



**COPYRIGHT
INFRASTRUCTURE
TASK FORCE**

The Proposition

8 January 2024

Timeline

PUBLIC



Characteristics of a data framework

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 information. A successful data framework helps to streamline current processes and trigger innovative businesses. It is of potential benefit to everybody in the ecosystem.

Definition of the copyright infrastructure

The **copyright infrastructure** is the set of rules, technologies and institutions that **frame**¹ data management practices in the **creative industries** to improve authoritative and updated information on rightsholders, terms and conditions, and licensing opportunities.

An intellectual property action plan to support the EU's recovery and resilience, COM(2020) 760

¹: For the sake of clarity, this framework includes neither data nor systems.

Objectives of the Digital Decade Policy Programme

Objectives of the Copyright Infrastructure

- Inclusive, transparent, and open digital environment
- Interoperable digital technologies and services
- Secure and accessible digital and data infrastructures
- Digital regulatory environment
- Sustainable, resilient, and efficient digital infrastructures
- Coherent coordination of policies and programmes



The **Copyright Infrastructure Task Force** set up by the Governments of Estonia and Finland proposes to collaborate with the **EDIC Europeum** and other EU initiatives to support three areas of the **Digital Decade Policy Programme**:

- a) European common data infrastructure and services, in relation with
- g) Connected public administration, and
- h) European blockchain services infrastructure.

OPEN RIGHTS DATA EXCHANGE

technical layer providing mechanisms for trust and interoperation of identification and rights metadata



OPEN RIGHTS DATA FRAMEWORK

semantic layer addressing the challenges of declaration, discovery and conflict where the meaning of identification and rights metadata is made clear and consistent



ACQUIS COMMUNAUTAIRE

foundational legal layer defining authors' rights by conventions, treaties, directives, acts and regulations

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 set an **innovative data architecture able to integrate rights and content data of any kind or complexity.**

Set of technologies

The Copyright Infrastructure Task Force will promote the advances of technology programmes supporting the **interoperability**, **searchability**, and **trustworthiness** of rights management information.

These cutting-edge technologies are the base of the Digital Decade comprehensive framework: **blockchain**, **digital wallet**, **digital twins**, **high-performance computing**, and **5G**.

Set of institutions

Building a pervasive and inclusive ORDF at European and then global level requires **cooperation** and **coordination** between standards organisations, existing repositories, rightsholders and rights users.

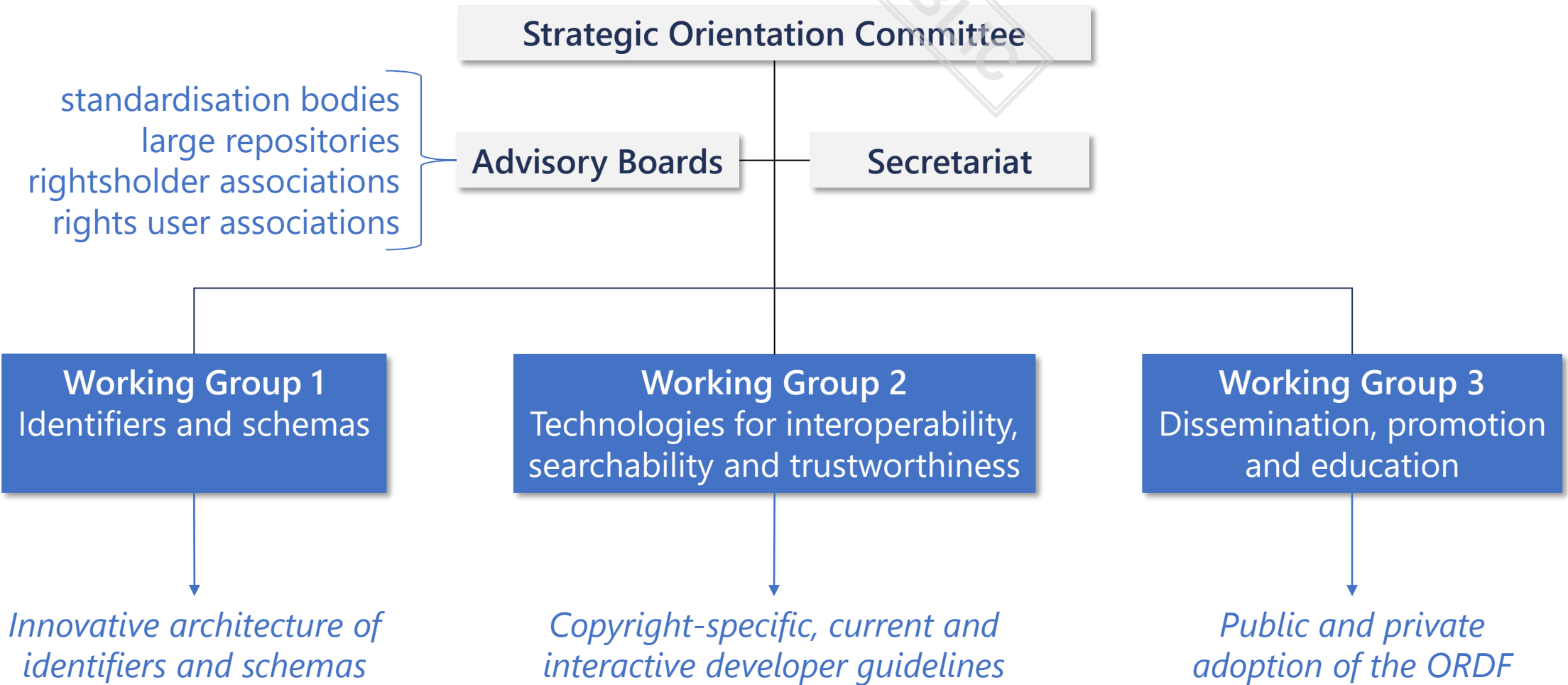
The CI Task Force does not publish standards itself but coordinate **requirements** and **resources** to foster the creation and evolution of standards within standards organisations. It assures the **governance** of the ORDF and liaise with the **European Data Innovation Board**.

