**Council of the
European Union**

Brussels, 20 December 2019
(OR. en)

15016/19

PI 156

NOTE

From:	Presidency
To:	Working Party on Intellectual Property (Copyright)
No. prev. doc.:	13515/19
Subject:	Developing the Copyright Infrastructure
	- Stocktaking of work and progress under the Finnish Presidency

Delegations will find attached a note that summarises and takes stock of inputs collected, over the past six months, from Member States and from stakeholders in the creative sectors, on the Finnish Presidency priority 'Developing the copyright infrastructure'. This note also draws on the outcome of the Data Economy Conference that took place in Helsinki on 25-26 November 2019.

Section IV of the note ('Stocktaking of the Presidency') includes the Finnish Presidency's suggestions for possible next steps for developing the copyright infrastructure in the European Union.

———————————

DEVELOPING THE COPYRIGHT INFRASTRUCTURE

- Stocktaking of work and progress during the Finnish EU Presidency

I.   INTRODUCTION

The priorities of the Finnish Presidency in the copyright field focused on the sustainable growth and competitiveness of the Union. At the dawn of a new decade, the use and practices regarding rights management information – i.e. data to identify works and authors, generally referred to as metadata – is considered highly topical and a way to support unleashing the digital potential of Europe's creative sectors. In particular, the Presidency aimed at drawing attention to the challenges related to the efficient use of such data for the identification of works and authors in the digital environment.

In general terms, the problems relating to licensing markets in Europe are caused by the fact that works and rightholders are often not being properly identified in the digital environment. This is due to missing or erroneous metadata. There are many International Standardisation Organisations (ISO) and industry-specific identifiers for works and rightholders, however they lack interoperability in a broader context. This results in high transaction costs for the industry and lost revenue for rightholders.

Investing in the copyright data infrastructure, encompassing standardised metadata entries in connection with digital copies of works and where relevant, registrations with Collective Management Organisations (CMOs), can/would improve the efficiency of licensing, and allow automated processes for distribution of revenue streams. Although the situation has improved a lot in certain sectors, such as music and publishing, the copyright infrastructure is not yet fully functional and there are still no overall best practices for all creative sectors. An improved copyright data infrastructure would also support the crediting of authors and performers and tackle the problem of the 'disappearing author' in the digital environment. Formats used for submitting and exchanging metadata, and Application Programming Interphases (API) used by digital services, are the responsibility of the distributors or users, that authors can seldom change.

Without uniform practices for entering Rights Management Information (RMI), the risk of omissions and errors arises. Correct entry of metadata, and ensuring that it is updated throughout the lifecycle of the work, i.e. 70 years from the death of the author, would be crucial. This requires training, coordination and transparency among the different players.

The aim of the Presidency was to identify and structure the issues that are of general relevance for a functioning copyright infrastructure, bearing in mind that these issues are part of a broader context, a context that includes, *inter alia*, the fast development of the data environment in general, the increasing use of Artificial Intelligence (AI), and blockchain technology. Given this vast field of issues and the obvious time limitations of a 6-month Council Presidency, the Finnish Presidency was conscious that the work on this topic could not be finished during its term. The main goal of the Finnish Presidency was therefore, above all, to start the exercise by drawing attention to, and taking stock of, all the work relating to the development at the metadata level of the copyright system that has already been done at stakeholder level and at Member State/EU level, so as to kick off a discussion, through a series of introductions to the topic, including from different perspectives. On this basis, it would be possible to consider what lessons could be learned from past initiatives. In this way, the Presidency hoped to contribute to finding a practical way forward to benefit European content markets in a broad manner, covering all different creative sectors, including publishing, music, images and audiovisual.

The Presidency believed that the approval of the Digital Single Market (DSM) Directive[1] in April 2019 provided a suitable moment to focus on these technical issues in greater detail, so that the information gathered could contribute to the national implementation work on the Directive.

Following, in particular, presentations at the High-Level Data Economy Conference in Helsinki on 25-26 November 2019, the term 'copyright infrastructure' has been broadly accepted as a suitable term to circumscribe the relevant issues at stake.

Sections II to IV of this note, along with the two Annexes, summarise the initiatives taken during the Finnish Presidency to generate discussion at EU-level on this topic. This involved gathering views from Member States and from the creative sectors (the publishing, music, images and the audiovisual sectors), as well as the inputs provided at the Data Economy Conference.

---

[1] Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC, OJ L 130, 17.5.2019, p. 92.

II.   BASIC ASPECTS IN DEVELOPING THE COPYRIGHT INFRASTRUCTURE

1.   Use of data

One of the key elements of the copyright infrastructure is the use of data by the parties of a specific creative industry. Data is used in different ways across the creative sectors. Until quite recently, reports on the use of works were submitted manually, for instance through excel files. Now many players have taken up more interoperable systems.

Article 7(2) of Directive 2001/29/EC on the Information Society, adopted almost 20 years ago, defines RMI as any information provided by rightholders, which identifies the work or other subject-matter, the author or any other rightholder, or information about the terms and conditions of use of the work or other subject-matter (protected by Directive 2001/29/EC or Directive 96/9/EC), and any numbers or codes that represent such information.

Recital 55[2] of Directive 2001/29/EC envisaged that technological development would facilitate the distribution of works and render the management of rights attached to works easier. This would entail, however, the need for rightholders to identify better the work or other subject-matter, the author or any other rightholder to provide information about the terms and conditions of use of the work or other subject matter. Moreover, the recital specifies that the data should be in place 'when putting works or other subject-matter on networks'.

In legal terms, metadata and identifiers are considered rights management information. Usage data, provided by platforms or other users, are data required under the Directive on collective rights management organisations (Directive 2014/26/EC)[3]. A challenge to address therefore is, how to encourage the creative sectors to provide correct data and encourage the use of existing identifiers, and how users can be offered easier means for reporting usage data.

