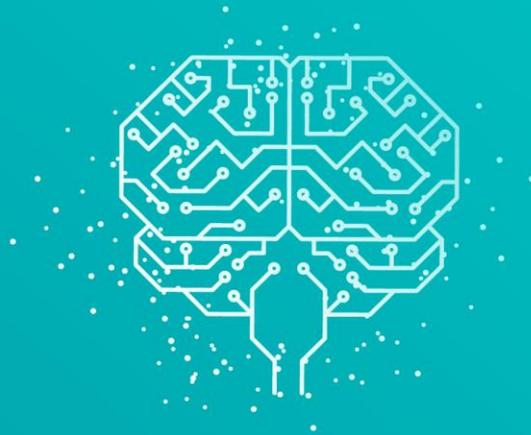
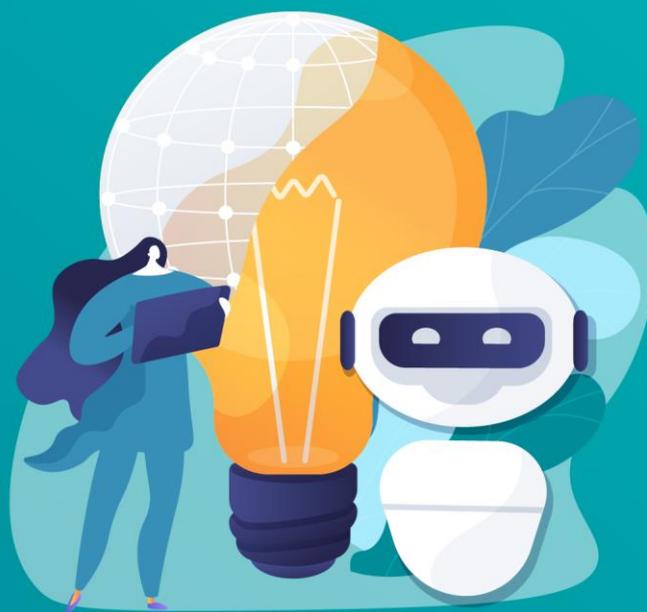




February 2022



# Study on copyright and new technologies: Copyright data management and artificial intelligence



## Executive Summary (EN)

In October 2020, DG CNECT commissioned a consortium of Technopolis Group, Philippe Rixhon Associates, UCLouvain, Crowell&Moring and IMC University of Applied Sciences with the execution of a 9-month “Study on copyright and new technologies: copyright data management and artificial intelligence”.

The study addresses two topics related to the copyright system and new technologies, dealt with in two dedicated parts of the study.

- The **first part** of the study systematically takes stock of the current situation with respect to rights metadata in different creative industries. It attempts to **identify and describe the economic impact of the current situation related to rights metadata**. It also compiles information on the **most important ongoing initiatives** to address some of the identified problems. Finally, the study indicates broad avenues which could contribute to improving functioning of the **copyright data ecosystem**.
- The **second part** of the study focuses on a) **uses** of copyright-protected content as input to feed AI technologies and b) the *copyright implications* of the **production of cultural outputs by or with the assistance of AI**. Furthermore, the study discusses possible **policy scenarios** which might be needed to react to these developments.

All information and views set out in this publication are those of the authors and do not necessarily reflect the official opinion of the Commission.

### Summary Part One

The **first part** of this project is focusing on the topic **rights metadata or “rights management information”** as defined in the European Directive 2001/29/EC on the harmonisation of copyright in the information society:

*“Rights Management Information means any information provided by rightsholders which identifies the work or other subject matter [ ], the author or any other rightsholder, or information about the terms and conditions of use of the work or other subject-matter, and any numbers or codes that represent such information. [This] shall apply when any of these items of information is associated with a copy of, or appears in connection with the communication to the public of, a work or other subject matter [ ]”<sup>1</sup>.*

It analyses the extent to which different **challenges** related to rights metadata can be empirically substantiated, including the availability of rights metadata attached to content, the interoperability between different systems for exchanging metadata, or the authority (i.e., trustworthiness) of sources. In doing so it provides supporting information in the context of the stocktaking document on developing improving the [Copyright Infrastructure](#) issued by the Council of the European Union under the Finnish Presidency

---

<sup>1</sup> European Community (2001). “Directive 2001/29/EC of the European Parliament and of the Council on the harmonisation of certain aspects of copyright and related rights in the information society”. Article 7 §2.

in December 2019 and the [Action Plan on Intellectual Property](#) adopted by the European Commission in November 2020.

The **empirical basis** for this part of the study consists of interviews and surveys among industry stakeholders as well as analysis of existing studies and further secondary research by the study team. We conducted more than 80 interviews covering three creative industries (film & television, music, and publishing, i.e., books, press, journals, and images) in different Member States of the European Union, Canada, the United Kingdom, and the United States. The interviewees were stakeholders of the digital value network in these creative sectors. The surveys were rolled out in spring 2021 and targeted the main trade associations of four creative industries (music, publishing, film, and TV broadcasting) as well as their members. The surveys consisted of two modules: one general part for industry experts without detailed insights into metadata challenges (mostly open questions on the respondents' perspective on the topic) and one detailed part targeted at metadata experts. Due to the high complexity of the topic, but also to the varying relevance of the topic in different industries, the number of respondents responding to the industry surveys differed widely (between 7 responses for broadcasting and 124 for publishing, with further item non-response for specific questions). Due to the low number of responses, quantitative estimations on the study topics were difficult to obtain. This led the study team to mainly concentrate on reporting qualitative indications.