The CI Task Force will define, promote, and govern an Open Rights Data Framework to facilitate the **fast, simple, and accurate management, licensing, enforcement, and remuneration of copyrights** and related rights.

For this purpose, the CI Task Force will carry out the following activities:

- a) define an **innovative architecture of identifiers and schemas** to integrate rights and content data of any kind or complexity,
- b) develop current and interactive user **references to technologies supporting interoperability, searchability, and trustworthiness of rights management information**, and
- c) disseminate and promote the results from activities (a) and (b) among **all stakeholders of the European creative industries**.

Organisation



Liaisons

Creative Industries

- Creators
- Management organisations
- Online platforms
- Users

Standards Developing Organisations

- ETSI
- CEN
- ISO
- W3C

Copyright Infrastructure Task Force

technologists
jurists
business modellers

Council of the European Union

- Working Party on Intellectual Property (Copyright)

European Commission

- IP Action Plan
- DSM Contact Committee
- Digital Decade Policy Programme
 - EDICs
 - **Europeum (EBSI)**
 - **Alliance for Language Technologies (ALT)**
 - Etc.
 - Data Spaces Support Centre
 - Other initiatives

- 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 the **EDIC Europeum**.

Article 19 of the EDIC Europeum Statutes: The EBSI Use Case Groups

1. The Assembly of members may set up EBSI Use Case Groups to drive the governance of EBSI use cases. The EBSI Use Case Groups shall follow the mandate given to them by the Assembly of members.
2. An EBSI Use Case Group shall be convened by one or more persons appointed by the Assembly of members.
3. The EBSI Use Case Groups shall:
 - a) Support the deployment of a use case or a family of use cases on EBSI;
 - b) Work on the technical specifications and first implementation of a specific EBSI use case, or ensure that solutions are provided by service providers or technical providers;
 - c) Facilitate the adoption of a use case.
 - d) Coordinate their work with other use case groups and other EDIC Europeum governance groups.
4. The Assembly of members may change the mandate of the EBSI Use Case Groups and may decide when an EBSI Use Case Group has completed its mandate, and if any EBSI Use Case Groups should be merged or dissolved.
5. Implementing Rules may indicate further details of the mission and functioning of the EBSI Use Case Groups.

Member States can join the EDIC Europeum and contribute to the working groups of the Copyright Infrastructure Task Force.

As **Members** or **Observers** of the EDIC Europeum, they –

- a) **appoint** a representing entity (entities),
- b) **participate** 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) **contribute** in-kind to the budget of the EDIC,
- g) **attend** the Assembly of Members of the EDIC.

As **Members** of the EDIC Europeum, they also –

- a) **contribute** financially to the budget of the EDIC,
- b) **participate** in and **vote** at the Assembly of Members of the EDIC.

Copyright Infrastructure Task Force and EDIC Europeum, current approach

The Copyright Infrastructure Task Force will be officialised –

- **Host:** an existing Finnish entity
- **Strategic Orientation Committee:** The Working Party on Copyright (Council of the EU)
- **2024 budget:** €167,000.00 funded by the Finnish Government for copyright work

In 2024, the Copyright Infrastructure Task Force will –

- Complete its **planning** by 20 December 2024,
- Launch the **AI & Copyright Use Case**, completing the legal and technical requirements by 28 June 2024, and the design by 20 December 2024, and
- Join forces with the **EDIC Europeum** by collaborating on the AI & Copyright Use Case, which will become a **Use Case Group** under conditions that remain to be defined.

Note – Member States and/or their Ministries can contribute to the EDIC Europeum as members or observers.

Let us plan together (for discussion only)



5 February	Next meeting of the Working Party on Copyright
15 February	Founding General Assembly of EDIC Europeum
29 March	Setting of the Copyright Infrastructure Task Force AI & Copyright Use Case – Initial scope (creative sector and Member States) Use Case Group proposal to EDIC Europeum
8-9 April	Conference of the Belgian Presidency on AI and DSM
28 June	AI & Copyright Use Case – Legal and technical requirements
20 December	Completion of the planning phase AI & Copyright Use Case – Design



The AI & Copyright Use Case

8 January 2024

AI & COPYRIGHT USE CASE

Requirements

- **Information:** standardise interoperable identifiers and opt-out declarations
- **Traceability:** identify content that has been generated by AI
- **Transparency:** document the training and generation algorithms used to produce AI-generated content

Regulations

- **EU directive 2019/790 on copyright**
Articles 3 and 4 (text and data mining)
- **Artificial Intelligence Act**
- **Etc.**