---

[2]   Recital 55 of Directive 2001/29/EC states: 'Technological development will facilitate the distribution of works, notably on networks, and this will entail the need for rightholders to identify better the work or other subject-matter, the author or any other rightholder, and to provide information about the terms and conditions of use of the work or other subject-matter in order to render easier the management of rights attached to them. Rightholders should be encouraged to use markings indicating, in addition to the information referred to above, inter alia their authorisation when putting works or other subject-matter on networks.'

[3]   Links to legislation and other documents relevant to developing the copyright infrastructure are set out in Annex I to this document.

2. Format of data

Missing standards for data exchange (format) cause submitted metadata to disappear, possibly unintentionally. But for example in the images sector, it was reported that intentional removal of metadata continues, as 96% of newspapers and in particular on the social media platforms lack metadata on images. This intentional removal or alteration of metadata on different services or platforms is commonly called 'data-stripping'.

The format of data is therefore the most relevant element for the interoperability of data that identifies the works and its rightholder. There are some best practices and standards for the interoperability of this data within the creative sectors.

The membership of CMOs is key to exercise rights in an efficient manner, in particular for music. The creative sectors, CMOs and other rightholders, such as publishers, have standard practices and use multiple different identifiers in order to identify the works as part of a repertoire, or for the purposes of managing a catalogue. Typically, they are based on the Digital Object Identifier (DOI) system, standardised by the ISO[4]. The images sector does not have an ISO identifier, but the International Press Telecommunications Council (IPTC) has just published a Photo Metadata Standard that enables identifying the copyright owners of multiple photos integrated into a single composite image.

Rightholders are also complementing the identification process with watermarks and content recognition technologies. Identifiers of authors and rightholders as clients of a CMO – such as the Interested Party Information (IPI) the unique identifier number (IPN) etc. – are often not ISO-standards, but serve in the internal exchange of data between CMOs or the industry.

These standards are developed for setting a standard language and vocabulary to express rights in machine readable format. In order to fully understand the impact and potential of these practices, they need to be reflected as part of a bigger infrastructure.

---

[4]   Such identifiers are ISBN (books), ISSN (serials), ISWC (works), ISRC (recordings), ISTC (text) and ISAN (audiovisual productions). An identifier that complements the existing ones is the ISNI (public identity).

Still, the lack of interoperability of the formats used in the data exchange of identifiers, and other metadata, causes problems and often leads to duplication of effort. One important issue is that while CMOs hold very comprehensive data repositories, they are not complete, as not all creators of works belong to a CMO and therefore the CMO does not have all the relevant information.

An example of new data exchange formats is the EU-funded data exchange standards used within the ARDITO[5] project that formed a Digital Rights Network. The background to the project, dating back to 2013, was that while copyright was applicable online, it was still difficult to implement it due to the scale, speed and complexity of the issues related to publishing, access and re-use of works on the internet. The project developed a technical framework required to facilitate automated communications between rightholders and those who wish to use content in text, video, images and multimedia. The purpose of the project was to ensure interoperability between identifiers and, where interoperability did not exist, to complement them with a globally resolvable new identifier.

There are also numerous industry standards for data exchange developed by the standardisation organisation Digital Data Exchange (DDEX). Anyone 'with a business interest in digital media content' can apply membership in the DDEX. The DDEX has grown to cover a broad range of record labels, and artists, as well as Music Licensing Companies (MLCs), such as digital service providers.

Finding common standards for data exchange in all creative sectors therefore remains the main challenge for the use of formats.

3.   **Quality and authoritative source of data**

This section focuses on how existing standards or the lack of standards impact the quality and hence the authoritative source for data, and ultimately the operation of the copyright system. Technologies may enhance the quality of data, but to manage copyright efficiently and securely for both rightholders and users, users should also be able to verify rights-management information and metadata against an authoritative source (at a given time).

As previous attempts to build centralised databases on repertoires for specific sectors, for instance the Global Repertoire Database (GRD), have failed, it essentially comes down to solving the challenge of how comprehensive, up-to-date and verified metadata could be ensured.

---

[5]   https://www.ardito-project.eu/

III. **WORK CARRIED OUT DURING THE FINNISH PRESIDENCY**

1. **Consultation of Member States and creative sectors**

In a first step, the Presidency requested the government level as well as industry level representatives through a questionnaire[6] to provide information on initiatives related to work and author data (metadata), best practices or standards on metadata, the use of blockchain-based services or AI enabled techniques for identification of works/rightholders etc., as well as on the state of play of, and possible views on, alternative approaches to develop the copyright infrastructure. Ten Member States' replies to the questionnaire were received[7].

Furthermore, in the context of brainstorming discussions held in the Council Working Party on Intellectual Property (Copyright), the Presidency organised a side-event with representatives of the creative sectors on 14 October 2019, on the margins of the Working Party, to present to the Member States the state of play and the stakeholders' views with regard to developing the copyright infrastructure in practice[8].

Member States and the Commission considered that it was the right time to start discussing issues relevant for developing the copyright infrastructure at EU level. In the discussions, Member States also underlined the technical nature of the issue and that they were still in initial stages of their analysis and views.

To get as complete a picture as possible, the Presidency also aimed to gather information on relevant legislation and/or court cases etc. involving new technologies, rights management information of works and other subject matter.

---

[6]   WK 8951/2019.

[7]   WK 9824/2019 +ADD 1 +ADD 2.

[8]   WK 11086/2019 +ADD1 contain the programme of the event and the stakeholders' presentations.