For the **literature review**, we thoroughly analysed more than 20 core documents, including recent research and working papers, studies, position articles, and communications from the European Commission. In the selection process, we chose recent documents that cover different industries, technologies, and perspectives.

The study team additionally compiled an impressive **list of EU and industry initiatives** that currently play a role, or propose to play a role, in the **content rights infrastructure**, be it on issues of concrete data access and exchange (interoperability), global standards, and identifiers of parties or content or the overall governance of the copyright infrastructure. Any serious future action in the area of the copyright infrastructure must take these into account, either because (like the identifier and metadata standards and many of data access/exchange systems) they already play an essential and established role which must be integrated with others where necessary; or because (like frameworks, working groups and reports) they provide guidance, tools or potential tools for solving aspects of the interoperability challenge. The majority of these initiatives come from specific content or cultural sectors and are thus not cross-sectoral by design.

Key findings of the report suggest that rights data management is – for many reasons – challenging in all creative industries. Based on interviews, surveys and literature review, it can be summarised that the analysed creative sectors are facing data-related challenges in the following four areas:

- **Costs**, in the area of *rights management*
- **Efficiency issues**, in the domain of *licensing*
- **Challenges concerning payments processes**, in the field of *rights remuneration*
- **Risks of misappropriation and other rights infringements**, in the sphere of *rights enforcement*

To be able to describe the challenges in a nuanced way, the study team took a **sectoral approach** in analysing metadata challenges, holding a differentiated perspective on the film, music, and publishing industry. The reason for this is that there are large **differences between industries, but also within industries**, with respect to the identified

importance of metadata challenges and how high they are currently ranked on the industry or policy agenda. The same applies to the current status quo of metadata initiatives in the different industries.

As an example, in the music industry findings from existing evidence (information from e.g. CMOs) and expert interviews for this study among various stakeholder groups suggest that – while the work of many initiatives (for example CISAC and DDEX) has improved data exchange processes – challenges of imperfect or disputed rights metadata information are still a challenge. This implies, for example, that at least in some cases the music industry “spends an inordinate amount of time correcting errors and resolving disputes which hold up payments to music creators”<sup>2</sup>. This is less the case for newer recordings since awareness of the importance of rights metadata has improved, but in any case still seems to be an issue for older works. However, stakeholders from the recorded music sector also signalled that rights metadata issues are for them nowadays less problematic than, for example, the inaccuracy or lack of usage metadata provided by online platform services.

In the publishing sector, the results of interviews and surveys suggest that in areas such as the digital news publishing sector, challenges are more prevalent than in book publishing. In the former, issues such as a lack of granular attribution of copyright ownership for photographs used on news websites can lead to challenges regarding the remuneration of rightholders. In the latter, rightholder identification (e.g., the author of a book) is more straightforward and rights management processes therefore easier. The granularity and the degree to which copyright-protected content is embedded in creative works as well as rightholder structures of a creative work (through iterative contributions or co-authorships) make a substantial difference.

In the audiovisual sector, problems of compatibility of descriptive as well as rights management data were identified – despite recent developments to increase compatibility through a harmonisation of the registration process between the standard content identifiers developed within the industry, EIDR and ISAN. Also, a transparent exchange of usage metadata seems currently not always to be a given. Niche players in the film and television industry (independent film producers), but not the major studios we spoke to, raised concerns in this regard. In film production, stakeholders mention interoperability issues in rights management systems. Moreover, in most cases, there is no obligation to use the EIDR or ISAN standard identifiers and not all players are using them to identify their works. Such standard identifiers do not cover the rights, which does not facilitate licensing.

The study concludes that different avenues for future action could help improve the current situation: **raising copyright awareness** in general (e.g., on the side of creators and rightholders, and on the side of users and consumers) would help clarify the importance the copyright system in general for creatives in all industries. More specifically, initiatives to help raising **awareness and skills specifically related to rights metadata** seem to be important. The study results suggest that individuals in the creative content industries (especially creators themselves) have a relatively weak understanding of what metadata are and how to handle them. This lack of expertise and attention ends up in non-exhaustive metadata at various distribution stages of works and related subject matters. A “metadata education programme” could therefore be promoted in order to improve metadata right from the start of the creative process. In order to improve metadata quality and rights management processes the study authors also suggest to further take into account new

---

<sup>2</sup> <https://www.prsformusic.com/m-magazine/news/auddly-launches-metadatapays-campaign/>

technological developments such as Artificial Intelligence and Distributed Ledger technologies (for example via the European Blockchain Services Infrastructure). Finally, the study authors are of the opinion that, in the long run, a **cross-sector rights data network** could bridge gaps between standard content identifiers such as ISRC, ISWC, ISBN or ISAN and digital manifestations of the content they denote. This would increase interoperability also between different media or content sectors. The ultimate objective of this endeavour would be to break the silos between different creative industries and improve the efficient rights data management and licensing across sectors. It could help to release even more of the digital potential of Europe's creative sectors.