COPYRIGHT INFRASTRUCTURE TASK FORCE

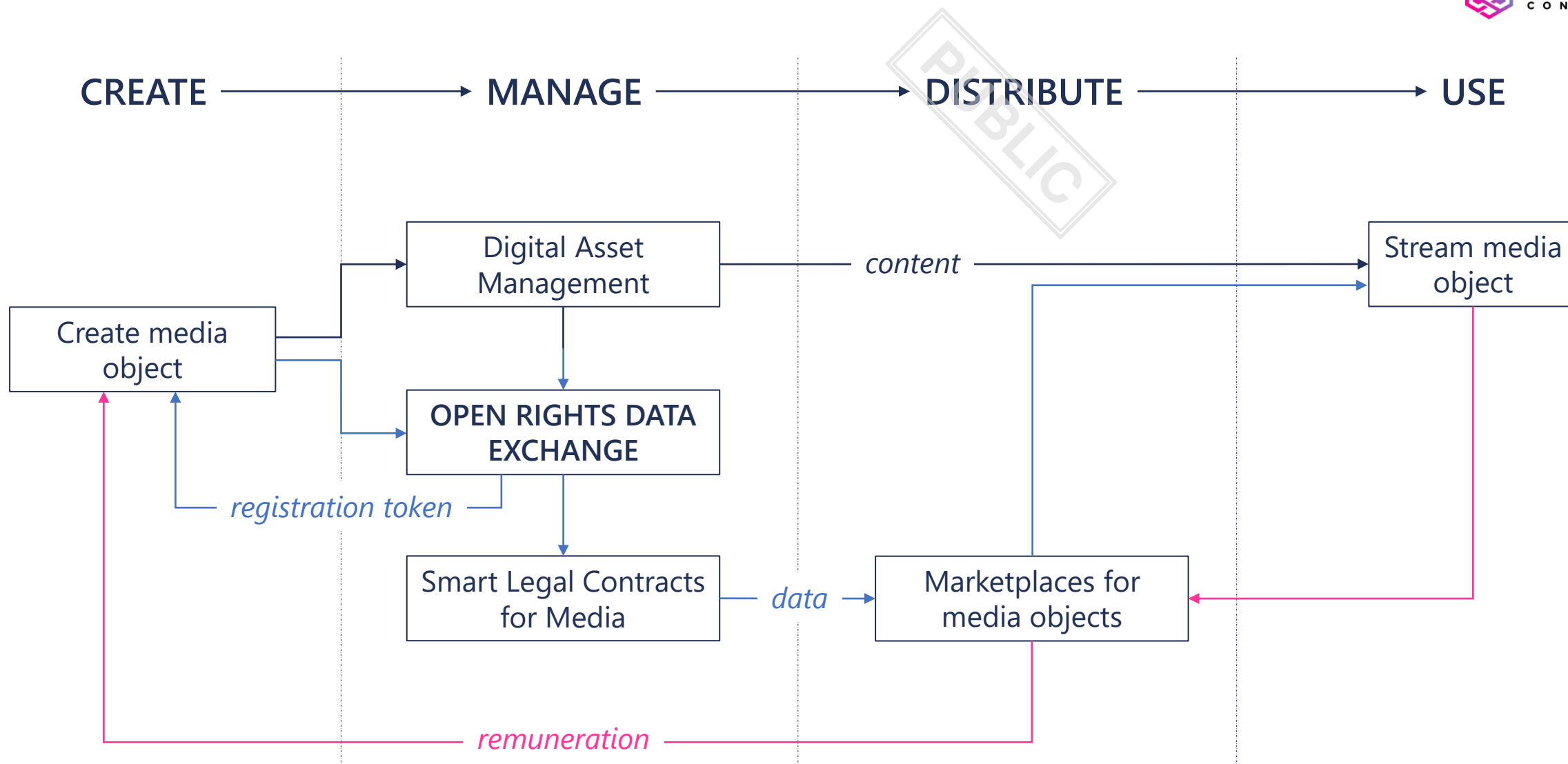
Market assessment
Prioritised use case items
Definitions
Requirements and key criteria
Engagement plan
Liaison agreements