2. Main issues raised and concerns expressed relating to use, format and quality/authoritative source for data

a) *Considerations and views with regard to use of data*

– *Views from Member States*

In their replies to the questionnaire, Member States confirmed that the main issue of concern was the incompleteness or inaccuracy of work and author metadata. They emphasised it was necessary to develop an easy way of retrieving data on works and authors. Even though it was also emphasised by them that these issues fall essentially within the realm of the private business interests of rightholders, many took the view that identifiers used by industry are either not always adapted to the digital environment or not always mutually interoperable. It was also stressed that metadata issues are of a global nature and, therefore, issues around metadata should preferably be dealt with at the international level. This being said, Member States considered that the EU level is an appropriate place to start working on the matter, considering that most global identifiers used today are developed by the creative sectors in Europe.

The responses from the Member States drew attention, in particular, to the following issues:

Numerous disadvantages have been found in identifiers, though many identifiers have been updated during recent years. For instance, no complete list of identifiers is currently available, as sources only list identifiers that are relevant for their own sector. One less known identifier, the International Standard Link Identifier (ISLI), was mentioned in the reply by the Czech delegation. Developers of digital services would benefit from transparency with regard to these different identifiers and the standards used for data exchange.

Studies[9] in the UK and Finland have shown that even correctly submitted metadata can get lost, or stripped, due to the use of different formats during the 'path' from creation to publishing of a work.

Data delivered by users on the use of works ('reporting') is often not uniform in format, with large deviations in quality and quantity. Digital tools could assist users to report usage in a more efficient and uniform way.

---

[9] Please see, in Annex I, 'Other documents'.

Attribution is a formal part of moral rights in most copyright laws under the Berne Convention, although it is not harmonised in the EU *acquis*. Reliable and up-to-date metadata can also identify works that are either free to use after expiry of the term of protection or are accessible to the public under different open access systems (such as 'creative commons' licensing).

Replies from Member States suggest that metadata quality could be improved by forming unified practices at EU level. In summary, inadequate practices for submitting metadata in works or the lack of uniform use of identifiers or metadata by users was identified by Member States as a subject for further work, or even regulation. As highlighted by the Croatian delegation, according to Croatian national stakeholders in the creative sectors, possible regulation of issues in relation to identifiers and metadata would need to be followed by deployment of new technologies that directly support the use of identifiers. This primarily refers to using and processing standard formats for data delivery (some of which have already been developed within industries themselves). In this context, it should be noted that certain other technologies, such as fingerprinting, are used ever more frequently.

The replies contained suggestions for minimum information (of musical works) that should be provided. This information should comprise the name, (possibly) the performer, the producer (and/or publisher), the author(s) and – a globally resolvable identifier such as – the International Standard Recording Code (ISRC). Here, an indication of the role of the author and possible performer in the work or version would be beneficial and in the interest of clarity. Some identifiers already describe the author's participation in the whole work, while others in a specific layer, e.g. direction, scenario, music composition, camera footage etc.

– *Views from rightholders and industry*

Representatives of the creative sectors that took part in the side-event organised by the Presidency on 14 October 2019 expressed a preference for the use of voluntary standards on interoperability of existing identifiers. However, the development of identifiers was considered an internal activity of and between CMOs. The representative of the audiovisual sector described that the purpose of the International Performers Database (IPD) is to declare and manage the mandates CMOs have been given by the performers that have been assigned a unique identifier number (IPN). This reliable database was created in 1997 to identify individual performers in audio recordings and audio-visual works and to provide evidence in relation to other CMOs of the repertoire that the IPD manages.

The purpose of the IPD is thus not to deliver information on works and authors to third parties or to consumers.

It was the general view of the stakeholders that voluntary and open standards developed in co-operation with government could be beneficial to European content markets. With investments in mutually beneficial initiatives, Europe could lead the way towards a rights-data infrastructure for content. It was emphasised that licensing activities, i.e. rights management, should be kept separate from the data management, because this is essential to preserve open competition in the rights supply chain. Some suggested regulating the use of identifiers, following positive experiences in Switzerland.

Another example, coming from France that relates to the use of data is the code of best professional practice, that was adopted by several press unions or federations, press agencies and photographers, following mediation by the French Ministry of Culture. One of the main purposes of this code was to ensure trustworthy information about the photographs. Photographers and agencies have committed to introduce and transmit the metadata on photographs according to a defined common standard. The work entailed a shortlist of minimum data entries that include: the name; the 'quality' (i.e. role) of the author of the photograph; the kind of picture (e.g. institutional, professional); the place and date of shooting; and the identification of the people or elements depicted and any restriction on use. In return, newspaper publishers committed to make sure that the metadata was kept and not stripped or lost due to format shifting.

The *Association Littéraire et Artistique Internationale* (ALAI), an organisation for academic research in copyright, participated in the side-event to share experiences of an ALAI survey launched in 2016 covering a range of online practices (concerning photography, visual arts, music, audiobooks, radio, film, TV, video, e-zines, e-books, blogs, news, and news aggregators). The survey was to conduct research in order to find out how the presumption of authorship provided for in Article 6bis(1) of the Berne Convention is fulfilled in the digital environment. There was an initial assessment already in 2016 that data-stripping is taking place, as authors are seldom in charge of the formats in which their works are published, which makes the position of the author very weak. Following a relaunch of the survey, ALAI will publish in 2020 a summary of its survey on the online practices for crediting authors.

The Presidency has identified a couple of key takeaways from a study commissioned by the UK Intellectual Property Office ('Music 2025: the music metadata dilemma: issues facing the music industry') that also addresses many of the elements described above, and remains relevant to the agenda on developing copyright infrastructure. The study recognised, among other things, that it is urgent to address, in the music industry field, certain specific issues related to metadata and identifiers. The study came to the conclusion that, if the importance of data is not acknowledged, the emergence of a 'Do-it-yourself' (DIY) culture – also known as 'prosumers' – and diverse ways of creating, will continue to increase and compete with professional content. A study commissioned by DuPho, the Dutch Photographers association, came to a similar conclusion.

