The Proposition

This document proposes to formally establish the Copyright Infrastructure Task Force instigated by the Governments of Finland and Estonia and kickstart its operations.

Table of Contents

1) The Copyright Infrastructure Task Force (CI Task Force)

The CI Task Force has been established by the Governments of Estonia and Finland in September 2022.

The concept of Open Rights Data Framework (ORDF) emerged from –

- the stocktaking document "Developing the Copyright Infrastructure" published by the Council of the European Union under the Presidency of Finland in December 2019
- the study "Copyright and New Technologies" published by the European Commission in March 2022

The CI task Force has since then continued to explore the setup of a Multi-Country Project (MCP) around the Open Rights Data Framework (ORDF) as suggested in the above-mentioned study.

September 2022	Establishment of the CI Task Force by Estonia and Finland
January 2023	Report to the Secretaries of State
March 2023	Submission of interest in EDIC to the Commission
March-May 2023	Latvia joins Estonia and Finland; the three Member States prepare a pre-notification (technical description and proposed statutes) with the support of the Commission; Germany and Portugal observe the deliberations of the CI Task Force.
June 2023	Announcement of the pre-notification by the Commission
September 2023	Lithuania joins Estonia, Finland and Latvia; the four Member States prepare an application (technical description, budget, and proposed statutes) with the support of the Commission; Germany observes the deliberations of the CI Task Force. The Commission recommends the CI Task Force to work closely with Europeum EDIC.
January 2024	The CI Task Force and Europeum EDIC consider the opportunity for the CI Task Force to become a Use Case Group of Europeum EDIC and collaborate on an AI and Copyright use case.
April 2024	The CI Task Force presents the AI and Copyright Use Case at the Belgian Copyright Conference.
May 2024	Europeum EDIC is established.
June 2024	Italy joins the CI Task Force.

Figure 1: Timeline of the CI Task Force

2) Rationale and objectives of the Multi-Country Project

In *An intellectual property action plan to support the EU's recovery and resilience, COM(2020) 760*[1], the Commission announced that it will further work with relevant stakeholders to promote the quality of copyright data and achieve a well-functioning "**copyright infrastructure**" – the set of rules, technologies and institutions that **frame** data management practices in the creative industries[1] – to improve authoritative and updated information on rightsholders, terms and conditions, and licensing opportunities.

In *Developing the Copyright Infrastructure, 15016/19*[2], the stocktaking of work and progress under the Finnish Presidency, the authors noted that there are many ISO (International Organization for Standardization) and industry-specific identifiers for works and rightsholders, however they lack interoperability in a broader context. This results in high transaction costs for the industry and lost revenue for rightsholders. Investing in a **copyright infrastructure**, encompassing **standardised metadata** entries in connection with digital copies of works and where relevant, registrations with Collective Management Organisations (CMOs), can improve the efficiency of licensing, and allow automated processes for distribution of revenues.

In the study on *Copyright and New Technologies, SMART 2019/0038*[3], published by the Commission in March 2022, the authors argue that a well-functioning **copyright infrastructure** would require to **open** and integrate the **rights data framework** – a set of rules and technologies – to:

- Support releasing much more of the digital potential of Europe's creative sectors and contribute to the development of a single market for data,
- Address interoperability issues and make rights management simpler, more accurate, faster, and more affordable for all stakeholders on the content value network,
- Provide trustworthy rights information which can then be relied upon for rights licensing and rights enforcement, as well as for a fair, appropriate, proportionate, and transparent rights remuneration,
- Restore a level-playing field between major actors and the European small and medium sized creative enterprises through an inclusive approach catering for interests of any rightsholder, stakeholder, incumbent or new intermediary.

[1] Disambiguation: as the definition of the IP Action Plan explicitly mentions, the copyright infrastructure **frames**, it is a **framework**. Similarly, the electronic Identification, Authentication and Trust Services (eIDAS) is also a **framework** that ensures that electronic interactions are safer, faster, and more efficient, no matter the European country they take place in. The eIDAS regulation created one single **framework** for electronic identification (eID) and trust services, making it more straightforward to deliver services across the European Union. The eIDAS framework or the copyright infrastructure include neither data nor systems – these infrastructures are "orgware" on which various systems can be built. Differently, the European Blockchain Services Infrastructure (EBSI) is a network of distributed blockchain nodes across Europe – this infrastructure is one system composed of hardware and software.

Opening and integrating the rights data framework is necessarily a **Multi-Country Project (MCP)** because:

- This development cannot be achieved by one or a few players but covers a multitude of different rightsholders of different sizes and their representatives as well as rights users of both commercial and non-commercial nature – **in 27 jurisdictions**.
- A successful development requires cooperation between holders and users of copyright data, intellectual property offices, copyright registrations, recordation of transfers, legal deposits, and standardisation bodies – **across the Union**.

3) Link to the Digital Decade Policy Programme 2030 general objectives and digital targets

3.1. Corresponding area of activity in the annex of the DDPP decision

The CI Task Force supports the following areas of activity listed in the DDPP Annex:

(a) European common data infrastructure and services, in relation with

(g) Connected public administration[2], and

(h) European blockchain services infrastructure (EBSI)[3]

Following these headings the CI Task Force will

(a) define the Open Rights Data Framework – the set of rules (identifiers, schemas, and architecture) to start building a Common European Copyright Data Space.

(g) take steps to ensure that the Open Rights Data Framework could be used to interconnect European public administrations such as national copyright registration, transfer recordation, and legal deposit systems, as well as national libraries and archives, and intellectual property offices.[4]

[2] The ORDF will indeed facilitate the EU-wide interconnection of National Voluntary Copyright Registration Systems, Recordation of Transfers of Rights, Legal Deposits, Archives, Libraries, and Intellectual Property Offices.

[3] The Multi-Country Project considers indeed the EBSI as a key underlying technology.

[4] In connecting these repositories, we must consider all Commission recommendations. They are also users of IP in other words, users of the CI:

- COMMISSION RECOMMENDATION (EU) 2021/1970 of 10 November 2021 on a common European data space for cultural heritage covers also national libraries and legal deposit systems. IT recognises still a lot of interoperability issues. It recommends among others to use the Europeana Data Model that covers a set of standardised rights statements that can be used by cultural heritage institutions to communicate the copyright and reuse status of digital objects to the public. There should be a reference to compatibility efforts here. See DataSpaceCultHeritage_inclCopyright_EN_TXT.pdf
- COMMISSION RECOMMENDATION of 1 March 2023 on a Code of Practice on the management of intellectual assets for knowledge valorisation in the European Research Area. It recommends for R&I actors (all categories of actors involved in R&I such as intermediaries, individual researchers, innovators and their teams, and

(h) promote the advances of technology programmes supporting the interoperability, searchability, and trustworthiness of rights management information. These cutting-edge technologies include blockchain, digital wallet, digital twins, high-performance computing, and 5G.

3.2. General objectives

The Multi-Country Project is directly linked to the following general objectives of the Digital Decade Policy Programme 2030 (DDPP) listed in the Article 3 of the Decision of the European Parliament and of the Council[4]:

(a)	*promoting a human-centred, fundamental-rights-based, inclusive, transparent and **open digital environment** where secure and **interoperable digital technologies and services** observe and enhance Union principles, rights and values and are **accessible to all**, everywhere in the Union;*	The MCP aims at defining for rights data an **open digital environment** based on **inter-operable digital technologies and services** that is **accessible to all**.
(c)	*ensuring the Union's digital sovereignty in an **open manner**, in particular by **secure and accessible digital and data infrastructures** capable of efficiently **storing, transmitting and processing vast volumes of data** that enable other technological developments, supporting the competitiveness and sustainability of the Union's industry and economy, in particular of SMEs, and the resilience of the Union's value chains, …*	The MCP aims at defining an **open, secure and accessible digital and data infrastructure** capable of efficiently **storing, transmitting and processing vast volumes of rights data**.

organisations including universities, public and private R&I organisations, businesses of all sizes, research and technology infrastructures, public administrations, and civil society representatives should be encouraged to follow this Recommendation to identify ownership of IP as soon as possible and to make a clear IP strategy that allows for open science and open investments. The creation of an environment where intellectual assets management practices are clearly defined, communicated, and implemented is the first step to facilitate their valorisation in the R&I ecosystem. It also recommends awareness and education on sources for information such as to consult "IP databases" for instance. See https://research-and-innovation.ec.europa.eu/news/all-research-and-innovation-news/commission-adopts-recommendations-codes-practice-management-intellectual-asset-and-standardisation-2023-03-07_en