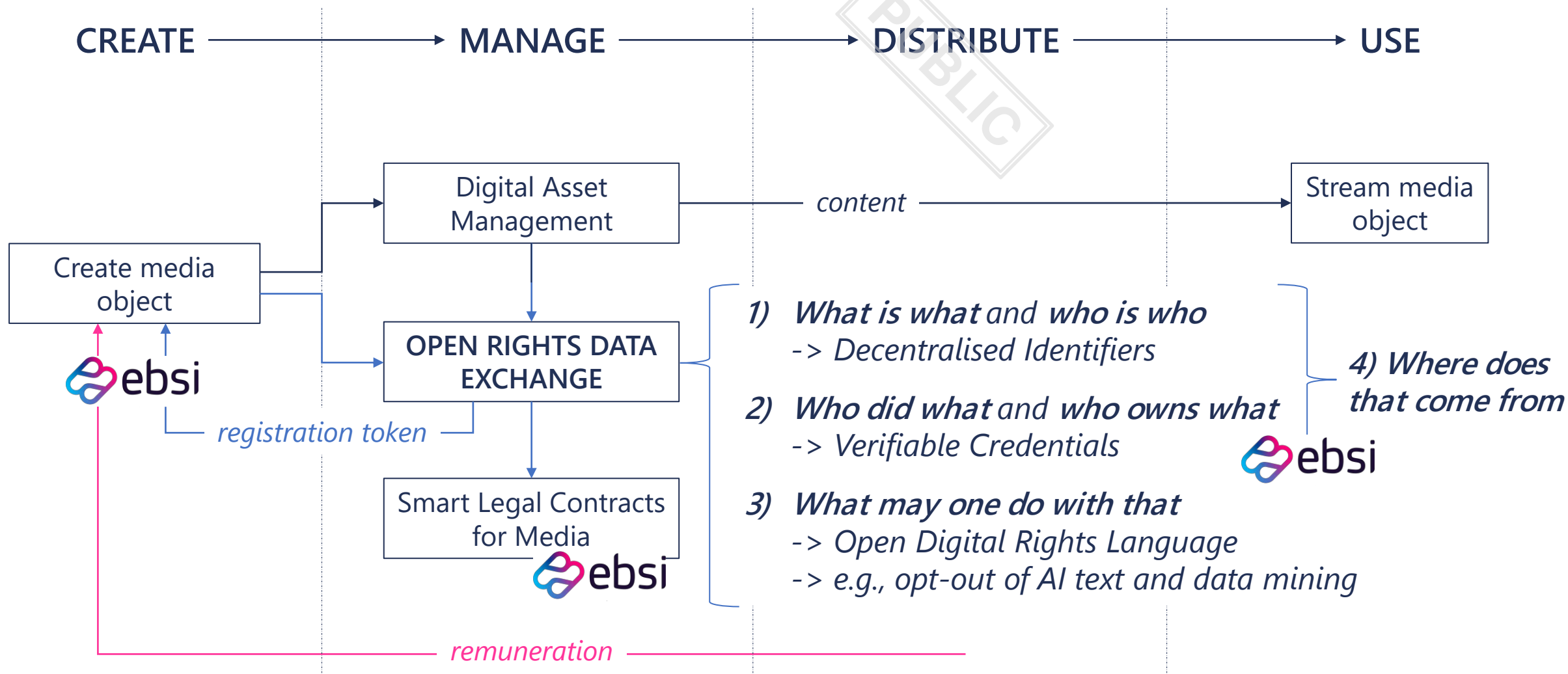
*Innovative architecture
of identifiers and
schemas*