The images industry was one of the first industries that felt the impact of having its content used by sharing platforms, as regards both images taken by professional photographers without any metadata, and those shot by non-professionals and used in blogs and newspapers. The music publishing sector is already contributing to developing the copyright infrastructure, by standardising 'cue sheets' in digital form that are completed with the composer, performer, label and other relevant information by those who are involved in music production. This way the accuracy and interoperability of data submitted with a registration, made at a performing rights organisation, can be ensured.

Another key recommendation of the above-mentioned Music 2025 study is the conclusion that 'no upload without identity' should take place. In this context, it is worth noting that, pursuant to Recital 55 of the Information Society Directive, rights management information should be in place when the work is published. It follows from this recommendation that ensuring that works are not published without 'identity' continues to be the key issue, which should therefore be embraced by the entire European content sector as soon as possible. It was also suggested in this UK study that links between identifiers should be enabled, by ensuring that all relevant databases are interconnected. It was also stressed that the music industry cannot do this work on its own, due to vested interests and power imbalance. Finally, to reduce administration costs and increase productivity, multi-layered collaboration and communication on a regular basis is recommended.

The Commission considered that the overall impact regarding any single initiative will remain small unless there is more coordinated action among different players within the industry.

b) *Considerations and views with regard to format of data*

– *Views from Member States*

The responses gathered from Member States suggest that it is not necessarily new metadata or other standards that are needed, but rather a facilitation of the data exchange architecture for copyright information for all sectors. The creative sectors are using data differently and have different levels of take-up of data standards. For instance, the German delegation's reply to the Presidency questionnaire noted that the scientific publishing sector seems to be espousing a strong industry-wide adoption of metadata standards related to individual articles and authors, whereas the music industry has employed several descriptive tools and numerous identifiers. In contrast to the publishing and music sectors, the audiovisual sector has developed without a standard identifier, until the recently adopted identifiers that are not uniformly applied throughout the audiovisual producers, who are mostly small and medium-sized companies.

The experiences from the academic and research fields in France suggest that the encouragement of widespread use of the ORCID identifier (Open Researcher and Contributor ID) for scientific authors would improve mechanisms – like automated linkages – for enforcement of proper attribution of their publications. Such mechanisms also contribute to moral right protection and the fight against plagiarism in scientific publications. As particularly emphasised by the French delegation, that complying with moral rights is equally important in the digital environment. Other similar projects can be found for instance in the E-Cultura Digital Library in Romania that is cataloguing standards, standards for technical metadata and standards for digital images interoperability.

According to the reply by the German delegation, the publishing rights organisation VG WORT has been calling for the introduction of obligatory use of the International Standard Name Identifier (ISNI) for all authors who have a contract of representation with VG WORT. In the long term, the internal VG WORT number will be substituted by ISNI. Furthermore, a joint venture by German, Finnish, Swedish and UK CMOs, the International Copyright Enterprise (ICE) is working on joint documentation of music rights, licensing of digital music providers and processing of digital usage reports.

It can be considered that improved copyright infrastructure could help overcoming the shortcomings of a fragmented and disorganized infrastructure and thereby contribute significantly to economic growth of the European digital market, as even the value of independent or 'prosumer' content could be recognised.

–   *Views from rightholders and industry*

As was strongly emphasised by the representatives of the creative sectors at the Presidency side-event, there is a need to start working for standardisation throughout the creative sectors in Europe. A good start would be to form a unique persistent identifier and a way to resolve the identifiers globally and to begin setting data exchange standards for all types of content. Trusted certifications for authority services could be provided in co-operation with all stakeholders.

An example on work with a 'unique persistent identifier' is the ARDITO project mentioned above. The development of improved user access to rights data via the Copyright Hub also took into consideration user-generated content. ARDITO helps users be directed to the entity offering licences for the use of images and other content.

Furthermore, the representatives of the audiovisual sector promoted an effective use of standard-IDs, as usage reports are key to revenue paid to the authors/performers. As the systems need to be able to process huge volumes of data from platforms and prepare also for micro-licensing at some point, effective use of data would cut down the cost of basic rights processing and would allow CMOs to focus on conflict resolution etc. in the future. In addition to broad dissemination of the ISAN, data exchange standards similar to DDEX specifications are needed. All such suggested improvements and standards should be made under ISO compliant regulation and if possible under EU regulation that guarantee open system and free access to data, long term availability of the system, mandatory use of standards in key areas to ensure 'exhaustivity' and a level playing field.

A study in the image sector (by Imatag) shows that metadata is missing from 85% of images. Copyright related data is available only in 3% of images. The representative of the image sector reported that Google Images and the Centre of the Picture Industry (CEPIC) had agreed to show credits for images by using the IPTC Photo Metadata in the Google Images Search. While this is an improvement, most newspapers, and all web pages and social media sites, are still not complying with metadata standards. Imatag is one company offering services to the images field through their content recognition and watermarking technology. To estimate the extent of the problem, the number of images they have been able to identify with the technology is normally only a fraction

(perhaps 10%) of the amount of images by a specific client that have been uploaded to the internet without authorisation and without metadata.

The RDx, the data exchange platform for performing rights for the recording industry was published in late 2019. It is an international producer's database and the first in-house application using the data exchange system developed by DDEX. While working also on their own data management practices, the International Confederation of Music Publishers (ICMP-CIEM) (including also CISAC, the International Confederation of Societies of Authors and Composers) will be adhering to this message exchange suite, as declared during the side-event.

c)   *Considerations and views with regard to quality and authoritative source of data*

– *Views from Member States*

Rightholders have expressed that the authoritative source for correct data on the holder of copyright is 'the rightholder'. While this is true, none of the existing databases are exhaustive.