- IP Offices are part of public administration infrastructure work. These are already very developed in some MS of the EU but in Finland for instance the Health Data Space stumbled on difficulties because of lack of interest to share data in public databases. Similar problems have arisen also within data space development/ data economy pilots the education area. See short presentation in English (I do not know the current status) https://vm.fi/en/opening-up-and-using-public-data; links with the EDIC public administration led by Greece. The issues relate often to lack of funding to engage in sharing of data as this is not stipulated as a task of the public entity. Still there are initiatives to build common standards between public and private entities in the Nordics and Baltics. Se short presentation in English. See https://dvv.fi/-/uusi-kasikirja-tukee-tietojen-liikkumista-maiden-valilla-tavoitteena-sujuvampi-kansainvalinen-arki?languageId=en_US

(d)	*promoting the deployment and the use of digital capabilities with a view to reducing the geographical digital divide and **granting access to digital technologies and data on open, accessible and fair terms**, in order to achieve a high level of digital intensity and innovation in Union enterprises, in particular start-ups and **SMEs**;*	The MCP aims at defining an open framework as a standard for systems to **grant access to rights data on open and fair terms**, in order to achieve a high level of digital intensity in the creative industries – i.e., in **SMEs**.
(e)	*developing a comprehensive and sustainable **ecosystem of interoperable digital infrastructures**, where high performance, edge, cloud, quantum computing, artificial intelligence, data management and network connectivity work in convergence, to promote their uptake by businesses in the Union, and to create opportunities for growth and jobs through research, development and innovation, and ensuring that the Union has a competitive, secure and sustainable data cloud infrastructure in place, with high security and privacy standards and complying with the Union data protection rules;*	The MCP aims at defining an open framework as the basis for a **comprehensive and inter-operable rights data infra-structure** leveraging the convergent advances of programmes such as EBSI, eIDAS, TEMS, and EuroCloud, and facilitating the growth of the creative industries, whilst protecting data privacy and business confidentiality.
(f)	*promoting a Union **digital regulatory environment** to support the ability of Union undertakings, especially that of SMEs, to compete fairly along global value chains;*	The ORDF will support a secure, interconnected, interoperable, and **regulated Digital Single Market**.
(h)	*ensuring that **digital infrastructure and technologies, including their supply chains, become more sustainable, resilient, and energy- and resource-efficient**, with a view to minimising their negative environmental and social impact, and contributing to a sustainable circular and climate-neutral economy and society in line with the European Green Deal, including by promoting research and innovation which contribute to that end and by developing methodologies for measuring the energy and resource efficiency of the digital space;*	Whilst defining underlying **digital infrastructure and technologies** for the ORDF, the MCP will ensure that they fulfil **resilience and sustainability objectives**.
(j)	*ensuring that all policies and programmes which are relevant to achieving the digital targets set out in Article 4 are **taken into account in a coordinated and coherent way** to fully contribute to the European green and digital transition, while avoiding overlaps and minimising administrative burdens.*	The MCP will **coordinate its work with the relevant policies and programme** and liaise with the **European Data Innovation Board** established by the Data Governance Act 2022.

3.3. Digital targets

The Multi-Country Project is directly linked to the following digital targets of the Digital Decade Policy Programme 2030 listed in the Article 4 of the Decision of the European Parliament and of the Council:

(3) *the digital transformation of businesses, where:*	
(a) at least 75 % of Union enterprises have taken up one or more of the following, in line with their business operations: (i) ***cloud computing services****, (ii)* ***big data****, (iii) artificial intelligence;*	The ORDF layer of the Copyright Infrastructure will enable **innovative scale-ups** of the Union to share the advance of **cloud computing** and **big data** with the multitude of **individuals and SMEs** that compose the creative industries and contribute to raising their **digital intensity**.
(b) more than 90 % of Union SMEs reach at least a basic level of ***digital intensity****;*	
(c) the Union facilitates the ***growth of its innovative scale-ups*** *and improves their access to finance, leading to at least doubling the number of unicorns;*	
(4) *the digitalisation of public services, where:*	
(a) there is 100 % ***online accessible provision of*** *key* ***public services*** *and, where relevant, the possibility for citizens and businesses in the Union to* ***interact online with public administrations****;*	The ORDF will enable the open **access to and online interaction with public services** related to rights data, see footnote 1 on page 19.
(c) 100 % of Union citizens have access to ***secure electronic identification (eID)*** *means that are recognised throughout the Union, enabling them to have* ***full control over identity transactions and shared personal data****.*	The ORDF will make **secure electronic identification** a prerequisite to user and rights-holder authentication and put them in **control of their data**.

4) Detailed description of the Multi-Country Project

A successful **data framework** is minimally prescriptive but maximally inclusive. It supports many different solutions, past, present, and future. It enables numerous ways in which individuals and organisations can cooperate in creating, enriching, governing, and distributing trusted copyright information. It helps to streamline current processes and trigger innovative businesses. It is of potential benefit to everybody in the ecosystem.

Parts of an international **rights data framework** already exist, but it has significant gaps and weaknesses and lacks interoperability across different media or content sectors. Its structure includes a number of established standards and technologies which underlie the exchange of

rights management information. The existing framework supports a fragmented network of rights declarations, attributions, verifications, and queries in the digital era. It varies greatly from sector to sector and relies too much on labour-intensive human interactions which can no longer cope with the volume of available content and types of use.

Opening and integrating the **rights data framework** in close collaboration with relevant parties will boost the value of the content sectors significantly over time. **The task is to make the framework trustworthy, automatically interoperable, and as accessible and comprehensive as possible. To do so it must be made extensible, able ideally to support any business model for any sector in the any jurisdiction of the European Union and beyond**.

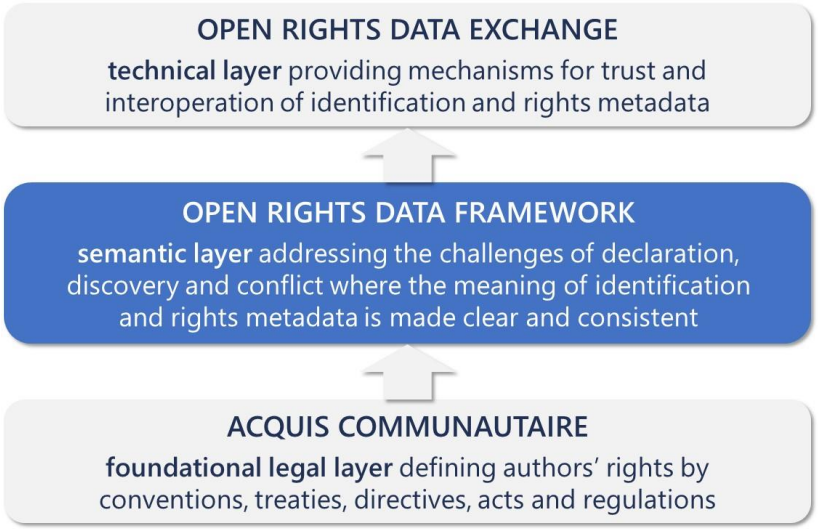

Figure 2: Position of the Open Rights Data Framework

Therefore – and coming back to the definition of the copyright infrastructure – the **Copyright Infrastructure Task Force, representing ideally all EU Member States and other public and private entities, will define the sets of related rules, technologies, and institutions, promote these definitions, and govern them**.

4.1. Set of rules

The semantic layer brings data together. Its building blocks are identifiers and schemas (the formats in which metadata is captured). There are many types of these in use, standard and proprietary. The ORDF will be designed to combine any and all of them. Based on some of the best work done in metadata standards in the last 25 years, the ORDF will have an innovative data architecture able to integrate rights and content data of any kind or complexity.

The design of this data architecture will also expose a remarkable black hole at the heart of the current rights data framework: existing standards and systems do not identify a right as a distinct data entity. Why does this matter? It means that unlike creations, parties or licences, rights do not

have identifiers – and computers cannot talk unambiguously about things unless they have identifiers. The availability of right identifiers will enable the automation of licensing, distribution, and remuneration and make these processes simpler, faster, more accurate and less expensive.

4.2. Set of technologies

Augmenting the interoperability, searchability, and trustworthiness of rights management information is also the matter of underlying technologies.

As mentioned, the CI Task Force will leverage the convergent advances of technology programs such as EBSI, eIDAS, TEMS, and Eurocloud, and more generally cutting-edge technologies listed at the base of the Digital Decade comprehensive framework: digital twins (to digitally represent physical content such as printed books), high-performance computing (to handle the very high volume and frequency of rights data transactions), digital wallet (to authenticate users, rightsholders and rights users whilst restoring their data sovereignty), blockchain (to safely record and distribute digital information) and 5G (to facilitate the capture and search of rights data).