*Copyright-specific,
current and interactive
developer guidelines*

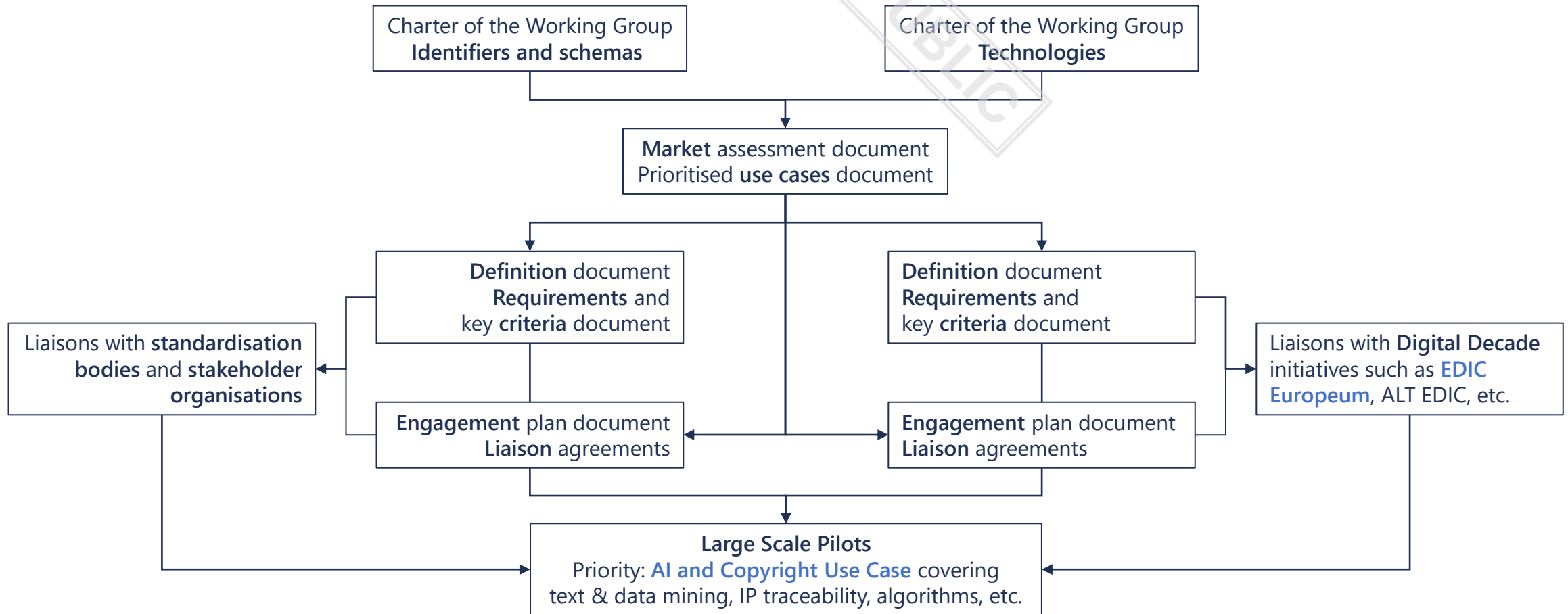
OPEN RIGHTS DATA EXCHANGE





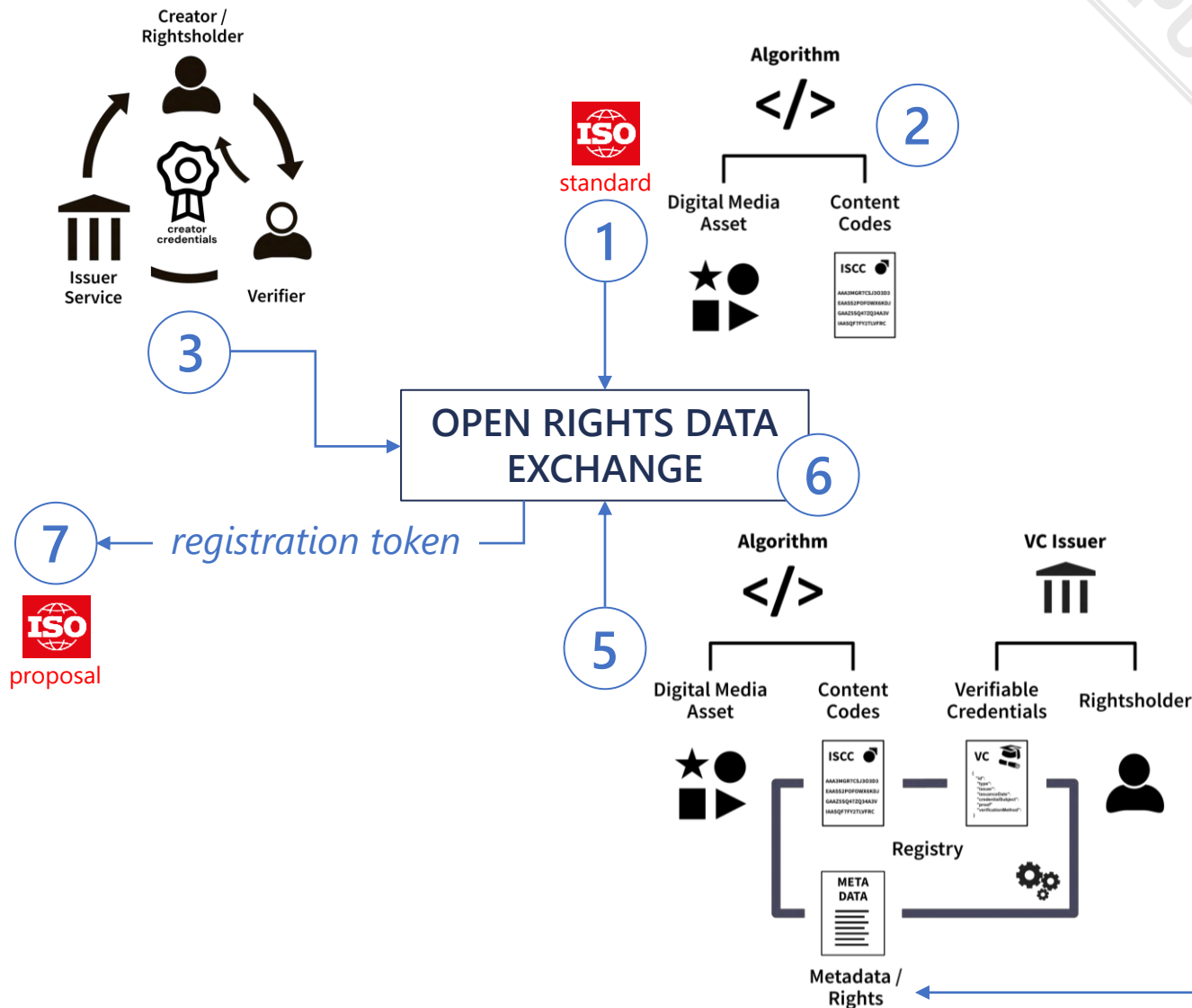


Plan



The Copyright Infrastructure Task Force is adopting a methodology based on the [OpenStand principles](#) and convening a team of experts to address the AI and Copyright Use Case.

European technologies



- 1) International Standard Content Code
- 2) Liccium App
- 3) Creator Credentials
- 4) Opt-Out Policy, e.g., TDM·AI Protocol
- 5) Verifiable Declaration
- 6) Open Rights Data Exchange 
- 7) Tokenisation of Media Registration

"Content **must not** be used for AI training purposes." (TDM·AI protocol, according to the Directive EU/2019/790 on Copyright, Article 4)

A comparison of opt-out protocols for discussion

	TDM-AI	C2PA	IPTC	TDMrep	Robots.txt	Comments
Schema definition	✓	✓	✓	✓	✓	Suggested or specified schema
Web-published content	✓	✓	✓	✓	✓	Can be used if content is published on websites
Opt-out embedded in asset metadata	✓	✓ (Selected file formats*)	✓ (Visual content)	(potentially EPUB, PDF)	✗	Opt-out is bound to the asset
Certificate embedded in asset metadata	✓	✓	✗	✗	✗	Verification of declarer's identity embedded within metadata.
Opt-out resilient when content is shared, metadata or watermarks are removed	✓	✗	✗	✗	✗	Opt-out persists in case of social sharing and removal of metadata or watermarks
Certificate resilient when content is shared, metadata or watermarks are removed	✓	✗	✗	✗	✗	Attribution remains variable regardless of social sharing and removal of metadata or watermarks
Verifiable timestamp	✓	✗	✗	✗	✗	Declaration includes a verifiable timestamp
Opt-out resilient to content manipulation	✓	✗	✓	✓	✓	Opt-out remains accessible when content is altered or modified
Certificate resilient to content manipulation	✓	✗	✗	✗	✗	Verifiable attribution certificate remains accessible when content is altered or modified
Comments	In development, based on upcoming ISO standard ISO/DIS 24138	New industry standard proposal	Industry standard	W3C Community Group Report	Global standard (RFC 9309)	

* <https://opensource.contentauthenticity.org/docs/c2patool/#supported-file-formats>

PUBLIC

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



Council of the European Union
General Secretariat

Brussels, 30 January 2024

WK 1252/2024 INIT

LIMITE

PI

This is a paper intended for a specific community of recipients. Handling and further distribution are under the sole responsibility of community members.