Discussion only started on what steps would be required to achieve to guarantee the quality and authoritative source for metadata, but a general conclusion that can be drawn from the feedback received is that the EU content markets would benefit from minimum or common standards for rightholders to declare their rights to content in an authorised, machine-readable, accessible form. New technologies and services can support better data management by rightholders, which will subsequently support the functioning of licensing markets and the remuneration of rightholders.

Germany's experience suggests that a holistic copyright metadata infrastructure should not (attempt to) draw on creating an individual database as a single source of authoritative information on metadata. Attempts to build 'copyright hubs' have shown that a centralised, globalised service appears too expensive to build, too unwieldy to manage, and offering too little return on the initial investment to ensure that the responsible body or entity has a sufficient incentive to manage it.

Replies and materials provided by Member States suggest that the reason for the failure of a single source database (GRD) could have been due, among other things, to the fact that no agreement could be reached on the standardisation of the data to be submitted to the database and/or that there was disagreement on the financing of the maintenance of the registry and lack of willingness to co-operate.

Instead, verifying the accuracy of metadata should rather be a process using several available sources. Such approach would also promote the access to databases with accurate data, and give an opportunity to estimate the value of data from one source. A possible solution might be a kind of 'metadata search and enhancement engine', operated by a trusted non-profit organisation and based on an open and fully transparent AI algorithm.

Among public actors, national libraries and other public legal deposit institutions already play a very important role in the referencing of works (books and works on other media), with a very high level of metadata quality. With their multiple expertise (in the field of quality control of metadata, as well as of 'semantic web' standards like RDF – resource description framework, for instance) and with their own sets of metadata to share, they could in the future be partners, in one way or another, in the global copyright infrastructure.

Another idea put forward was the establishment of an ad-hoc expert group established with a mandate by the Council to facilitate the exchange of best practices and experiences among Member States in order to develop the copyright infrastructure. This could pave the way for future work in the field, including for the implementation of Article 17 of the DSM Directive.

It was also noted that were more detailed studies to be commissioned by individual Member States, the European Parliament and the Commission, they could look into the various ways that might facilitate the development of a common copyright infrastructure.

The Commission announced that it is planning to launch a public procurement procedure to commission a study following up on the work of the Finnish Presidency.

– *Views from rightholders and industry*

CMOs cannot verify the data in all cases, because not all rightholders are registered as clients of a CMO. This means that the CMO's data is incomplete. As described by a songwriter at the Data Economy conference, it is possible that available data on a same work is different in different CMOs. The Commission's Recommendation in 2005 to allow authors to be members of a CMO of their choice has resulted in different registrations of one and the same work in different CMOs. CMOs then have to match these different registrations when they analyse the usage data from the user reports. Therefore, payments of collected remuneration at times have been very slow, or have been left non-distributed due to lack of registration.

The representative of the audiovisual sector considered a hub of authoritative source databases to be appropriate for a copyright system, which required, as a first step, and possibly through legislation, to make the interoperability of pivot standard identifiers a condition.

Representatives of the publishing sector also recalled that governments should not overestimate the difficulty of building such an infrastructure; in the view of this sector, it is achievable. It was also mentioned, reflecting what was emphasised by Recital 55 of the Information Society Directive, that 'reliability' comes from a distributed data network of rights data, which implies a distributed *responsibility* of rightholders (i.e. creators and other rightholders and their representatives) to declare their rights and consequently justify the *liability* for users of works, in practice, to seek permission to use works. This implies that the publishing sector has considered a user-friendly approach to the application of copyright law.

There also exist some data management practices for libraries responsible for cultural heritage, audiovisual, films and computer programmes. In France, several feasibility studies have been published in an open register for 'cultural metadata', following up on the mission undertaken by Pierre Lescure[10]. A synthesis study is expected to be published in due course.

At international level, WIPO is currently compiling, and will publish in early 2020, a summary of the views of WIPO Member States on voluntary registration of copyright.

The digital metadata ecosystem allowing works to be identified includes multiple actors, both private and public. Existing standards and practices require new digital tools and technologies, as do ongoing projects to increase trust in data available to digital service providers or online content-sharing service providers regulated under Directive (EU) 2019/790 on the DSM.

d) Other considerations

In order for users to trust metadata included in digital files of works, it is important to ensure that data is not altered or removed. The Presidency asked the Member States about experiences and possible court cases relating to the practical application of the national legislation transposing Article 7 of Directive 2001/29/EC. In the replies received, only the German delegation reported about a court case that related to the application of the provision on the protection against removal or alteration of rights management information. In this case, Facebook had been ordered to refrain from removing metadata from uploaded graphic files pursuant to §95c German Copyright Act

---

[10]   Please see, in Annex I, 'Other documents'.

(ruling by the Hamburg Regional Court, 15 January 2016 – 308 I 48/15). The lack of further examples from other Member States seems to indicate that not many rightholders have used this provision/piece of legislation to enforce their rights against unauthorised removal of metadata. The Italian delegation pointed out that even if stronger punishments would decrease the alteration and removal of metadata, data-stripping would still take place across sectors (without other supporting measures).

The replies by Member States listed studies addressing the opportunities and challenges related to AI and blockchain in different fields. The challenges brought about by AI and blockchain technologies are numerous. Some Member States mentioned having started developing their first national strategy in the field of AI. Though none of these strategies seem to be specifically linked to the development of the copyright infrastructure, at this stage, they relate to issues that could be cross-fertilisers. They include, among other things, the interoperability of standards, as well as the collection and validation of metadata. From a broader perspective, the ethical use of AI technologies was also mentioned.