## Summary Part Two

The **second part** of the study analyses the implications for copyright and the related rights of the increasing use of AI technologies within the cultural and creative industries. Building on a review of several experiences with AI tools and concrete use cases, on a traditional study of the legal sources, on interviews with legal experts and stakeholders and lastly on a Delphi survey among legal experts and industry stakeholders, this part of the study focuses on the challenges raised by those tools and use cases for the EU copyright and related rights framework.

Over the last few years, AI solutions have been deployed across **different industries** and in a **wide range of applications**. The cultural and creative sector is no exception: some AI tools assist or complete the highly human process of creation; more often, they are used for improving the production of successful cultural artefacts or the consumer experience, e.g. through well-targeted recommendations. The reliance on AI technologies *for or during the creative process* might yet challenge copyright and/or related rights. The study distinguishes upstream or **input issues**, i.e. those related to the use of protected content as inputs for an AI application, from the downstream or **output issues**, i.e. those related to the musical, graphical, audiovisual or other cultural content that results from the use of the AI application. On the input side, AI applications may be trained with large datasets of creative content enjoying protection under copyright and related right, which prompts the question whether the rightholders' authorisation is needed for such use. On the output side, the AI applications can generate cultural content without any significant human contribution, which raises the question whether such outcome is protected under copyright or a related right. Other issues come to mind: Is there a need for additional incentives (in the form of copyright-like rights) to use AI tools for generating cultural outputs? Should the related investments in AI solutions be protected by an exclusive right or just promoted through funding? Are there authorship or ownership issues?

**The study is structured as follows.** It first shows how some AI applications are used in practice (part 3.2). The illustrations help to understand how this developing practice may impact the various stakeholders (creators, artists, producers, distributors, etc.) in the cultural and creative industry. This assessment is done in four creative sectors, namely visual arts, music, audiovisual & film, and video games. This allows then to identify possible issues with and challenges to copyright and related rights (part 3.3). Finally, some policy options are examined (part 3.4) to address these challenges.

Concerning the deployment of AI solutions in the cultural sector, the research demonstrates on the basis of concrete examples, that the overall reliance on these tools is increasing even if the **degree of adoption of AI solutions** varies significantly from sector to sector, for instance with a clear use of AI tools for upgrading video games, as well as for generating photos and faces for advertisements or "elevator music". Furthermore, AI solutions are not only deployed for repetitive or mechanical duties but also for tasks leading

to outputs which appear original and imaginative – and therefore, traditionally considered as within the sole realm of humans. The study focuses on the creations that are produced autonomously by the AI solution, with no or little human intervention.

Most of the AI applications appear to be **marketed online “as a service”** (one might even refer to the nascent market for “Creation as a Service” or “CaaS”). This business model confers factual control of the use of the AI solution and the production of AI output to the AI developers, who can consequently protect their revenues (and terminate the service in case of non-payment). The features of this business model should be considered when assessing the need for protecting the AI output under copyright or related rights.

On the **input side**, several challenges are examined, including the limits of the exclusivity conferred by copyright and the related rights.

Firstly, **the scope of the reproduction right** is still in the process of being defined by the European courts, especially when purely technical or intermediate copies are made such as within the process for training an AI algorithm through the analysis of protected elements. The recent teleological interpretation of the reproduction right and of the extraction right under the database right by the Court of Justice of the EU (CJEU) opens new avenues for the interpretation and application of those exclusive rights. It remains to be seen how the case law will interpret the notion of reproduction and whether it applies to intermediate and technical copies made during the process of feeding an AI tool. The scope of the reproduction right might indeed be resized so as to permit some uses that do not lead to an output in which the protected elements contained in the input are visible or audible.

Secondly, it appears from the consultation of the stakeholders that the expected application of the **TDM exceptions**, in particular of the TDM exception for other purposes than scientific research, raises some interrogations, especially concerning the way opt-out decisions should be communicated. In any case, the transposition of the TDM exceptions within the national laws should be carefully monitored to avoid diverging interpretations. Clarification as to the means and modes for expressing the opt-out under Article 4 DSM Directive might result from the development of good practices when the TDM exceptions will come into force.

Thirdly, the **moral rights** attached to copyright (and to the performers’ right) have not been harmonised at EU level. This could lead to diverging situations where some member states allow authors and performers to exercise their moral rights (in particular the right of integrity) to oppose the use of their works or performances as AI inputs. One way to address this is to clarify that the moral rights cannot block the application of harmonised exceptions (such as for TDM). A more ambitious approach would be to (partly) harmonise the moral rights.

On the **output side**, the AI-generated output is not protected under copyright **in the absence of human creative choices**. The research, interviews, and surveys conducted within the study indicate firstly that no incentive for the use of AI tools in the creative process in the form of additional exclusive rights appears necessary. The already broad deployment of AI tools in the creative context confirms this. Also, the feedback received seems to indicate that an additional right in favour of machine-generated outputs might have negative impacts on the traditional creative sectors. The study concludes that a new related right for AI-generated outputs is not desirable.

Secondly, even if advanced AI applications are increasingly capable of approximating the style of human-made works or performances, the scope of copyright should not be extended to offer protection to the **style** of an author or of a performer, unless some significant and recognizable features of a protected work or performance are reproduced in the AI output. The protection of a creator's style would indeed amount to a significant extension of copyright scope and would limit in a disproportionate way the freedom of expression and freedom of art downstream in the artistic process. Under national law, other remedies (e.g. image rights, personality rights or unfair practices) may be available and, some harmonisation at EU level of the claim of parasitism as an unfair commercial practice could be considered.