4.3. Set of institutions

The ORDF semantic model and ontology will be designed to represent any rights transaction or declaration, including those under different jurisdictions or schemes which conflict with one another.

Building a pervasive, open, and inclusive rights data framework at a European and then at a global scale will require cooperation and coordination between –

- **international standards organisations**, including the European Telecommunications Standards Institute (ETSI), the International Organization for Standardization (ISO), the World Wide Web Consortium (W3C)[5], Creative Commons, etc.,
- **industry standards organisations** such as the International Confederation of Societies of Authors and Composers (CISAC), Digital Data Exchange (DDEX), ONIX, the International Press Telecommunications Council (IPTC), etc.,
- IP offices[6], collective management organisations, national libraries, voluntary copyright registration systems, and legal deposits schemes, which all maintains **large repository of rights management information**,
- associations of **rightsholders**, and
- associations of **rights users**.

[5] The CI Task Force will particularly consider their Open Digital Rights Language (ODRL), a policy expression language that provides a flexible and interoperable information model, vocabulary, and encoding mechanisms for representing statements about the usage of content and services.

[6] The CI Task Force will also consider public domain works, orphan works, and out-of-commerce works and how their existing repositories could be interlinked through a standardised protocol to make these works discoverable.

The CI Task Force will **complement** their work and build – with and for them – a common fundament to assure interoperability and searchability of rights management information across sectors, jurisdictions, and devices, hence, to support the emergence of a true Digital Single Market. The CI Task Force will not create standards itself but will **coordinate requirements and resources** to foster the creation and evolution of standards within standards organisations working in relevant domains.

The CI Task Force will **assure the governance** of the ORDF and collaborate with the European Data Innovation Board. The semantic governance will be overseen by experts from participating standards bodies (e.g., CISAC, DDEX, DOI, IPTC, ISO/TC 46/SC 9, and ONIX), rightsholders and platforms. It will direct the semantic policy.

5) Main tasks, activities, and deliverables

The CI Task Force will define, promote, and govern the Open Rights Data Framework (ORDF).

5.1. Definition

This task will consist of the following activities:

- define use cases and requirements,
- identify, name, and describe necessary –
    - content, party, and rights identifiers,
    - metadata sets,
    - data exchange protocols,
- identify, name, and describe technologies fostering the interoperability, searchability, and trustworthiness of rights management information,
- identify the relevant standardisation bodies,
- participate in the relevant working groups of these standardisation bodies at observer or contributor level.

5.2. Promotion

This task will consist of the following activities:

- set up and maintain a comprehensive interactive website publishing the results of the definition task,
- publish in relevant journals,
- plan and deliver awareness campaigns,
- produce education materials and liaise with education providers,
- organise a yearly conference to be attended in person or online,
- participate in relevant conferences.

5.3. Governance

This task will consist of the following activities:

- maintain and develop the ORDF identifiers, schemas, and vocabularies (including the right identifier, right attributes, and right declaration).
- maintain the ORDF ontology, including authorise mappings of schemas/vocabularies; mappings will also have to be authorised by the management of the schema/ vocabulary; following the successful model of DDEX, the ontology will store all terms used in ORDF schemas and vocabularies.
- assure that ORDF standards will remain neutral in respect of jurisdiction and commercial interests.
- assure that ORDF standards will be able to represent the detail of any and all statutes or agreements governing content rights.

5.4. Deliverables

The Working Group on identifiers and schemas will deliver an **innovative architecture of identifiers and schemas**.

The Working Group on technologies for interoperability, searchability, and trustworthiness will deliver **copyright-specific, current, and interactive developer guidelines**.

The Working Group on dissemination, promotion and education will ensure the **public and private adoption of the ORDF**.

Each Working Group can define focused **Task Groups** to cater for the size and complexity of the tasks at hand, and/or to consider the specificity of some creative sectors.

6) Implementation strategy

6.1. Objectives, use cases, KPIs, and risks

6.1.1) Objectives

The CI Task Force shall implement an ORDF, i.e., the semantic layer of a copyright infrastructure[7] addressing the challenges of declaration, discovery, and conflict; the layer where the meaning of identification and rights metadata is made clear and consistent.

The CI Task Force shall define, promote, and govern an ORDF that is minimally prescriptive but maximally supportive and inclusive; a framework that will allow many solutions to be used while enabling numerous ways in which one can cooperate in originating, enriching, governing, and distributing trusted rights management information, helping streamline current processes and trigger innovative business models.

The ORDF will be conducive of a **real-time access to reliable, exhaustive, current, and interoperable rights data** that will support rights **management, licensing, enforcement, and remuneration**.

For this purpose, the CI Task Force shall carry out the following **phased** activities:

1) define an innovative architecture of identifiers and schemas to integrate rights and content data of any kind or complexity[8],
2) develop current and interactive developer guidelines about technologies supporting interoperability, searchability, and trustworthiness of rights management information,
3) disseminate and promote the results from activities 1) and 2),
4) govern the ORDF defined in activity 1).

[7] The set of rules, technologies and institutions that frame data management practices in the creative industries.
[8] This is already technically feasible, but it still requires much work at the semantic level of the Copyright Infrastructure to cover challenging issues such as the dynamic collaborative creation, the graph representing multiple contributions used in multiple configurations, and the qualification of trusted service providers.

6.1.2) Exemplary use cases

News publishers – European press articles are consolidated on American online platforms and used to trained large language models. What are the necessary interoperable content identifiers for articles, texts, pictures, titles, and graphics and the standard datasets for opt-outs, terms, and conditions to facilitate the enforcement of rights related to Articles 4 (**Text and data mining**) and 15 (**Protection of press publications**) of the EU directive on copyright? Is the documentation on the training and generation algorithms used to produce AI-generated content transparent? Are easy and efficient opt-out mechanisms available? What about the identification of content that has been generated by AI? Need for intervention: high, urgency of intervention: high.

Videogames – A videogame is a complex construct of graphics, sound, characters, content, programming (goals and objectives, rules and instructions, interaction, conflict, competition, challenge, opposition, outcomes, and feedback), and more. Most of these individual items are produced by different individual contributors. How can the graph of components be represented to facilitate the **management** of related rights across creative sectors and jurisdictions? Need for intervention: high, urgency of intervention: high.

Spatial computing – With a highly dynamic and interactive "metaverse" everybody is a co-creator. The distinction between authors and users becomes fuzzy. One needs to cater for the registration of collaborative rights. Avatars can be digital twins. Who owns them? Who can do what with them? Need for intervention: high, urgency of intervention: high.

Educational content – The value of a licence for one teacher to use once one document to teach one class might be close to €5.00 but the cost of a licence, even with good IT support, will not be less than €20.00. How can the author of the document be fairly, appropriately, proportionally, and transparently remunerated in a timely fashion through a sustainable **licensing** process that prevent the emergence of black boxes and the use of inaccurate flat rate levies? Need for intervention: medium, urgency of intervention: medium.

Music and film – Swedish and Portuguese songwriters create a song in Germany. The song is performed by Croatian and Slovenian musicians in a Spanish recording studio. The recording is used by filmmakers in Italy. The film is broadcast – among others – in Greece. All of this happens within 2 months. The Swedish, Portuguese, Croatian and Slovenian music makers are not members of any management organisation in Greece. How can they be fairly, appropriately, proportionally, and transparently **remunerated** in a timely fashion? How will **Article 17** of the EU directive on copyright be practically implemented for everyone? Need for intervention: medium, urgency of intervention: medium.

Out-of-commerce works – are millions of works that are still protected by copyright but end up being considered out-of-commerce, such as literary works, audio-visual works, phonograms, photographs, and unique works of art. Providing access to this rich European cultural heritage relies on swift and affordable rights clearance and licensing process. Need for intervention: medium, urgency of intervention: medium.

6.1.3) Relationships with Europeum EDIC, Alliance for Language Technologies (ALT EDIC), eIDAS, Trusted European Media Data Space (TEMS), EuroCloud, etc.

The EU Strategy for Data proposes the deployment of Common European Data Spaces. Some of them will support directly or indirectly the creative industries. Various EDICs and initiatives of the Digital Decade will contribute to the emergence of these Data Spaces.

The Working Group "Technologies for interoperability, searchability and trustworthiness" of the CI Task Force will collaborate closely with initiatives such as Europeum EDIC, ALT EDIC, eIDAS, TEMS, and EuroCloud to create copyright-specific, current, and interactive technical guidelines for:

- data solutions that are neutral, incorruptible, and trustworthy,
- party authentication,
- the advance of self-sovereign identity and data sovereignty,
- the respect of individual privacy and business confidentiality,
- backward-compatibility and futureproofing.