The representative of the photographers organisation DuPho described the PicaPro – the digital blockchain app for images. It has all five important features needed in the digital economy. It enables smart contracts, establishes transparent P2P transactions, promotes efficient licensing systems and a reputation system, and micromanages and micro-monetizes content. This service has been tested with photographers and the initial feedback is positive. One reason has been that it uses real currency not crypto currencies. Also, the French content industry is working on blockchain based technologies.

It was suggested by the German and Portuguese delegations that it would also be important to coordinate the issues at hand in a broader context, taking into account regulatory activities in the IT and telecommunication sector and data protection (MyData etc.).

3.   **Issues related to access to and sharing of data**

The data economy is one of the strategic focus areas during the Finnish Presidency of the Council. Though sharing of data and access to data involves IP related issues and rights to restrict access to data, these aspects were not discussed in-depth at Council Working Party level. The use, management, and protection of data was the main theme at the High-Level Conference on Data Economy in Helsinki, 25-26 November 2019. The main outcomes from this conference were

reflected in the common *Principles for a human-centric, thriving and balanced data economy*[11], based on the abilities of companies, the public sector as well as individuals to access, share, act, innovate trust and learn from accessible and interoperable data[12].

At the side-event, the music rights CMO Teosto considered that managing data in silos will slow down innovation in the field of music management, make it costlier, and ultimately stifle it. Instead of radical opening of data for free access to anyone interested, data should be put in the centre of business development of CMOs. Opening data should be a means, not a goal. Where open data is not appropriate, data sharing solutions should be adopted. By enriching their data and providing a better positioning in the data value chain CMOs may succeed better in the data economy. All the investments should ultimately benefit the rightholders.

In the same vein as the 'inspirational speech' delivered by David Osimo on 17 July 2019 on the margins of a Council Working Party meeting[13], Teosto proposes the following key drivers for the copyright infrastructure: Requirements for interoperability and real-time processing of data. Distributed ledger systems (blockchain) can also speed up payments to rightholders, which can be a differentiating factor for new competitors entering the rights management business.

IV. <u>STOCKTAKING OF THE PRESIDENCY</u>

On the basis of the input collected and the discussions held during the second half of 2019, the Finnish Presidency is convinced that the challenges of, and potential solutions for, the development of the copyright infrastructure will continue to be highly topical in the years to come. More work needs to be done to support the relevant industries in respect of submitting metadata and enhancing the quality of metadata, as well as the format for the exchange of such data, and identifiers to ensure interoperability across the relevant fields.

Next steps for developing the copyright infrastructure could include:

- the preparation of guidelines setting out minimum requirements and/or best practices at EU level for submitting rights metadata in works when they are published, so that it corresponds to the data submitted to the CMOs (where applicable);

---

[11]   https://dataprinciples2019.fi/wp-content/uploads/2019/09/Dataprinciples_web_1.0.pdf
[12]   A summary of the presentations and discussions at the conference relating to copyright is set out in Annex II to this document.
[13]   WK 12386/2019.

- measures to allow verification of such data from an authoritative source;

- improvement of metadata quality by encouraging open standards and interoperable identifiers with a view to ensuring uniform application in the use of digital tools, such as Application Programming Interfaces (APIs);

- exploring the metadata capacity of legal deposit libraries;

- further consideration as to whether the building of a hub for European content, with related data on works and rightholders, should continue to be a goal of the process.

Stakeholders from the publishing, music, images and audiovisual sectors emphasise that the effective use of open metadata standards, interoperable identifiers and co-operation is important. They suggest it would also be of key importance to ensure process automation capacity through data exchange standards, and to develop real-time data processing tools and decentralised systems.

The public sector could play an important role in the further examination of the issues identified above, *inter alia* through studies or consultation with stakeholders, by fostering stakeholder dialogues. Investing in political, technological and economic analyses has been suggested, as well as the identification of what sources for funding could be provided for conducting comprehensive studies on metadata and identifiers, including the opportunities provided by new distributed ledger technologies.

The development of the copyright infrastructure also links to broader phenomena, like the European Data Economy, open data policies, ePrivacy and the use and impact of AI for the benefit of the better functioning of the IP system. The roles and tasks of CMOs, individual rightholders and users (both commercial and private) in the licensing of works and in the copyright infrastructure should be kept separate. It is hence important to separate rights management from data management. Finally, involving all stakeholders, including users and 'prosumers' (i.e. individuals who are both consuming and producing copyright protected content) should be promoted.

The Finnish Presidency hopes that its endeavours in relation to the copyright infrastructure, and this stocktaking paper, will provide useful input to incoming Council Presidencies and to the 2019-24 European Commission.

Documents relevant to developing the copyright infrastructure

1.   EU and international legal framework:

–   Articles 5(2), 15(1) and 6*bis*(2) of the Berne Convention

–   Article 12 of the WIPO Copyright Treaty (1996)

–   Article 19 of the WIPO Performers and Phonograms Treaty (1996)

–   Relevant articles of the Directive on the legal protection of databases (Directive 96/9/EC)

–   Article 7 of the Directive on the harmonisation of certain aspects of copyright and related rights in the information society (Directive 2001/29/EC)

–   Article 5(2) of the Directive on the enforcement of intellectual property rights (Directive 2004/48/EC)

–   Commission Recommendation on collective cross-border management of copyright and related rights for legitimate online music services (2005/737/EC)

–   Relevant article and annex of the Directive on certain permitted uses of orphan works (Directive 2012/28/EU)

–   Several articles of the Directive on collective management of copyright and related rights and multi-territorial licensing of rights in musical works for online use in the internal market (Directive 2014/26/EU)

–   Relevant articles of the Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (Regulation (EU) 2016/679)

–   Relevant articles of the Directive on copyright and related rights in the digital single market and amending Directives 96/9/EC and 2001/29/EC (Directive (EU) 2019/790)

