobre 2024, aff. « Cœur à Cœur »).

Aux **États-Unis**, l'approche est plus prudente. Les juges américains considèrent qu'une IA ne peut pas être reconnue comme auteur d'une œuvre protégeable par le droit d'auteur. Ils considèrent que la seule utilisation de prompts ne suffit pas pour revendiquer un droit d'auteur sur une œuvre générée par IA. En revanche, une intervention humaine créative avant ou après la génération pourrait ouvrir droit à protection ; il faut donc être en mesure d'identifier dans les résultats une contribution humaine « suffisante ».

**Pour en savoir plus** : Consultez le Guide de l'IA : § 152 à § 159

## b. Protection par un droit de propriété industrielle

- **Droit des dessins et modèles** : un visuel généré par IA pourrait être protégé par le droit des dessins et modèles s'il reflète notamment un apport créatif humain.
- **Droit des brevets** : une invention assistée par IA pourrait être protégée, si elle est nouvelle, inventive et susceptible d'application industrielle, et si un inventeur humain ayant contribué à la réalisation de l'invention peut être identifié.
- **Droit des marques** : un logo ou un signe distinctif généré par IA peut être enregistré à titre de marque, sous réserve qu'il satisfasse aux conditions classiques de protection.

**Pour en savoir plus** : Consultez le Guide de l'IA : § 160 à § 162

### 3.2.2. VIOLATION DE DROITS DE TIERS

L'utilisation de l'intelligence artificielle est susceptible de violer les droits de tiers.

Les atteintes potentielles sont diverses (droits de propriété intellectuelle, données personnelles, droits de la personnalité, secret des affaires ou confidentialité, contrats, concurrence déloyale ou parasitaire, etc.).

Les actions adéquates pourront être mises en œuvre au titre de l'entraînement de données et du contenu lui-même généré.



**Point de vigilance** : Attention à ne pas porter atteinte aux droits des tiers !

Les contenus générés par IA ne peuvent être utilisés, sans vérification préalable que ces derniers ne portent pas atteinte à un ou plusieurs droits reconnus à des tiers, dont notamment :

- **Le droit d'auteur, un droit voisin du droit d'auteur, le droit des marques, le droit des dessins et modèles, etc.**
- **La vie privée, les données personnelles, le secret des affaires, etc.**

Par exemple, s'inspirer d'un style n'est pas sans risque (exemple : générer des images dans le style des « Studios Ghibli »). Le style n'est pas en soi protégeable au titre du droit d'auteur. Comme les idées, les concepts, les méthodes, le style est dit « *de libre parcours* » ; il ne peut faire l'objet d'une protection privative. Mais si le contenu généré par une IA générative dépasse la simple imitation d'un style générique et reproduit les caractéristiques originales et reconnaissables d'une œuvre de l'esprit protégée au titre du droit d'auteur, alors c'est l'empreinte de la personnalité de l'auteur qui est reprise et qui pourra être sanctionnée sur le terrain de la contrefaçon de droit d'auteur.

En cas d'atteinte aux droits de tiers, la responsabilité de la personne fautive pourra être engagée ; sachant que selon les cas, plusieurs personnes pourraient être considérées comme responsables (l'utilisateur ayant généré le prompt, le fournisseur du système d'IA ayant commis une faute dans la conception du système, etc.).



**Le saviez-vous ?** S'agissant des « deepfake » (hypertrucage), le Règlement européen sur l'IA encadre ces pratiques. Ainsi, les contenus manipulés par IA qui ressemblent à des images ou vidéos réelles (personnes, lieux, événements) et qui peuvent induire en erreur, doivent être clairement identifiés comme ayant été générés par une IA.

**Pour en savoir plus** : Consultez le Guide de l'IA : §163 à §171