INFORMATION

From:	the Finnish delegation
To:	Working Party on Intellectual Property (Copyright)
N° prev. doc.:	WK 15993/2023
Subject:	Copyright Infrastructure Task Force – Information from the Finnish delegation

Delegations will find attached the draft text of a proposition developed by the Copyright Infrastructure Task Force that consists of an appointed expert to the European Commission and FI, EE, LV and LT as members and 7 other Member States as observers.

The draft text of this proposition is accompanied by a PPT presentation (also attached) that will be presented by the Finnish delegation as a basis for discussions under AOB (CM 1391/24) of the Working Party on Intellectual Property (Copyright) on 5 February 2024.



The Proposition

This document presents the proposition of the Copyright Infrastructure Task Force instigated by the Governments of Finland and Estonia and is accompanied by the slide deck *Copyright Infrastructure Task Force – Proposition – Slides – 20240108 PR v17.pdf*.

Table of Contents

- 1) The Copyright Infrastructure Task Force (CI Task Force)..... 3
- 2) Corresponding area of activity in the annex of the DDPP decision..... 4
- 3) Rationale and objectives of the Multi-Country Project 5
- 4) Link to the DDPP general objectives and digital targets..... 6
 - 4.1) General objectives 6
 - 4.2) Digital targets..... 8
- 5) Detailed description of the Multi-Country Project..... 8
 - 5.1) Set of rules..... 9
 - 5.2) Set of technologies..... 10
 - 5.3) Set of institutions..... 10

6)	Main tasks, activities, and deliverables of the CI Task Force.....	11
6.1)	Definition	11
6.2)	Promotion.....	11
6.3)	Governance.....	12
6.4)	Deliverables	12
7)	Implementation strategy, including implementation modalities.....	13
7.1)	Strategy.....	13
7.2)	Planning phase (2024).....	16
7.3)	Reaching out activities	17
7.4)	Modalities	18
8)	Tasks, roles, and responsibilities.....	19
8.1)	Management of the CI Task Force	19
8.2)	Member States.....	20
8.3)	The Strategic Orientation Committee.....	20
8.4)	The Advisory Boards.....	21
8.5)	The Secretariat	21
9)	Indicative timeline, milestones, and targets.....	22
9.1)	Possible timing.....	22
9.2)	Project duration	22
10)	The AI & Copyright Use Case.....	23
11)	Indicative budget and sources of funding	27
11.1)	Resources required for the operations of the CI Task Force	27
11.2)	First estimation of how the project would be funded	27
11.3)	2024 budget.....	28
11.4)	First estimation of the 2025 budget.....	29
12)	Publicly available information.....	29
13)	Contacts	29

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

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 been established by the Governments of Estonia and Finland in September 2022 to explore the set up a Multi-Country Project (MCP) around the Open Rights Data Framework (ORDF) suggested in the above-mentioned study. In September 2023, the Commission recommended the CI Task Force to work closely with the EDIC Europeum in formation.

- 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 the **EDIC Europeum**.

Figure 1: Progress of the CI Task Force

2) 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 administration¹, and
- (h) European blockchain services infrastructure (EBSI)²

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

(g) 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.³

¹ 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.

² The Multi-Country Project considers indeed the EBSI as a key underlying technology.

³ 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 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. See short presentation in English. See https://dvv.fi/-/uusi-kasikirja-tukee-tietojen-liikkumista-maiden-valilla-tavoitteena-sujuvampi-kansainvalinen-arki?language=en_US

(h) The CI Task Force will 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) 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 industries⁴ – 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,

⁴ 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.

- 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.

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**.

4) Link to the DDPP general objectives and digital targets

4.1) General objectives

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

- | | |
|---|---|
| <p>(a) <i>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;</i></p> | <p>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.</p> |
| <p>(c) <i>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, ...</i></p> | <p>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.</p> |

- (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 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.

4.2) 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;*

(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;*

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**.

(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**;*

(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 enable the open **access to and online interaction with public services** related to rights data, see footnote 1 on page 19.

The ORDF will make **secure electronic identification** a prerequisite to user and rights-holder authentication and put them in **control of their data**.