Thirdly, the absence of copyright protection for an AI autonomously generated output could leave **artistic performances** without protection under the law of certain member states, which may require that a 'work' protected under copyright be performed for the performance to be protected under the related rights. The human performance of an AI creation would not meet that condition and consequently the human performer would then be left without protection. To avoid this, the study suggests to consider a harmonised definition of 'performance' as the subject matter of the performers' right. This definition would not require the performed subject matter to be protected under copyright (while making sure that the scope of the performers' protection is not stretched to cover activities with little cultural interest).

Fourthly, even if autonomously generated outputs are excluded from copyright protection, they might, however, enjoy protection **under the related rights of phonogram producers, film producers, or press publishers** (if the outputs respectively qualify as sound or audiovisual fixation or as press publication), even if the production of an AI output has little to do with the traditional activities of producers or publishers and does not necessarily require a similar investment. To avoid that some of those related rights are used to circumvent the copyright policy trade-off, the study proposes to impose a condition of investment so that only the fixations for which a certain investment (possibly with a de minimis threshold) was made could trigger the application of the related rights of phonogram and film producers or publishers.

Fifthly, in the (still rare) cases of AI autonomously generated outputs, the **false attribution of authorship** to a human might, in practice, allow to circumvent the absence of copyright protection for this type of output. It might indeed suffice to claim authorship (by mentioning a person as the author) for the person to enjoy (unlawfully) the presumption of authorship and consequently, in fact, copyright protection over an AI-generated creation, knowing that this presumption is difficult or even impossible to rebut. However, the status quo concerning the authorship presumption should reasonably be maintained. A restriction or abolishment of this presumption would indeed be excessive and could have negative effects for the human creators on whom the burden of establishing authorship would lie.

Lastly, the study reflects on the evidence that the fact of knowing that an art piece is created by a human or by an AI system might affect the perception and the experience of the public. The study nevertheless points to the conclusion that no **information obligation** concerning the use of an AI solution for the development of the work should be added *within* the copyright framework. Such a legal obligation would indeed raise issues regarding its scope and its impact on the creators' artistic freedom and their personality rights. The study does not enquire about the adequacy of imposing an information obligation in other bodies of law, such as consumer law.

## Executive Summary (FR) / Résumé

En octobre 2020, la DG CNECT a confié à un consortium composé de Technopolis Group, Philippe Rixhon Associates, UCLouvain, Crowell&Moring et IMC University of Applied Sciences Krems la réalisation d'une « Étude sur le droit d'auteur et les nouvelles technologies : gestion des données du droit d'auteur et intelligence artificielle » d'une durée de 9 mois.

L'étude aborde deux sujets liés à l'écosystème du droit d'auteur et aux nouvelles technologies, traités en deux parties :

- La **première partie** de l'étude fait le point sur la situation actuelle de l'information en matière de droits dans différentes industries créatives. Elle tente d'**identifier et de décrire l'impact économique de la situation actuelle en ce qui concerne les métadonnées**. Elle compile également une **liste d'initiatives importantes** visant à résoudre certains des problèmes identifiés. Enfin, l'étude indique de grandes pistes qui pourraient contribuer à améliorer le fonctionnement de **l'écosystème des données sur les droits d'auteur**.
- La **deuxième partie** de l'étude se concentre sur a) **l'utilisation** de contenus protégés par le droit d'auteur pour alimenter les technologies d'IA et b) les *implications en matière de droit d'auteur* de la **production de biens culturels par ou avec l'aide de l'IA**. En outre, cette partie examine des **scénarios politiques** qui pourraient être nécessaires pour encadrer ces développements.

Toutes les informations et opinions présentées dans cette publication sont celles des auteurs et ne reflètent pas nécessairement l'opinion officielle de la Commission.

### Résumé de la première partie

La **première partie** de ce projet se concentre sur le sujet des **métadonnées de droits ou « informations sur le régime des droits »** telles que définies dans la directive européenne 2001/29/CE sur l'harmonisation de certains aspects du droit d'auteur et des droits voisins dans la société de l'information :

*« On entend par « information sur le régime des droits » toute information fournie par des titulaires de droits qui permet d'identifier l'œuvre ou autre objet protégé [ ], l'auteur ou tout autre titulaire de droits. Cette expression désigne aussi les informations sur les conditions et modalités d'utilisation de l'œuvre ou autre objet protégé ainsi que tout numéro ou code représentant ces informations. [Ceci] s'applique lorsque l'un quelconque de ces éléments d'information est joint à la copie ou apparaît en relation avec la communication au public d'une œuvre ou d'un objet protégé [ ] »<sup>3</sup>.*

Elle analyse en quelle mesure différents **défis** liés aux métadonnées de droits peuvent être empiriquement définis, notamment la disponibilité des métadonnées de droits et leur rattachement aux contenus, l'interopérabilité entre différents systèmes d'échange de

---

<sup>3</sup> Communautés européennes (2001). Directive 2001/29/CE du Parlement Européen et du Conseil sur l'harmonisation de certains aspects du droit d'auteur et des droits voisins dans la société de l'information. Article 7 §2.

métadonnées ou l'autorité (c'est-à-dire la fiabilité) des sources. Ce faisant, elle fournit des informations contextuelles se rapportant au [document d'inventaire sur l'amélioration de l'infrastructure du droit d'auteur](#) publié par le Conseil de l'Union européenne sous la présidence finlandaise en décembre 2019 et au [plan d'action en faveur de la propriété intellectuelle](#) adopté par la Commission européenne en novembre 2020.