The close collaboration between the CI Task Force, EDICs, and other initiatives of the Digital Decade can lead to the deployment of a Common European **Copyright** Data Space, i.e., technical infrastructures. The CI Task Force will provide a copyright-relevant data architecture and contribute to the selection of the most appropriate technologies to realise these technical infrastructures.

Practically, this close collaboration can lead to consortiums[9] conducting **Large-Scale Pilots** (LSPs) under current data regulations[10] in different creative sectors / data spaces and with different technological solutions to pave the way to technical infrastructures, gather feedback, and fine-tune the Digital Policy Programme concept, system, and data requirements.

6.1.4) Key performance indicators

Numbers of –

- Member States participating actively in the working groups of the CI Task Force,
- international or industry standardisation bodies participating in the working groups,
- creative sectors engaged in and impacted by the activities of the CI Task Force,
- copyright registrations based on the Open Rights Data Framework of interoperable identifiers of work, related subject matters, authors, rightsholders, and rights,
- connections of existing data repositories such as national libraries, copyright registries based on the Open Rights Data Framework.

Timely delivery of –

- an innovative architecture of identifiers and schemas,

[9] Like the consortiums EUDI Wallet Consortium (EWC), NOBID, Potential, and DC4EU carrying Large Scale Pilots for eIDAS.

[10] The new data laws or data space initiatives cover indeed areas that have already established rules.

- copyright-specific, current, and interactive user references, and
- promotion and education collaterals.

Public and private adoption of the ORDF.

6.1.5) Risk assessment

The main risk is the lack of commitment and availability of governmental and private experts. The probability is medium, and the impact would be important. Therefore, we are scheduling a thorough planning phase.

6.2. Modalities

6.2.1) Best practices

The CI Task Force will act as a **Standards Forum**, define use case and requirements, and liaise with **Standards Development Organisations** such as ETSI. The standards triggered and/or recommended by the CI Task Force will be voluntarily adopted on a global level. They will serve as building blocks for the Open Rights Data Framework to meet the needs of creative industries and users. The use cases, requirements and recommendations of the CI Task Force will be developed through an open, participatory process, support interoperability, and foster global competition.

The CI Task Force will adopt and adapt **best practices** implemented by other standards **forums** such as eIDAS or the Metaverse Standards Forum.

The CI Task Force will adhere to the five key **OpenStand** principles[5]:

1) **Cooperation**: respectful cooperation between standards organisations, whereby each respects the autonomy, integrity, processes, and intellectual property rules of the others,
2) **Adherence to the principles** of due process[11], broad consensus, transparency, balance, and openness,
3) **Collective empowerment** striving for standards that based on technical merit; provide global interoperability, scalability, stability, and resiliency; enable global competition; serve as building blocks for further innovation; and contribute to the creation of global communities, benefiting humanity,
4) **Availability** on fair terms that may vary from royalty-free to fair, reasonable, and non-discriminatory terms,
5) **Voluntary adoption** whereby success is determined by the market.

[11] Including votes.

The CI Task Force will create '**role books**' that specify 'who-does-what-by-when, what-does-it-take, and what-does-it-cost' and harmonise national 'role books' to improve coordination and communication. Role books will be particularly useful in clarifying the interactions with new agencies and bodies, such as the European Data Innovation Board or European Data Infrastructure Consortiums (EDICs). For these to have a significant impact at the EU level and improve member state-to-member state communication, 'role books' will be interoperable, give up on the one-size-fits-all approach and respond to the needs of sectoral data spaces.

6.2.2) Large-Scale Pilots

The CI Task Force will leverage Large-Scale Pilots to anchor its work in the reality of the creative industries and immediately test its recommendations.

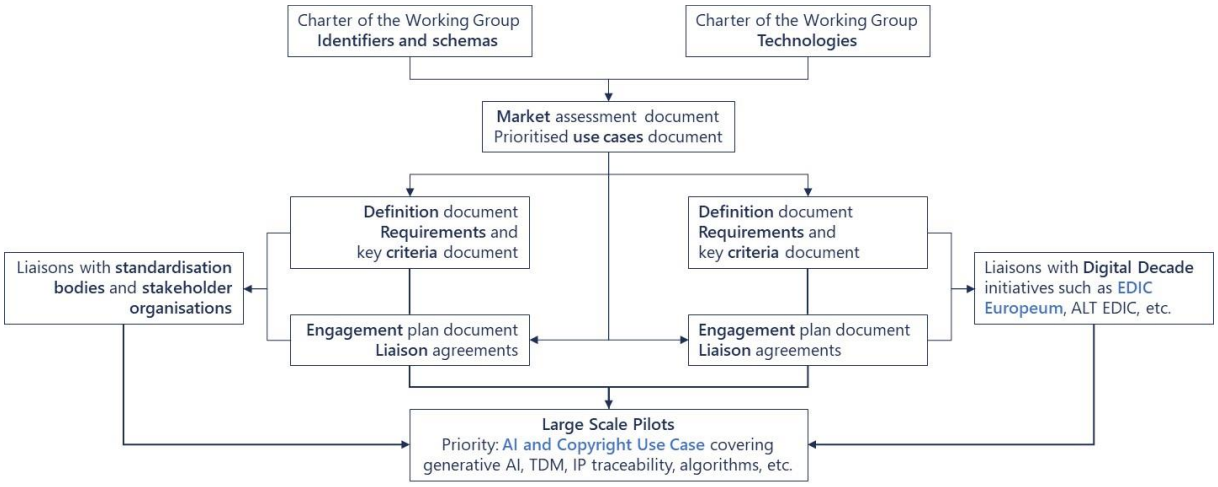

Figure 3: Setting up a Large-Scale Pilot

6.2.3) Reaching out activities

Reaching out activities started from the moment the Estonian and Finnish Government established the CI Task Force in Fall 2022. From Day One, delegates from the Portuguese and Slovenian Governments joined the task force as observers. These Member States are members of Europeum EDIC.

Supported by the Commission, reaching out activities included meetings with delegates from the French, German, Latvian, and Lithuanian Governments in Spring 2023.

Whilst preparing the pre-notification, the Finnish Ministry of Culture and the Estonian Patent Office reached out to management organisations[12].

[12] For example, the Finnish Ministry of Culture reached out to Gramex (phonogram producers), Kopiosto (educational institutions, businesses, and public administration), Kuvasto (visual arts), and Teosto (authors and composers). All of them welcomed the CI Task Force initiative.

During the pre-notification and application phases, the CI Task Force reached out to –

- the Member States delegates at the Copyright Working Party of the Council,
- standardisation bodies,
- organisations maintaining large repositories of rights management information in each of the applying Member States, and
- Digital Decade initiatives such as Europeum EDIC and eIDAS.

That activity aims at involving these parties in the scoping of the Strategic Orientation Committee, Advisory Boards and Working Groups, and in the definition and prioritisation of use cases.

6.3. The AI & Copyright Use Case

This section describes legal and technical aspects of the use case at hand of the Open Rights Data Framework (ODRF) by a Use Case Group (UCG) proposed by the Copyright Infrastructure Task Force (the CI task Force) to Europeum EDIC.

Considering –

- the need and urgency to equip the creative industries with adequate tools to face challenges arising from Large Language Models and Generative AI, and
- the current work at EBSI, whereby Intellectual Property is a use case for the Pre-Commercial Procurement and an Open Rights Data Exchange is an application scenario for TRACE4EU, the EBSI traceability project,

the CI Task Force suggested leveraging an AI & Copyright Use Case to –

- launch its operations,
- conduct a Large-Scale Pilot, and
- work closely with Europeum EDIC while potentially becoming one of its Use Case Groups.

6.3.1) EU Directive 2019/790 on Copyright in the Digital Single Market (DSM directive)

The use of copyright protected works and other subject matter for reproduction or communication to the public requires the authorization of the rightsholder concerned unless relevant copyright exceptions and limitations apply.

Articles 3 and 4 of the Directive (EU) 2019/790 (the DSM Directive) provide for two exceptions to the exclusive right of reproduction under copyright and the right to prevent extraction under the database right when those acts are committed for the purpose of text and data mining (TDM) as defined in Article 2(2) DSM Directive. The copyright on protected works and the related rights on other subject matter (including databases and press publications) are subject to those exceptions. The directive has been implemented into national law by almost all EU Member States.

The definition in Article 2 of DSM on text and data mining is *"text and data mining' means any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations"*.

Under Article 3, TDM is allowed for scientific research by research organisations and cultural heritage institutions under the condition of lawful access to the protected work or subject matter.