5) 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.**

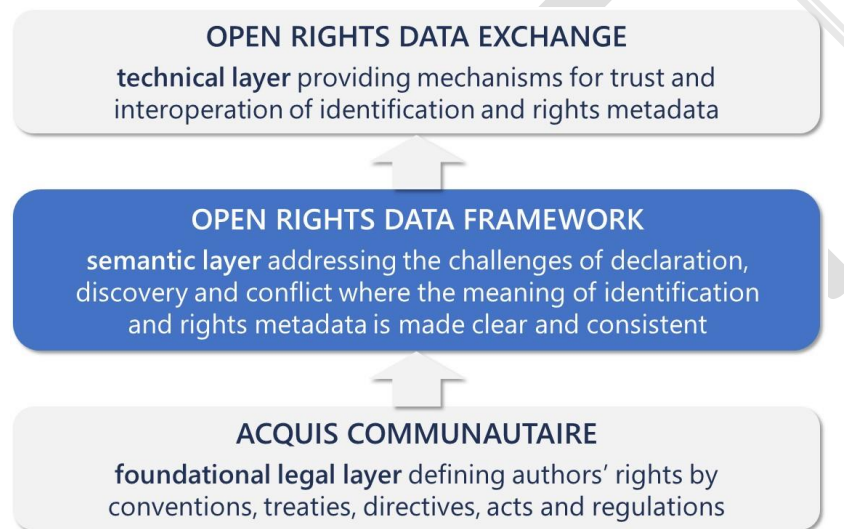


Figure 2: Position of the Open Rights Data Framework

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.

5.1) 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.

5.2) Set of technologies

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

As mentioned here above, the CI Task Force will leverage the convergent advances of technology programmes such as EBSI, eIDAS 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).

5.3) 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 offices⁶, 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.

6) Main tasks, activities, and deliverables of the CI Task Force

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

6.1) 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.

6.2) 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.

6.3) 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.

6.4) 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.

7) Implementation strategy, including implementation modalities

Within six months⁷ from the entry into force of the Commission decision establishing the EDIC Europeum, the Secretariat of the CI Task Force shall propose to their Assembly of Members the establishment of a specific Europeum Use Case Group to implement the strategy of the CI Task Force⁸.

7.1) Strategy

Objectives

The CI Task Force shall implement an ORDF, i.e., the semantic layer of a copyright infrastructure⁹ 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 complexity¹⁰,
- 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 Preferably before the meeting organised by the Belgian Presidency of the EU in Namur/Belgium on 8-9 April 2024.

⁸ We will need to have one person leading the CI Task Force and another person leading the Use Case Group.

⁹ The set of rules, technologies and institutions that frame data management practices in the creative industries.

¹⁰ 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.

Exemplary use cases

News publishers – European press articles are consolidated on American online platforms and used to train 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.

Relationships with the EDIC Europeum, Alliance for Language Technologies (ALT EDIC), eIDAS, 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 the EDIC Europeum, ALT EDIC, eIDAS, 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, the other task forces of the EDIC Europeum, other 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 consortiums¹¹ conducting **Large-Scale Pilots** (LSPs) under current data regulations¹² 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.

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, and
- 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,

¹¹ Like the consortiums EUDI Wallet Consortium (EWC), NOBID, Potential, and DC4EU carrying Large Scale Pilots for eIDAS.

¹² 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.

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.

7.2) Planning phase (2024)

Agenda

- 1) Formulate detailed objectives (e.g., interoperability and searchability) and values (e.g., trust, security, and inclusiveness) aligned with the Acquis Communautaire, Digital Decade Path, and IP Action Plan,
- 2) Define use cases, priorities, and scenarios (aligned with the Acquis Communautaire, Digital Decade Path, IP Action Plan, and the emergence of disruptive technologies and business models), whereby **use cases** will be detailed specific sectoral or cross sectoral, national or European data challenges, **priorities**¹³ will be justified by necessities to intervene, and **scenarios** will be discussed with stakeholders,
- 3) Identify and map the relevant standardisation bodies, existing repositories of rights management information, and associations of rightsholders and rights users, and reach out to them,
- 4) Develop the requirements (aligned with the Acquis Communautaire, Digital Decade Path, and IP Action Plan),
- 5) Set up the identity (logo, etc.) of the CI Task Force and deploy its website,
- 6) Identify relevant grants and submit applications.

Approach

- Convene the first CI Task Force Strategic Orientation Committee consisting of the Member States delegates to the Working Party on Intellectual Property (Copyright) at the Council of the European Union,
- Put the CI Task Force on the agenda of the upcoming Presidencies of the Council: Spain, Belgium, and Hungary,

¹³ The CI Task Force will choose which creative sector and which Member States come first. See the draft of a study that contains the laws on text and data mining and opt-out mechanisms in Austria, France and Germany.

- Meet twice in person and discuss copyright, once on 8-9 April 2024 in Namur, Belgium¹⁴, and once in the second semester (end of September / early October), whereby these meetings will take place just before or just after scheduled meetings of the Council Working Party on Intellectual Property (Copyright) and in the same city,
- Meet twice in person and discuss data and digital transformation, whereby these meetings will take place just before or just after meetings related to the Digital Decade organised by the Belgian and Hungarian presidencies and could be organised with the help of EDIC Europeum.

Expected outcomes

- Detailed plan, objectives, use cases, and requirements,
- Strong basis for national and European in-kind and financial commitments, and
- First draft of an eventual recommendation.

7.3) 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 play an active role in the upcoming EDIC Europeum.

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 organisations¹⁵.

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 the EDIC Europeum 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.

¹⁴ The CI Task Force will contact the organiser of the conference.

¹⁵ 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.

7.4) Modalities

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 process¹⁶, 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.

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 Consortia (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.

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.

¹⁶ Including votes.