La **base empirique** de cette partie de l'étude est constituée d'entretiens et d'enquêtes auprès des acteurs de l'industrie, d'une analyse des études existantes et de recherches supplémentaires menées par notre équipe. Nous avons eu plus de 80 entretiens couvrant trois industries créatives (cinéma et télévision, musique, et édition, c'est-à-dire livres, presse, revues scientifiques et images) dans les différents États membres de l'Union européenne, au Canada, au Royaume-Uni et aux États-Unis. Les interlocuteurs étaient des acteurs du réseau numérique de ces secteurs créatifs. Les enquêtes ont été menées au printemps 2021 et ciblaient les membres des principales associations professionnelles de quatre industries créatives (cinéma, télédiffusion, musique et édition). Les enquêtes se composaient de deux modules : une partie générale destinée aux experts du secteur qui ne disposent pas d'informations détaillées sur les défis liés aux métadonnées (essentiellement des questions sur le point de vue des répondants) et une partie plus détaillée destinée aux experts en métadonnées. En raison de la grande complexité du sujet, mais aussi de la variation de sa pertinence dans les différents secteurs, les nombres de répondants aux enquêtes sectorielles ont été très différents (entre 7 réponses pour la radiodiffusion et 124 pour l'édition). Vu le faible nombre de réponses, il a été difficile d'obtenir des estimations quantitatives sur les sujets étudiés. Cela a conduit les experts à se concentrer sur les rapports d'indications qualitatives.

Pour l'**examen de la littérature**, nous avons analysé de manière approfondie plus de 20 documents fondamentaux, y compris recherches récentes, documents de travail, études, articles de position, et communications de la Commission européenne. Nous avons choisi des documents récents qui couvrent différentes industries, technologies et perspectives.

L'équipe chargée de l'étude a également dressé une **liste non-exhaustive du nombre impressionnant d'initiatives de l'UE et de l'industrie** qui jouent actuellement un rôle, ou proposent de jouer un rôle, dans l'amélioration de l'infrastructure des métadonnées sur les droits, qu'il s'agisse d'actions concrètes sur l'accès et l'échange de données (interopérabilité), sur les identifiants des parties prenantes ou des contenus, ou sur la gouvernance de l'infrastructure. Toute action sérieuse dans le domaine de l'infrastructure des données du droit d'auteur doit tenir compte de ces initiatives, soit parce qu'elles jouent déjà un rôle essentiel qui doit être intégré à d'autres si nécessaire (comme les normes relatives aux identifiants, métadonnées et nombreux systèmes d'accès et d'échange de données), soit parce qu'elles fournissent des orientations, des outils ou des approches pour résoudre certains aspects d'interopérabilité (comme les cadres, les groupes de travail et les rapports). La majorité de ces initiatives viennent de secteurs spécifiques des médias et de la culture et ne sont donc pas intersectoriels.

Les principales conclusions du rapport suggèrent que la gestion des données relatives aux droits est, pour de nombreuses raisons, en chantier dans la plupart des industries créatives. Sur base des entretiens, des enquêtes et de l'analyse documentaire, on peut résumer que les secteurs créatifs analysés sont confrontés à des défis liés aux données dans les quatre domaines suivants :

- Les **coûts**, dans le domaine la *gestion informatique* des droits
- L'**efficacité**, dans le domaine de l'*octroi de licences*

- Les **procédures de paiement**, dans le domaine de la *rémunération des droits*
- Les risques de **détournement et autres atteintes aux droits**, dans le domaine de l'*application des droits*.

Pour décrire les défis de manière nuancée, les experts ont adopté une **approche sectorielle** différenciée dans l'analyse des industries du film, de la musique et de l'édition. En effet, il y a de grandes **disparités entre secteurs, et aussi au sein d'un même secteur**, en ce qui concerne l'importance et les priorités accordées aux métadonnées et à la place qu'elles occupent actuellement dans les agendas industriel ou politique. Il en va de même pour le statu quo des initiatives en matière de métadonnées dans les différents secteurs.

Par exemple, dans l'industrie de la musique, l'analyse de la situation, sur base d'informations communiquées par des sociétés de gestion collective et d'entretiens avec des parties prenantes d'horizons divers, suggère que trop d'informations sur les droits restent imparfaites ou contestées bien que le travail de nombreuses initiatives (par exemple DDEX) ait amélioré les processus d'échange des données s'y rapportant. Cela implique, par exemple, qu'au moins dans certains cas, l'industrie de la musique « passe un temps excessif à corriger des erreurs et à résoudre des litiges qui retardent les paiements aux créateurs et interprètes ». C'est moins le cas pour les enregistrements récents car la prise de conscience de l'importance des métadonnées s'est améliorée ; mais cela reste souvent un problème pour les œuvres plus anciennes. Par ailleurs, les représentants des maisons de disques ont signalé que les métadonnées de droits posent aujourd'hui moins de problèmes que, par exemple, l'inexactitude ou le manque de métadonnées d'utilisation fournies par les services de plateforme en ligne.