Under Article 4, TDM for any other purpose, including for commercial purpose, is allowed when the works and other subject matter are lawfully accessible. According to Article 4(3) the exception applies on condition that the use "*has not been expressly reserved by their rightsholders in an appropriate manner, such as machine-readable means in the case of content made publicly available online*". This is referred to as a possibility to declare an "opt-out" from the application of the exception or limitation.

Especially since the flush of generative AI tools made available online, rightsholders feel the need to reserve their rights and to opt-out of the exception in certain situations. The DSM Directive requires the reservation to be "expressly reserved by their rightsholders in an appropriate manner". Other than the reference to machine-readable means for online content, the provision does not detail how the opt-out should be declared. Recital 18 of directive (EU) 2019/790 states: "In the case of content that has been made publicly available online, it should only be considered appropriate to reserve those rights by the use of machine-readable means, including metadata and terms and conditions of a website or a service. […].

The Artificial Intelligence Act (AI Act) provides for directly applicable provisions in relation to copyright protected works. The Act was approved by the Council on 2 February and by the European Parliament on 13 March. The Act normally comes into force after a 12 month-period of its adoption, *i.e.* sometime in 2025. The regulation is expected to be finally adopted before the end of the legislative period. The Act also needs to be formally endorsed by the Council. It will enter into force twenty days after its publication in the official Journal, and be fully applicable 24 months after its entry into force, except for: bans on prohibited practices, which will apply six months after the entry into force date; codes of practice (nine months after entry into force); general-purpose AI rules including governance (12 months after entry into force); and obligations for high-risk systems (36 months).

According to the AI Act, providers of general purpose AI (GPAI) models shall "*put in place a policy to respect Union copyright, in particular to identify and respect, including through state of the art technologies, the reservations of rights expressed pursuant to Article 4(3)*" of the DSM Directive. "State of the art technologies" could be considered to mean blockchain or distributed ledger technology, which the EBSI – European Blockchain Service Infrastructure provides for.

Furthermore, providers of GPAI models must "*draw up and make publicly available a sufficiently detailed summary about the content used for training of the GPAI model, according to a template to be provided by the AI Office*". The summary obligation should be generally comprehensive in its scope (instead of technically detailed), and thus include the general, main content sources/sets for training the GPAI in order to facilitate parties with legitimate interests (including copyright holders) in exercising and enforcing their rights. This could for instance be achieved by listing the main private or public databases/archives used and by providing an account of the other sources used. The AI Office is expected to provide a template for the summary, which should be a simple

and effective way to provide the required summary in a narrative form. The AI Office was established by a Commission decision of 24 January 2024. It would be important to ensure collaboration between the AI Office and the CI Task Force at a suitable moment once it is fully operational.

6.3.2) Material scope of the AI & copyright use case

Text and data mining techniques may be used extensively for the retrieval and analysis of various types of content, which may be protected by copyright and related rights. Training of AI models require vast amounts of data. Most of the data composing works and other content is protected by copyright. The TDM exception is one important way to ensure that AI models can be trained with copyright protected content. Permission to use the content is provided through a provision in the law, based on certain caveats.

The AI & copyright use case (later "AI use case") will focus on the first of the obligations for the GPAI *i.e.* the identification and respect of the opt-out declared by rightsholders with regard to the and data mining, unless this is done for the purposes of scientific research. Ideally, the ORDF could be developed and piloted at a later stage with regard to the summary requirement or any other use case in relation to the proper functioning of the copyright system in the digital environment such as Art 17(4) or Art 8(4) of the DSM Directive.

The AI Act does not specifically provide that a new standard should be developed, but rather "a policy to respect Union law on copyright and related rights, in particular to identify and respect the reservations of rights expressed by rightsholders pursuant to Article 4(3) of Directive (EU) 2019/790". However, in practice, in order to deploy an opt-out mechanism in the Digital Single Market a standard needs to be set on EU level. This is proposed to be developed through the Open Rights Data Framework. This would also be very important from legal certainty as well as growth and competitiveness point of view of the European companies active in training of AI.

Piloting the opt-out mechanism is a complex matter.

For example, to know if copyright applies in the first place and whether there is a need to seek permission, the mechanism would need to establish the "copyright status". In implementing these rules, this would be "true": "in copyright" or "false": "out of copyright" or "in public domain"). Status should, ideally, be based on the date of death, if applicable, linked with the term of expiration of the rights.

Only then some rights apply and can be managed by the rightsholder in accordance with Article 7 of Directive 2001/29/EC[13].

[13] Article 7. Obligations concerning electronic rights management information. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32001L0029

Terms and/or exceptions must also relate to a specific right, in this case the right to reproduction. This could translate to "reproduction" true: yes/false: no. Then an exception to the exclusive right could be determined; "TDM" true: yes / false: no; The same would go for context "commercial" (true: yes / false: no) and finally the opt-out use based on exception "AI training" true: yes / false: no.

Taking all these elements into consideration shows how the use case is complex.

While legal aspects should not make the piloting of the opt-out mechanism overly complex, neither should the use case be conducted merely as a technical endeavour. In developing the requirements for the use case certain key aspects were presented as a basis for discussion at the CI Task Force meeting on 26 March. Please find below some considerations on the scope of the use case.

Firstly, the use case should focus on the opt-out declaration that is made in a machine-readable way. For online content, the opt-out must be expressed in a machine-readable way. This means that the opt-out declaration is a statement included in the metadata of the work that can be read by the search robot when searching for content to train an AI model.

Secondly, the use case will pilot the right to reserve TDM for other purposes than scientific purposes.

The initial purpose of the TDM exception in the proposal of the Commission was to allow the analysis of large masses of data for the purpose of scientific research. Only later it was extended to also cover other purposes, such as TDM for commercial purposes, including training of AI. Although the exception could cover instances of training of GPAI, it is not well suited for this purpose. This is particularly true with regard to generative AI systems, as these systems produce new content (and revenue) based on the content they were trained with.

Thirdly, the opt-out should be piloted keeping in mind the scope of the TDM exception itself, taking into account the so-called three-step-test[14].

Therefore, the opt-out mechanism should be limited to cases that are likely to be in line with the requirements under Art 7(2) of the DSM directive[15] when they were implemented[16]. In accordance with Article 5(5) of the Information society directive 2001/29/EC, use based on a limitation is allowed only in "*certain special cases which do not conflict with a normal exploitation of the work*

[14] See https://en.wikipedia.org/wiki/Berne_three-step_test

[15] Art 7.2 (2): Article 5(5) of Directive 2001/29/EC shall apply to the exceptions and limitations provided for under this Title. The first, third and fifth subparagraphs of Article 6(4) of Directive 2001/29/EC shall apply to Articles 3 to 6 of this Directive.

[16] In the doctrine interpreting the three step the conformity with the three steps test is assessed at adoption of the provision providing legal certainty for the user of the exception. If the threshold would be assessed in individual cases, the application of it would not be serving the very purpose of the exception itself.

or other subject-matter and do not unreasonably prejudice the legitimate interests of the rightsholder".

Hence, there are three cases in which an opt-out would not be relevant, even if it is expressed in a machine-readable way.

One would be the case where the conditions for the exception for scientific research are met and no opt-outs are allowed. This is confirmed also under Art 4(4), which states that the exception in Art 4 must not impact the application of TDM under Art 3.

Another case is when a) the training of the AI could not be considered text and data mining (under Art 4) or b) when the use is otherwise considered not to pass the three-step test.

So far the TDM exception has not been broadly analysed by courts on EU level. Some development is taking place in Germany[17] regarding application of the TDM exception for training of AI models that are used for offering generative AI systems. Such systems normally require the consent of the authors. On the other hand, in United States alone multiple court cases are addressing topics relating to generative AI and copyright[18].

Consequently, the use case would normally explore how the opt-out could technically be declared in a way that takes into consideration not only the opt-out but also the purpose of the use and its lawfulness under the exception.

The user does not need to be identified, only the purpose. Anyone can apply the TDM provision under Art. 4(3), subject to the opt-out by the rightsholders and in the case of GPAI models, subject to the identification and respect of the reservations of rights. However, it is a key element of the data economy and new technologies that each entity participating in the sharing of data, is also identified with an open and interoperable identifier.

6.3.3) Geographical scope of the pilot

The use case will focus on applying the opt-out on uses of works for training an AI model in the EEA. The four to six EU Member States, which would participate in the pilot must have transposed the DSM Directive in their national copyright act. Indeed it is the copyright law of a Member State that applies to the reproduction of a work, taking place in its national territory.