2. *Other Documents:*

– ‘Study on blockchain potential in the music industry’ (in German):
https://www.pwc.de/de/technologie-medien-und-telekommunikation/pwc-studie-nach-dem-streaming-kommt-die-blockchain-hype-oder-echte-chance-fuer-die-musikindus-trie.pdf

– ‘Derivative use of public domain content – film industry focus’:
https://euipo.europa.eu/tunnel-web/secure/webdav/guest/document_library/observa-tory/documents/publications/public_domain/Full_Final_Report_Public_Domain.pdf

– ‘Why success of the music modernisation act depends on open standards’:
https://static1.squarespace.com/static/56d5e44060b5e9e20a94b16c/t/5bd7199e8165f5a5e9241ae9/1540823454789/Open+Music-Music+Modernization-Open+Stand-ards+White+Paper.pdf

– ‘Why building more rights databases won’t solve the music industry metadata problem’:
https://www.hypebot.com/hypebot/2018/01/why-building-more-rights-databases-wont-solve-the-music-industry-metadata-problem.html

– ‘Music 2025: the music metadata dilemma: issues facing the music industry’:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/809492/Music-2025.pdf

– ‘Metadata of digital music files: summary’:
https://musiikintekijat.fi/sites/default/files/metadata_170925_eng_final.pdf

– Principles for a human-centric, thriving and balanced data economy’:
https://dataprinciples2019.fi/wp-content/uploads/2019/09/Dataprinciples_web_1.0.pdf

– ‘A symphony, not a solo – how collective management organisations can embrace innovation and drive data: sharing in the music industry’:
https://www.teosto.fi/sites/default/files/files/A%20Symphony%20Not%20A%20Solo%20Policy%20Brief%20Final%2009012019.pdf

– In 2018, the French High Council for literary and artistic property (CSPLA), an advisory board whose main role is to advise the Ministry of Culture and in the field of copyright and related rights, published a report on blockchain.

'Mission du Conseil supérieur de la propriété littéraire et artistique (*CSPLA) sur les chaînes de blocs* (blockchains)':

http://www.culture.gouv.fr/Thematiques/Propriete-litteraire-et-artistique/Conseil-superieur-de-la-propriete-litteraire-et-artistique/Travaux/Missions/Mission-du-CSPLA-sur-les-chaines-de-blocs-blockchains

–   The CSPLA published in 2018 a report relating to the implications of new legal concepts such as 'data' or 'digital content' for literary and artistic property.
'Mission du CSPLA sur les conséquences pour la propriété littéraire et artistique de l'avènement des notions de données et de contenus numériques':

http://www.culture.gouv.fr/Thematiques/Propriete-litteraire-et-artistique/Conseil-superieur-de-la-propriete-litteraire-et-artistique/Travaux/Missions/Mission-du-CSPLA-sur-les-consequences-pour-la-propriete-litteraire-et-artistique-de-l-avenement-des-notions-de-donnees-et-de-contenus-numeriques

Summary of copyright-related discussions at the High Level Conference on Data Economy

On 25-26 November 2019, a High-Level Conference on Data Economy brought together nearly 400 participants at Finlandia Hall in Helsinki. Jointly organised by the Ministry of Transport and Communications, the Ministry of Finance and the Ministry of Education and Culture, the Conference was the backdrop to the introduction of a set of data principles for the growth of the European data economy. The data principles drafted for the Conference aiming at defining and promoting a human-centric, successful and balanced European data economy were built on the abilities of everyone to *access, share, act, trust, innovate and learn* from interoperable and trustworthy data.

The Finnish Presidency wanted to raise awareness of the importance of data, including sharing and access to data for the benefit of European content markets. In particular, representatives of the games and music sectors and experts on metadata highlighted, together with the Commission, the importance of data for the growth and competitiveness of the digital single market.

The creative sectors were also invited to discuss the topic '*New revenue streams from the platform and data economy*', which focuses on developing the copyright infrastructure, that is the metadata level related to copyrighted works and authors in order to be able to develop automated licensing systems where applicable.

Speakers representing the games industry (Markku Ignatius), songwriters (Helienne Lindvall) and metadata experts from academia in the US (Bill Rosenblatt) at the Conference confirmed many of the issues described above.

Among them are the fact that huge revenue from the games sector has to do with direct investment in the developers, empowering them, instead of licensing their products to third parties. Also the design of products and services make consumers want to commit to one game, and make in-game purchases. The game changer was the mobile APP (Application) stores. The music is still bought separately for the games, and that is one big part of the competitiveness of the gaming industry, but this is also changing. Games are interactive, as opposed to music or other types of content. Ways to attract talent to Europe is their suggestion to ensure growth in the future. The social aspect of the games, and game platforms, are both developing. They are essentially malls, where the younger generation is doing everything it is doing in 'real' life, including going to concerts.

Missing or unauthoritative metadata causes loss of revenues for rightholders. This is the major challenge for the industry. The growth of the games industry does not show in the music industry's results. There are many challenges that make up the 'music industry data problem', because it is multifaceted and involves many entities. It is estimated that today more than 20% of publishing revenue is lost due to incomplete or unmatched metadata. In lack of attribution of the data, the revenue is distributed according to market share. DSPs should have access to a source from which song data could be verified, and labels should not be able to publish songs without data. The accurate data should come from the songwriter. It is not a territorial problem – for example EU wide problem – but a global challenge in a global market that the EU may help to solve. Some sort of regulation is needed for the users to know who are the rightholders. Education is also needed.

Furthermore, the topic of automated licensing of digital content was specifically discussed and with the view of clarifying what the EU should do to enable such licensing. One aspect was that licensing should be possible at internet speed because this is the expectation of the users of works in the digital environment. Otherwise they use more easily accessible content, that can also be from pirated sources.