Dans le secteur de l'édition, les résultats des entretiens et des enquêtes suggèrent que les défis sont plus fréquents dans le secteur de la presse numérique que dans celui de l'édition de livres. Dans le premier cas, l'absence d'attribution granulaire des droits pour les photographies utilisées sur les sites d'information peut, par exemple, entraîner des difficultés quant à la rémunération des ayants-droits. Dans le second cas, l'identification des ayants-droits (par exemple, l'auteur d'un livre) est plus simple et la gestion des droits est donc plus facile. La granularité et la modularité du contenu protégé par le droit d'auteur ainsi que les relations entre ayants-droits d'une œuvre créée par contributions itératives et collaboratives, définissent le niveau de ces défis.

Dans le secteur de l'audiovisuel, nous avons identifié des problèmes d'incompatibilité des données nécessaires à la description des contenus et à la gestion des droits malgré les récents développements visant à harmoniser par exemple les processus d'enregistrement EIDR et ISAN, les identifiants de contenu standards développés dans l'industrie. De même, l'échange transparent de métadonnées d'utilisation ne semble pas toujours acquis à l'heure actuelle. Ce sont les producteurs indépendants et non les grands studios avec lesquels nous nous sommes entretenus qui nous ont fait part de leurs préoccupations à cet égard. Dans la production cinématographique, les problèmes d'interopérabilité des systèmes de gestion de droits sont mentionnés. De plus, dans la plupart des cas, il n'y a pas d'obligation d'utiliser les identifiants standards EIDR ou ISAN et tous les acteurs ne les utilisent pas pour identifier leurs œuvres. Ces identifiants standard ne couvrent pas les droits, ce qui ne facilite pas l'octroi de licences.

L'étude conclut que différentes pistes d'actions pourraient contribuer à améliorer la situation actuelle : la **sensibilisation au droit d'auteur** en général (du côté des créateurs et des ayants-droits autant que du côté des utilisateurs et des consommateurs) aiderait à

démontrer l'importance des droits d'auteur pour les créateurs de tout secteur. Plus précisément, les initiatives visant à **développer les compétences spécifiquement liées aux métadonnées des droits** semblent importantes. Les résultats de l'étude suggèrent que les acteurs du monde des médias et de la culture (en particulier les créateurs eux-mêmes) ont une compréhension relativement faible de ce que sont les métadonnées et de la façon de les traiter. Ce manque d'expertise et d'attention se traduit par des métadonnées non exhaustives à différents stades de la distribution des œuvres et de leurs représentations. Un « programme d'éducation » pourrait donc être mis en place pour améliorer la qualité des métadonnées dès le début du processus de création. Afin d'améliorer la qualité des métadonnées et les processus de gestion de droits, les auteurs de l'étude suggèrent également de prendre en compte les nouveaux développements technologiques tels que l'intelligence artificielle et les technologies de registres distribués (par exemple via l'infrastructure européenne de services blockchain). Enfin, les auteurs sont d'avis qu'un **réseau transsectoriel de données sur les droits** pourrait relier identifiants standards tels que ISRC, ISWC, ISBN ou ISAN et les manifestations numériques du contenu qu'ils désignent. L'accessibilité et l'interopérabilité s'en trouveraient grandement accrues entre les différents secteurs des médias et de la culture. L'objectif ultime de cette entreprise serait de briser les cloisonnements entre différents secteurs et juridictions et ainsi d'améliorer l'efficacité de la gestion des données relatives aux droits et à l'octroi de licences. Elle pourrait contribuer à libérer encore davantage le potentiel numérique des secteurs créatifs européens.

## Résumé de la deuxième partie

La **deuxième partie** de l'étude analyse les effets sur le droit d'auteur et les droits voisins de l'utilisation croissante des technologies d'IA dans les secteurs culturels et de la création. S'appuyant sur un examen de plusieurs expériences avec des outils d'IA et des cas d'utilisation concrets, sur une étude traditionnelle de sources juridiques, sur des entretiens avec des experts juridiques et des parties prenantes et enfin sur une étude *delphi*, cette partie de l'étude se concentre sur les défis amenés par ces outils et cas d'utilisation pour le cadre européen du droit d'auteur et des droits voisins.

Au cours des dernières années, des solutions d'IA ont été déployées dans **différents secteurs** et dans un large éventail d'applications. Le secteur culturel et de la création n'y fait pas exception : certains outils d'IA assistent ou complètent le processus éminemment humain de la création ; ils sont généralement utilisés pour améliorer la production d'objets culturels ou l'expérience du consommateur, à travers des recommandations ciblées par exemple. Le recours aux technologies de l'IA pour ou pendant le processus de création peut néanmoins soulever des difficultés en matière de droits d'auteur et/ou de droits voisins. L'étude fait la distinction entre les **problématiques en amont** ou *d'input*, à savoir celles liées à l'utilisation de contenus protégés comme *inputs* pour une application d'IA, et les **problématiques en aval** ou *d'output*, à savoir celles liées aux contenus musicaux, graphiques, audiovisuels ou autres contenus culturels qui résultent de l'utilisation de l'application d'IA. En ce qui concerne l'*input*, les applications d'IA peuvent être entraînées avec de grands ensembles de données de contenu créatif bénéficiant de la protection du droit d'auteur et des droits voisins, ce qui soulève la question de savoir si l'autorisation des titulaires de ces droits est requise pour une telle utilisation. En ce qui concerne l'*output*, les applications d'IA peuvent générer du contenu culturel sans aucune contribution humaine significative, ce qui soulève la question de savoir si ces produits sont protégés par le droit d'auteur ou un droit voisin. D'autres questions se posent : Faut-il prévoir