Member States could include Finland, Estonia, and members of Europeum EDIC such as Italy, and also Latvia and Lithuania, subject to confirmation by each MS that they want to participate in the use case. Relevant provisions in the respective copyright acts are referred to below.

[17] Some development in this regard is taking place; see https://cepic.org/news/an-up-date-on-the-robert-kneschke-v-laion-e-v

[18] See https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html

For the obligation to respect opt-out declarations to be effective, however, it would be important to apply it regardless of the jurisdiction in which the copyright-relevant acts take place. Training of AI *i.e.* the reproductions for that purpose, are likely to first take place outside the EU (at least for the models developed by leading actors in the USA).

There are no harmonised rules on training of AI at international level and the issue remains open in many jurisdictions.

6.3.4) Creative sectors best suited for the pilot

Selecting the creative sector to be piloted with the ORDF is also important. Keeping in mind the previously expressed requirements, there are several appropriate candidates within the creative sectors for an AI & copyright use case.

The **book-publishing sector** has been active in developing the copyright infrastructure and there is a high percentage of open and interoperable identifiers (ISBN, ISSN, DOI, ISNI[19], ISCC[20]) in use in this sector. The book industry represents European stories and is of immense value for the European cultural heritage. The new mechanism to assist accessibility of out of commerce works online (out of commerce works portal) was promoted through the decision on the Cultural Heritage Data Space. According to preamble 17, Europeana has been key in strengthening cooperation and standardisation activities across borders, in the EU and beyond. Its standardised frameworks for sharing digital content and metadata online, in particular, the Europeana Data Model, Rights Statements and the Europeana Publishing Framework should be used to ensure access and to improve interoperability of works made available on Europeana. The open availability of works online could on the other hand make European cultural heritage still in-copyright vulnerable for reproduction for training of AI without a possibility to opt out. The ORDF is key in ensuring that these works would be safe from TDM for commercial purposes, still allowing TDM for scientific purposes.

We could also envisage a specific sub-sector of the publishing industry, like the scientific publishing sector. The **scientific journals** may have several identifiers. Some are identified with the ISNI, some with the ORCiD identifiers. The newly introduced, DLT-ready ISCC identifier could work very well for identifying works that have several versions such as articles and journals. These journals are valuable training material to LLM's focusing on the production of new content based on that information, such as ChatGPT more focused on re-use of data than on the content itself. This would be an interesting area as scientific articles would be largely used for research purposes but also for training GPAI. On the other hand, scientific authors are under obligations to go for Open Access publications and make their papers available online (on university repositories for example that are used for TDM). Most uses of the repositories will be for scientific research (not for commercial purposes). The opt-out in this sector is not likely to be declared by authors. It

[19]More on ISNI: https://isni.org/
[20] More on ISCC: https://iscc.codes/

would also not be consistent with the purpose for Open Access to promote declaring opt-outs by the publishers in this field.

The **photographic industry** is very vulnerable in the digital environment as to infringing use of images made available online and stripping of metadata indicating the intent to attest copyright. Photographs have normally only one creator. Photographers have also chosen to use commercial image banks or open source models to promote the use of their works. AI models like the ones providing new images based on existing images for advertising could cause significant harm to the industry. What kind of AI model would use photographs to train it but would also be justifiable under the TDM exception from the perspective of the rightsholders? The DSM Directive clarifies the status of works of visual art in the public domain with the objective to increase legal certainty. This means that for photographs, a point to consider is whether it would be necessary to ensure that metadata also captures the possibility of providing status of copyright (author is alive, or death date) including if it is in public domain. One interesting factor to consider is the benefit of the use case would have in relation with the prevention of fake news and dis- and misinformation.

Another possibility for a specific sector is **video games**. There, the rights are kept in one place, with the developers. On the other hand, video games are not identified with open and interoperable identifiers.

The **music sector** would also be an interesting area for the use case. Text and data mining is also made on music. AI applications are used by a broad array of users, including for commercial purposes. However, considering the complexity of the use case on an opt-out as such, the additional complexities could risk slowing down the progress of the use case. The sheer number of rightsholders concerned (e.g., authors, publishers, performers, producers) and the availability of various partly overlapping, not fully interoperable identifiers, may also limit the benefit of the use case in the larger context.

6.3.5) Defining legal requirements through examples

It is important to separate the actual reproduction of the works, *i.e*. the copyright relevant acts from the right of the author to use rights management information in order to manage their exclusive rights or provide machine-readable information about terms of use.

The below examples aim to clarify what kind of situations opt-out declarations would be best suited for. Application areas are being increasingly identified within different creative sectors and also within the public sector. For instance, the ALT EDIC Alliance for Language technologies aims to support the building of LLMs in the European languages. These instances may be non-commercial but could also cover AI systems developed on basis of the AI models for commercial purposes.

The use cases presented below are hypothetical, but they demonstrate that it may be important to specify whether and on what terms the information registered to an Open Rights Data Exchange (ORDE) may be used later, in another case of training of AI or for another purpose.

Exemplary pilot 1: training of a large language model (LLM) on scientific publications in [European] languages

An AI developer wants to train an AI model with scientific journals in selected European languages to provide a plugin to the national "ChatGPT" service. 1,000 articles of scientific journals in selected languages that are in copyright are registered at the ORDE. A minimum identifier and metadata set is agreed at the ORDF level (by the use case group) and realized at the EBSI-based ORDE level where data can be verified.

1. Each article is identified with the ISCC and linked with existing party and content IDs.
2. Each author of these articles obtains a creator credential from a Qualified Trust Service Provider (QTSP)[21]. The trust list is published by the EU and covers ISNI, DOI etc. providers).
3. A standard digital rights language is used to register the terms of use, including eventual opt-out expressions.
4. Verifiable credentials (VCs) are used and linked together; verifiable declarations are stored in a discoverable and interpretable distributed registry.
5. Digital declaration certificates are issued and can be stored in the authors' digital wallets.
6. New open standards are developed.

A chat service producing content in European languages would initially fall in the scope of the TDM exception. The purpose of the identification and registration of the works does not require reproduction of protected content. The purpose covers still commercial uses, i.e. the system that can be developed using the AI model can be used for a generative AI system used for/by also commercial actors. It should be noted that the above mentioned ISCC is based on a hash and is also derived (generated) from the digital content itself.

Exemplary pilot 2: training of AI with in-copyright images from [European] countryside

An AI provider wants to develop an AI model to produce images from European countryside. Photographers register 1,000 images/photographs of the countryside at the ORDE. A minimum identifier and metadata set is agreed at the ORDF level and realized at the EBSI-based ORDE level where data can be verified. Participants decide themselves whether to reserve their rights ("opt out") from TDM for commercial purposes (training of AI) or to allow reproduction for any purpose.

[21] Under the Regulation on electronic identification and trust services for electronic transactions in the internal market (eIDAS Regulation), national trusted lists have a constitutive effect. In other words, a trust service provider and the trust services it provides will be qualified only if it appears in a trusted list. Users, including citizens, businesses and public administrations, will benefit from the legal effect associated with a given qualified trust service only if the latter is listed as qualified in the trusted lists. Issuers of party identifiers (e.g., ISNI) or content identifiers (e.g., CISAC) could become QTSPs thereby raising the trustworthiness of their credentials. See https://digital-strategy.ec.europa.eu/en/policies/eu-trusted-lists

1. Each image is identified with the ISCC and linked with existing content IDs.
2. Each creator/photographer obtains a creator credential from a Qualified Trust Service Provider. The trust list is published by the EU and covers ISNI, DOI etc. providers).
3. A standard digital rights language is used to register the terms of use, including eventual opt-out expressions.
4. Verifiable credentials (VCs) are used and linked together; verifiable declarations are stored in a discoverable and interpretable distributed registry.
5. Digital declaration certificates are issued and can be stored in the creators'/photographers' digital wallets.

The AI model could be used for AI systems that generate images for advertising or for production of synthetic content on the metaverse. It should be further discussed on what terms (if any) such a model could be used for commercial AI services. With the help of the copyright infrastructure the follow-up terms of use or licensing of works would be much more streamlined and easy, whenever the purpose would not be allowed based on an exception alone.