Complete, precise, accurate, current data is required for content and rightholder identification (metadata) as well as protocols to exchange such data, and access to data repositories and registries. In this context the term 'copyright infrastructure' was broadly accepted to describe relevant issues been discussed on international level since the end of 1990s.

It was also highlighted that the metadata or the elements of the infrastructure was different from the perspective of the songwriter, performer and the publisher. One could even add the cover art designer too that is mostly not mentioned in streaming services. Each of them has a set of minimum data entries and corresponding identifiers. It was also noted that the number of collaborators in one song have also increased remarkably. One example is 'Bad Romance' having the work ID (ISWC T-903308670-3) and songwriters: Stefani Germanotta (Lady Gaga), IPI 519338442 and Nadir Khayat (RedOne), IPI 268458032.

The names in brackets refer to the publishers of each of the songwriters. They can and often are different, but there is no obligation for them to work in co-operation to serve the user in a licensing process (i.e. when someone wants to have permission to use a work). So they have their own IPIs – House of Gaga Publishing, IPI 664688789 and Songs of RedOne, IPI 606809345 as well as

Sony/ATV, IPI 187062752 – depending on if the rights are split between several publishers or there has been a transfer of rights due to a merger etc.

The recording of the track 'Bad Romance' has the ID (ISRC USUM70903859) and information about the (main) artist: Lady Gaga, the producer: RedOne and the record label: Interscope/Universal Music Group. This metadata could also hold IPNs identifying the performer etc. What furthermore is developed is the DDEX ERN-standard for the data exchange of sound recording metadata from the label to Spotify and the DDEX - DSR for the sound recording play (usage) data from Spotify to the label. Helienne Lindvall pointed out that her own rights management has suffered from this complexity, but also from the fact that rights can be split by not only several publishers, but also – following the Commission Recommendation in 2005 – several CMOs. This needs somehow to be streamlined, in order to ensure links and supportive activity of all parties involved.

The lessons learned are that technologies follow a lot of entrepreneurial activity and start-ups. Everything cannot be solved at once. Standards have to develop together. In addition, appropriate governance and dispute resolution mechanisms are needed.

Bill Rosenblatt from GiantSteps invited to work for a 'minimum viable' approach based on a SAUCE test: Scope, Adoptability, Urgency, Complexity and Equity. The aim should be for a clearly focused solution to an existing problem that works also with existing tools but is relatively simple and that creates a win-win-win for everyone in the industry. He also emphasised the need for financing and governance of any model developed. Scale will still most likely force change. It is impossible to process 40 000 uploaded songs, and transaction volumes of these works, on a track by track basis without automation. Awareness will create a competition and thereby a demand for precision and transparency in payments.

From the Commission, Roberto Viola and Gail Kent (DG CNECT) were present. Talking about the state of play within the Commission and DG CNECT on developing the copyright infrastructure, Ms Kent said that the data economy and this topic is, and continues to be, at the core of EU policies being developed, now that the new Commission is starting its work.

In the panel entitled '*What should the EU do next to develop the content market*', speakers addressed the issue of developing the copyright infrastructure, noting in particular that the growth in the use of music, which is bigger than ever, is not reflected in the revenue to songwriters. Therefore, it is crucial to develop incentives, and to provide financing for the solutions. Also, Helienne Lindvall noted that there are big players who benefit from the complexity and missing metadata.

The function of the CMOs should be to serve the songwriters, and not to hold on to their data on the excuse of the CMO's mandate and repertoire. The panelists also discussed what is the right governance model in relation to data in the copyright field, consider that this is not an easy task. The International Copyright Enterprises (ICE) is one model to license access to data. The more entities involved, the higher the administrative costs. On the other hand, competition is also important, as a means to keep services efficient. To the question as to whether all data should be available for free, it was uniformly agreed that not all data needs to be shared for free, but there needs to be transparency. The fact that data is not being shared causes many missed opportunities. Sometimes, the mere availability of data can create new services, and solve previously unsolved issues through entrepreneurial activity.

According to Helienne Lindvall, creators rather than CMOs should be the ones to decide on access to their data. She stressed that empowering the rightholder instead of the platform was favored by the Commission. Bill Rosenblatt preferred some more detail on the transparency of the data held by CMOs, as opposed to just a general recommendation of transparency of data. Power imbalance and conflict of interest results from major labels and publishers having positions in the boards of CMOs, and in many instances they even decide on the representatives of the authors.

One important highlight that was emphasised by the panelists was that data needs to be submitted at the beginning, before publishing, otherwise copyright will not work. Data is valuable to creators and individuals in general, now and in the future, as already discussed in context of MyData. The main challenge is educating people. Helienne Lindvall said that creators do not need to know everything on copyright, but they must understand that they need to register their songs. 'Session', the service developed to help creators manage their music, works as a platform for songwriters that can agree on the share of the creators contribution in songs, and can also make sure that everybody has the same information on rightholders and IPIs and that, subsequently, registrations to the CMOs will be correct.

In the Commission's view, the best way to develop the copyright infrastructure was to gather existing initiatives and lessons learned, and decide on the basis of that what is the value-added of the EU on standards and interoperability.

The Ministers of the three organising Finnish Government Ministries each delivered a speech at the Data Economy conference. Minister Hanna Kosonen, Minister of Science and Culture, summarised the thinking of the Finnish Government with the following remarks: 'we believe this is the right

time to begin discussing the importance of data and issues related to data on works and authors on the EU level. The European Union needs a clear vision of how we can guide this development to the benefit of the creative sectors and science'[14].

----

[14] Press statement: https://minedu.fi/en/article/-/asset_publisher/luovat-alat-osana-datataloutta