des incitations supplémentaires (sous la forme de droits similaires au droit d'auteur) pour utiliser les outils d'IA afin de générer des produits culturels ? Les investissements dans les solutions d'IA doivent-ils être protégés par un droit exclusif ou simplement encouragés par des financements ? Y a-t-il des problèmes de paternité ou de propriété ?

**L'étude est structurée comme suit.** Elle présente d'abord comment certaines applications d'IA sont utilisées dans la pratique (partie 3.2). Les illustrations permettent de comprendre comment cette pratique en développement peut affecter les différentes parties prenantes (créateurs, artistes, producteurs, distributeurs, etc.) de l'industrie culturelle. Cette analyse est réalisée dans quatre secteurs de la création, à savoir les arts visuels, la musique, l'audiovisuel & le cinéma, et les jeux vidéo. Cela permet ensuite d'identifier les potentiels problèmes et défis liés au droit d'auteur et aux droits voisins (partie 3.3). Enfin, certaines options de politiques sont examinées (partie 3.4) pour relever ces défis.

En ce qui concerne le déploiement de solutions d'IA dans le secteur culturel, la recherche démontre, sur la base d'exemples concrets, que ces outils sont utilisés plus fréquemment, même si le **degré d'adoption des solutions d'IA** varie considérablement d'un secteur à l'autre, avec par exemple une utilisation manifeste des outils d'IA pour l'amélioration des jeux vidéo, ainsi que pour la création de photos et de visages pour les publicités ou de « musique d'ascenseur ». En outre, les solutions d'IA ne sont pas seulement déployées pour des tâches répétitives ou mécaniques, mais aussi pour des tâches aboutissant à des résultats qui se révèlent originaux et imaginatifs - et donc traditionnellement considérés comme relevant du seul domaine des humains. L'étude focalise sur les créations produites de façon autonome par les outils d'IA, sans intervention humaine significative.

La plupart des applications d'IA semblent être **commercialisées en ligne « en tant que service »** (un marché naissant de la « *Creation as a Service* » ou « CaaS » peut être observé). Ce modèle commercial attribue le contrôle factuel de l'utilisation de la solution d'IA et de la production de produits d'IA aux développeurs d'IA, qui peuvent ainsi protéger leurs revenus (et mettre fin au service en cas de non-paiement). Les caractéristiques de ce modèle commercial devraient être prises en compte lors de l'évaluation de la nécessité de protéger le produit de l'IA par le droit d'auteur ou les droits voisins.

En ce qui concerne **l'input**, plusieurs défis sont examinés, notamment les limites de l'exclusivité conférée par le droit d'auteur et les droits voisins.

Tout d'abord, le **champ d'application du droit de reproduction** est toujours en cours de clarification par les juridictions européennes, notamment lorsque des copies purement techniques ou intermédiaires sont réalisées, comme dans le cadre du processus d'apprentissage d'un algorithme d'IA par l'analyse d'éléments protégés. L'interprétation téléologique récente du droit de reproduction et du droit d'extraction dans le cadre du droit des bases de données par la Cour de Justice de l'UE (CJUE) ouvre de nouvelles voies pour l'interprétation et l'application de ces droits exclusifs. Il reste à voir comment la jurisprudence interprétera la notion de reproduction et si elle s'applique aux copies intermédiaires et techniques réalisées au cours du processus d'alimentation d'un outil d'IA. Le champ d'application du droit de reproduction pourrait en effet être réaménagé de manière à permettre certaines utilisations qui ne

conduisent pas à un *output* dans lequel les éléments protégés contenus dans l'*input* sont visibles ou audibles.

Deuxièmement, il ressort de la consultation des parties intéressées que l'application prévue des **exceptions au TDM**, en particulier de l'exception au TDM à des fins autres que la recherche scientifique, soulève certaines interrogations, notamment en ce qui concerne la manière dont les décisions d'*opt-out* doivent être communiquées. En tout état de cause, la transposition des exceptions au TDM dans les législations nationales doit être surveillée de près pour éviter des interprétations divergentes. La clarification des moyens et des modes d'expression de l'*opt-out* en vertu de l'article 4 de la directive DSM pourrait résulter du développement de bonnes pratiques lorsque les exceptions au TDM entrent en vigueur.

Troisièmement, les **droits moraux** attachés au droit d'auteur (et au droit des artistes-interprètes) n'ont pas été harmonisés au niveau de l'UE. Cela pourrait conduire à des situations divergentes dans lesquelles certains États membres autorisent les auteurs et les artistes-interprètes à exercer leurs droits moraux (en particulier le droit à l'intégrité) pour s'opposer à l'utilisation de leurs œuvres ou de leurs prestations comme *input* de l'IA. Une façon de résoudre ce problème est de prévoir que les droits moraux ne peuvent pas empêcher l'application d'exceptions harmonisées (comme pour le TDM). Une approche plus ambitieuse consisterait à harmoniser (partiellement) les droits moraux.