6.3.6) Technical considerations

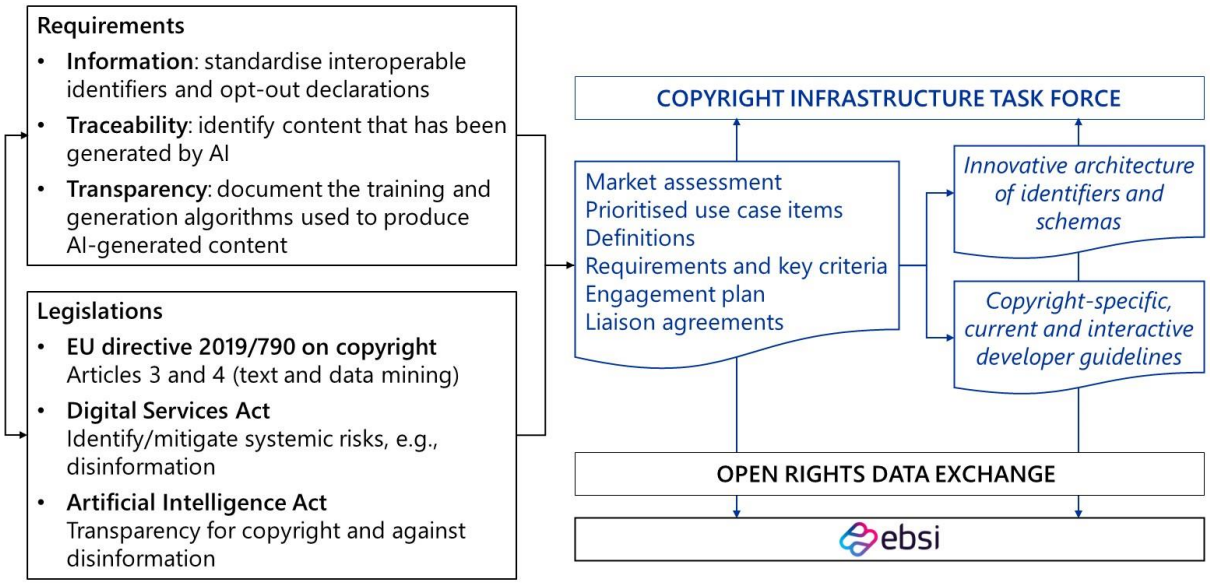

Figure 4: The AI & Copyright Use Case in relation with the CI Task Force and EBSI

The templates for the documents related to market assessment, prioritised use cases, definitions, requirements and key criteria, engagement plan, and liaison agreements are already available.

Requirements will be prioritised, and the development of a standardised opt-out application could be the first considered use case item.

The TRACE4EU project of Europeum EDIC is addressing the questions raised by the emergence of Large Language Models and the deployment of Generative AI applications through its application scenario "Open Rights Data Exchange" –

- Media asset: what is what, and who can tell,
- Parties: who is who, and who is accredited to claim work authorship or right ownership,
- Terms and conditions: what may one do with what, and how can an author or rightsholder opt-out of text and data mining, and
- Provenance and authenticity: where does that media asset come from.

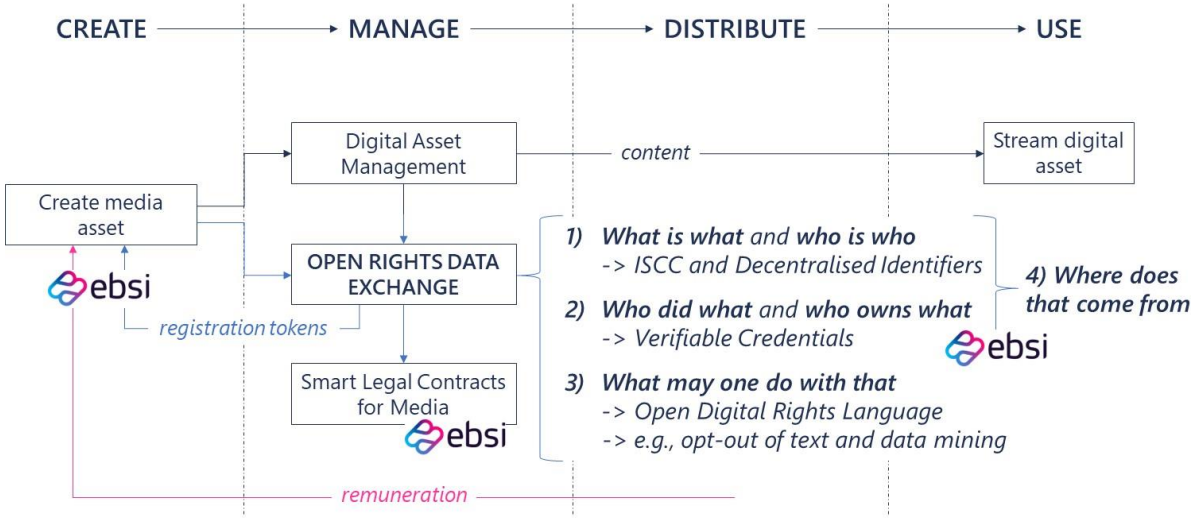

Figure 5: The questions answered by the Open Rights Data Exchange (EBSI)

For example, the opt-out declaration should be:

- machine-and-human-readable,
- based on open standards,
- inseparably bound to the content (i.e., resilient to content sharing),
- resilient to manipulation (i.e., resilient to the stripping of watermarks or metadata),
- able to provide verifiable attribution (e.g., using digital signatures or verifiable credentials),
- and timestamped.

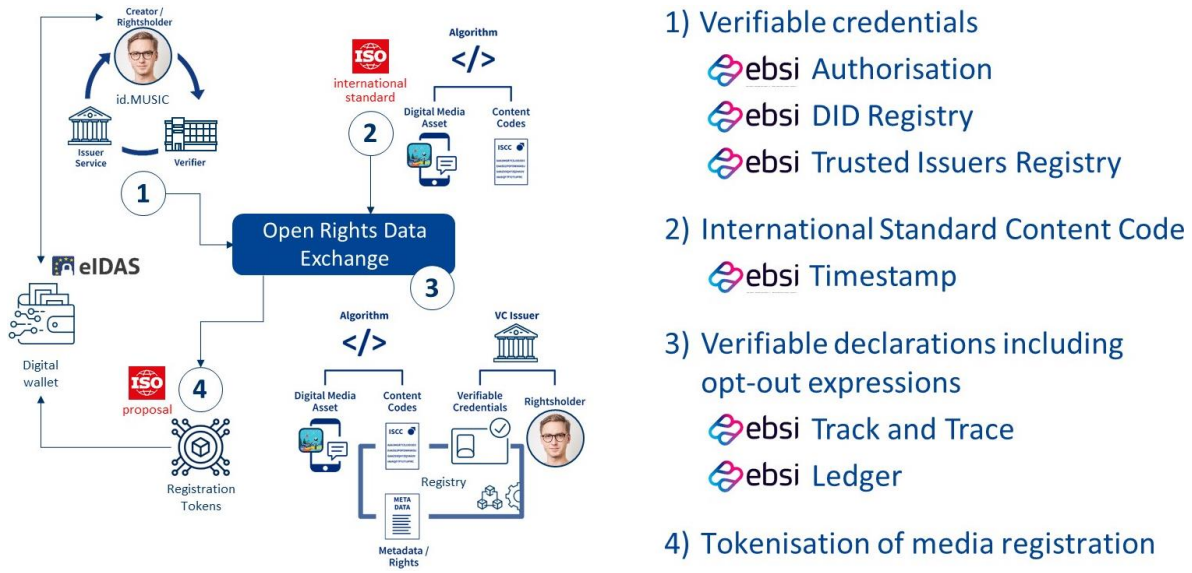

Figure 6: Leveraging EBSI APIs to store TDM opt-out expressions

The envisaged integration would rely on open standards – ISO standards and W3C recommendations.

The legal and technical requirements of the AI & Copyright Use Case could be completed by 31 January 2025, a mock-up of an approach and the design of a solution by 31 May 2025.

6.4. Timing

- Formal establishment of the CI Task Force: 1 October 2024
- Planning phase: 1 October 2024 – 31 May 2025
- Proposition to Europeum EDIC: 1 February 2025
- Operational phase: from 1 September 2025

6.5. Duration

Five years from the moment of full operational set up, i.e., the project will run until end of 2030 to deliver –

- An innovative architecture of identifiers and schemas,
- Current and interactive user references to technologies supporting the interoperability, searchability, and trustworthiness of rights management information,
- Dissemination, promotion, and adoption of the Open Rights Data Framework.

After that, the maintenance and governance of the Framework will need to be ensured to deal with the ever-changing content technologies, media business models, and related regulations. The Open Rights Data Framework will need a permanent home.

In the meantime, the close collaboration between the CI Task Force, the other task forces of Europeum EDIC, other EDICs, and other initiatives of the Digital Decade can lead to the deployment of a Common European **Copyright** Data Space whose sustainability and perennity would also need to be ensured.