En ce qui concerne l'**output**, les produits générés par l'IA ne sont pas protégés par le droit d'auteur **en l'absence de choix créatifs humains**. Les recherches, entretiens et enquêtes menés dans le cadre de l'étude indiquent premièrement qu'aucune incitation à l'utilisation d'outils d'IA dans le processus de création, sous la forme de droits exclusifs supplémentaires, ne semble nécessaire. Le déploiement déjà large des outils d'IA dans le contexte de la création le confirme. Par ailleurs, les réactions reçues semblent indiquer qu'un droit supplémentaire en faveur des produits générés par des machines pourrait avoir des effets négatifs sur les secteurs traditionnels de la création. L'étude conclut qu'un nouveau droit connexe pour les produits générés par l'IA n'est pas souhaitable.

Deuxièmement, même si les applications avancées de l'IA sont de plus en plus capables de se rapprocher du style d'œuvres ou de prestations créées par les humains, le champ d'application du droit d'auteur ne devrait pas être étendu pour offrir une protection au **style** d'un auteur ou d'un artiste-interprète, à moins que certaines caractéristiques significatives et reconnaissables d'une œuvre ou d'une prestation protégée ne soient reproduites dans le produit de l'IA. La protection du style d'un créateur équivaldrait en effet à une extension significative du champ d'application du droit d'auteur et limiterait de manière disproportionnée la liberté d'expression et la liberté artistique en aval du processus artistique. En vertu du droit national, d'autres moyens (par exemple, les droits à l'image, les droits de la personnalité ou les pratiques déloyales) sont envisageables et une certaine harmonisation au niveau européen de l'action en parasitisme en tant que pratique commerciale déloyale pourrait être envisagée.

Troisièmement, l'absence de protection du droit d'auteur pour un produit généré de manière autonome par l'IA pourrait priver certaines **prestations artistiques** de toute protection en vertu de la législation de certains États membres, qui exigent parfois qu'une « œuvre » protégée par le droit d'auteur soit interprétée pour que la prestation

soit protégée par les droits voisins de l'artiste interprète. L'interprétation humaine d'une création de l'IA ne remplirait pas cette condition, de sorte que l'artiste-interprète humain se retrouverait sans protection. Pour éviter cela, l'étude suggère d'envisager une définition harmonisée de « l'interprétation » comme objet du droit des artistes-interprètes. Cette définition n'exigerait pas que l'objet de l'exécution soit protégé par le droit d'auteur (tout en veillant à ce que le champ d'application de la protection des artistes-interprètes ne soit pas étendu à des activités présentant un faible intérêt culturel).

Quatrièmement, même si les produits générés de manière autonome sont exclus de la protection du droit d'auteur, ils pourraient toutefois bénéficier d'une protection **au titre des droits voisins des producteurs de phonogrammes, des producteurs de films ou des éditeurs de presse** (si les produits peuvent être qualifiés respectivement de fixation sonore ou audiovisuelle ou de publication de presse), même si la production d'un produit d'IA ne présente que peu de rapport avec les activités traditionnelles des producteurs ou des éditeurs et ne nécessite pas nécessairement un investissement similaire. Afin d'éviter que certains de ces droits voisins ne soient utilisés pour contourner l'arbitrage sous-jacent du droit d'auteur, l'étude propose une condition expresse d'investissement, de sorte que seules les fixations pour lesquelles un certain investissement (éventuellement avec un seuil minimum) a été réalisé puissent déclencher l'application des droits voisins des producteurs de phonogrammes et de films ou des éditeurs.

Cinquièmement, dans les cas (encore rares) de produits générés de manière autonome par l'IA, **la fausse attribution de la paternité** à un humain pourrait, en pratique, permettre de contourner l'absence de protection du droit d'auteur pour ce type d'output. Il pourrait en effet suffire de revendiquer la paternité (en mentionnant une personne comme auteur) pour que celle-ci bénéficie (illicitement) de la présomption de paternité et donc de fait de la protection du droit d'auteur sur une création générée par l'IA, sachant que cette présomption est difficile, voire impossible à réfuter. Toutefois, le *statu quo* concernant la présomption de paternité devrait raisonnablement être maintenu. Une restriction ou une suppression de cette présomption serait en effet excessive et pourrait avoir des effets négatifs pour les créateurs humains à qui incomberait la charge de démontrer la paternité.

Enfin, l'étude analyse le fait que savoir qu'une œuvre d'art est créée par un humain ou par un système d'IA peut affecter la perception et l'expérience du public. L'étude aboutit néanmoins à la conclusion qu'aucune **obligation d'information** concernant l'utilisation d'une solution d'IA pour le développement d'une œuvre ne devrait être ajoutée à la réglementation en matière de droit d'auteur. Une telle obligation légale susciterait en effet des questions quant à son champ d'application et à son incidence sur la liberté artistique et les droits de la personnalité des créateurs. L'étude ne porte pas sur l'opportunité d'imposer une obligation d'information dans d'autres branches du droit, comme le droit de la consommation.