7) Organisation and funding

7.1. Management of the CI Task Force

The CI Task Force is managed by –

- the Strategic Orientation Committee, formed by the Council working party for copyright
- the Advisory Boards, formed by the stakeholders form both standards organisations, organisations presenting right-holders, collective management organisations, large repositories and users.
- the Secretariat, formed by representatives from participating member states of the first project and Valunode OÜ. Work will be based on a clear charter of tasks.
- at least three Working Groups, which can be subdivided into task groups.

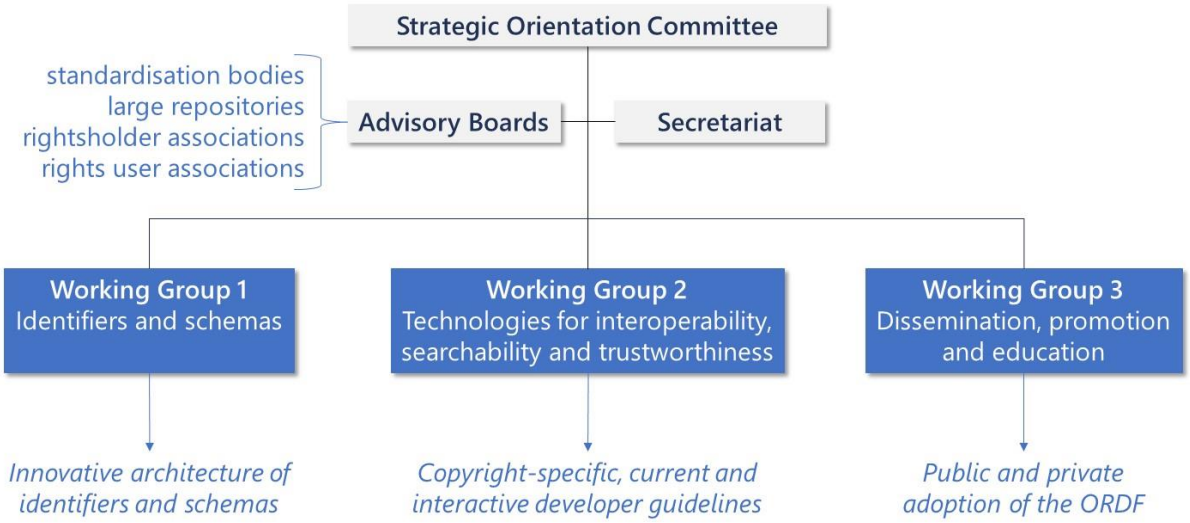

Figure 7: Organisation of the Copyright Infrastructure Task Force

7.2. Member States

The Member States –

a) appoint a representing entity (entities),

b) participate actively in the working groups of the CI Task Force,

c) support and promote the adoption of relevant standards,

d) promote uptake of the work and resources of the CI Task Force among relevant users and gather users' feedback,

e) facilitate the integration of the services of the CI Task Force,

f) provide the annual financial and in-kind contributions.

7.3. The Strategic Orientation Committee

The Strategic Orientation Committee could consist of the representatives of Member States participating in the Working Party on Intellectual Property (Copyright) at the Council of the European Union.

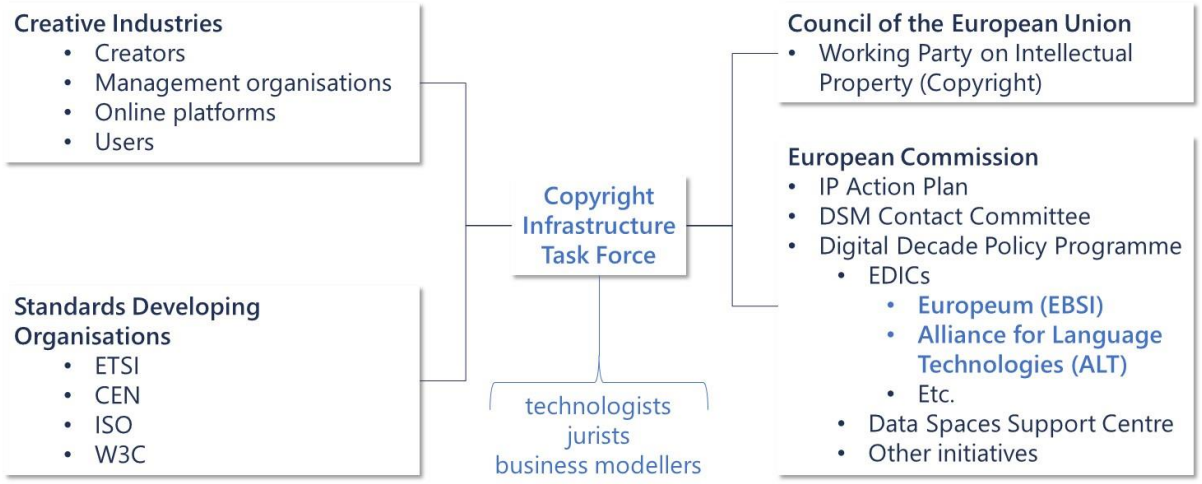

Figure 8: Liaisons

The Strategic Orientation Committee prepares –

- the Biennial strategic orientation for the work of the CI Task Force
- resolutions providing strategic or political advice for the implementation of the ORDF,

and present them to the Assembly of Members of CI Task Force and – potentially – to the Presidency of the Council.

Examples of necessary strategic and political advice will include –

- coordination between European Institutions and Member States,
- coordination within Member States, e.g., between various Ministries dealing with either rights, or data, or digitisation, or small and medium enterprises.

7.4. The Advisory Boards

There are four Advisory Boards –

- Standardisation, consisting of experts from standardisation bodies,
- Repositories, consisting of experts from organisations maintaining large repositories,
- Rights holding, consisting of experts from rightsholder associations,
- Rights usage, consisting of experts from rights user associations.

They **represent** stakeholders which are not Members and **answer** requests from the Assembly of Members, the Director, or the working groups. Among others, these experts advise on the selection and definition of use cases and on their practical implementations. They also advise on the governance of the CI Task Force.

7.5. The Secretariat

The Secretariat –

a) performs support functions of the CI Task Force on behalf of the Chair of the CITF informed by the Strategic Orientation Committee who is represented by the Copyright policy officers of the Member States of the European Union,

b) coordinates the activities of the advisory boards and working groups of the CI Task Force.

7.6. Necessary resources and sources of funding

Resources required for the operations

- Member States' experts working part-time in three groups –
  - Identifiers and schemas
  - Technologies for interoperability, searchability and trustworthiness
  - Dissemination, promotion and education,
- Delegates from international and industry standardisation bodies, organisations currently maintaining large repository of rights management information, associations of rightsholders and rights users contributing part-time to the working groups,
- Secretariat working part-time,
- Online collaboration platform,
- Legal, accounting, and auditing services,
- Promotion costs such as logo, website, graphics, collaterals, and
- Conferences.

First estimation of how the task force would be funded

The CI Task Force will be a not-for-profit organisation funded by –

- Seed funding from the Finnish Government (€0.167 million p.a. for 2 years 2024-25 and 2025-26)
- In-kind contributions of Member States focused on their country, e.g., personnel participation in the work of the CI Task Force and national promotion, or shared with the group, e.g., specific know-how contributions[22],
- In-kind contribution from the Secretariat covering their labour costs,
- Research and operation grants, e.g., from Horizon Europe,
- At a later stage: membership fees and proceeds from publications and conferences.

The planning phase will provide a strong basis for national and European in-kind and **financial** commitments for the year 2025-26 and beyond that will allow for the development and a secure operation of the CI Task Force covering among other personnel costs.

Note: ORDF vocabularies, ontologies, mappings, schemas, and protocols should be freely available for public use, except where proprietary schemas or protocols are developed for restricted use by a party or a group of parties.

8) Publicly available information

[1] European Commission, *An intellectual property action plan to support the EU's recovery and resilience*, COM(2020) 760

[2] Council of the European Union, *Developing the Copyright Infrastructure*, 15016/19

[3] European Commission, *Copyright and New Technologies*, SMART 2019/0038

[4] Decision of the European Parliament and of the Council establishing the 2030 Policy Programme "Path to the Digital Decade", COM(2021) 574 final, 2021/0293(COD)

[5] OpenStand principles, https://open-stand.org/

9) Contacts

Anna Vuopala, Senior Ministerial Adviser at the Finnish Ministry of Education and Culture
Email: anna.vuopala@gov.fi

Philippe Rixhon, Appointed Expert to the European Commission, Email: philippe@rixhon.net

[22] For examples: know-how from outreach experiences and the ISNI project in Finland, centralised copyright information system in Latvia, or anti-piracy and open academic publishing in Lithuania.