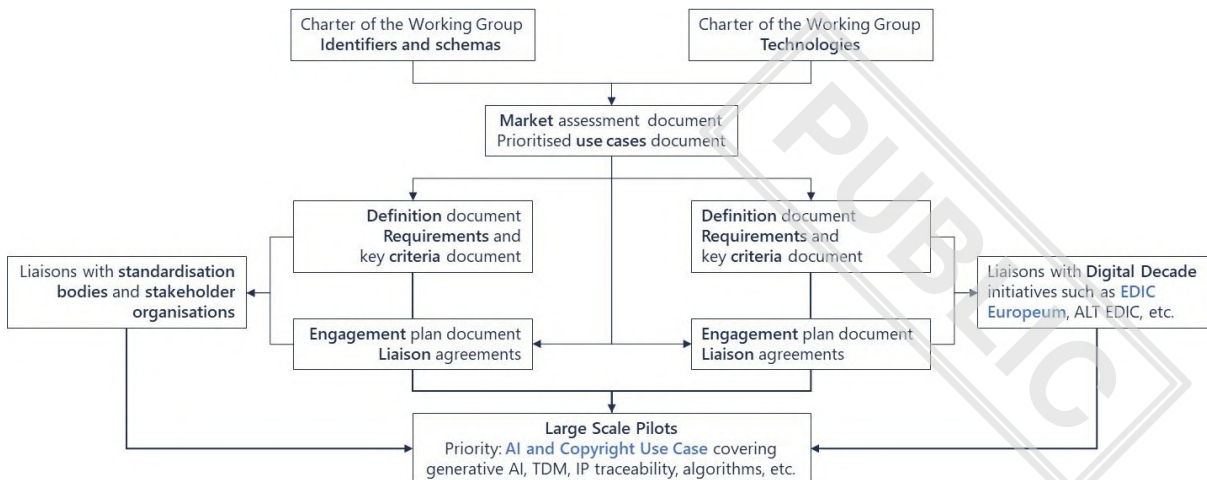


Figure 3: Setting up a Large-Scale Pilot

8) Tasks, roles, and responsibilities

8.1) Management of the CI Task Force

The CI Task Force is managed by –

- the Strategic Orientation Committee,
- the Advisory Boards,
- the Secretariat,
- at least three Working Groups, which can be subdivided into task groups.



Figure 4: Organisation of the Copyright Infrastructure Task Force

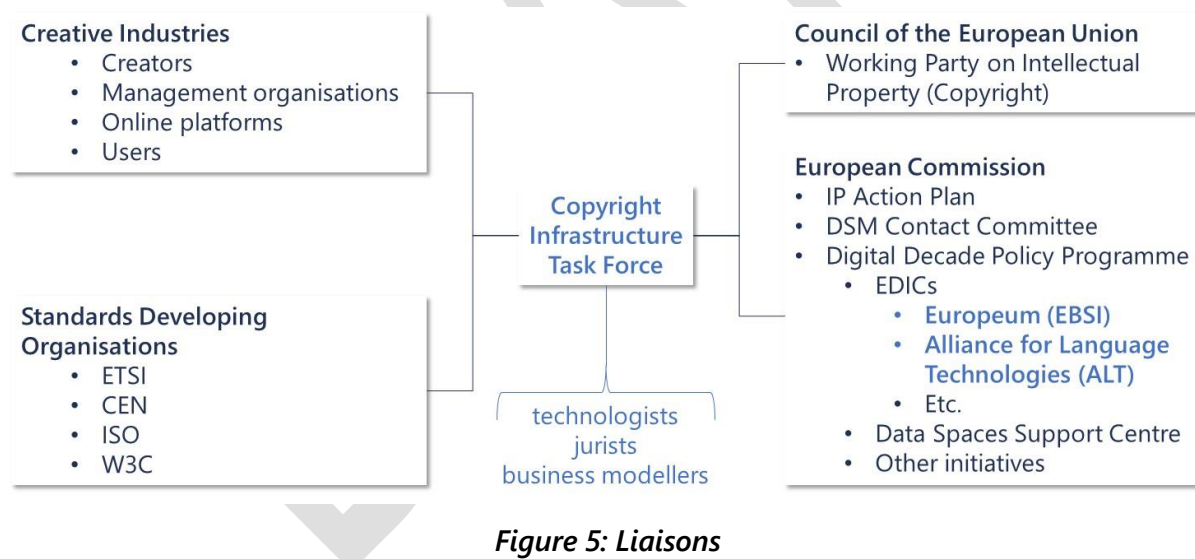
8.2) Member States

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.

8.3) 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.



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 the EDIC Europeum, 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.

8.4) 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.

8.5) The Secretariat

The Secretariat –

- a) performs support functions of the CI Task Force on behalf of the Chair of the Strategic Orientation Committee who represents the CI Task Force,
- b) coordinates the activities of the advisory boards and working groups of the CI Task Force.

9) Indicative timeline, milestones, and targets

9.1) Possible timing

- Proposition to the EDIC Europeum: 1 November 2023
- Establishment: 1 January 2024
- Planning phase: 1 January – 31 December 2024
- Operational phase: from 1 January 2025

9.2) Project duration

Five years from the moment of full operational set up, i.e., the project will run until end of 2029 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 the EDIC Europeum, 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.

10) The AI & Copyright Use Case

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 EBSI work, 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 suggest leveraging an AI & Copyright Use Case to –

- launch its operations in 2024,
- conduct a Large-Scale Pilot, and
- work closely with the EDIC Europeum while becoming one of its Use Case Group.

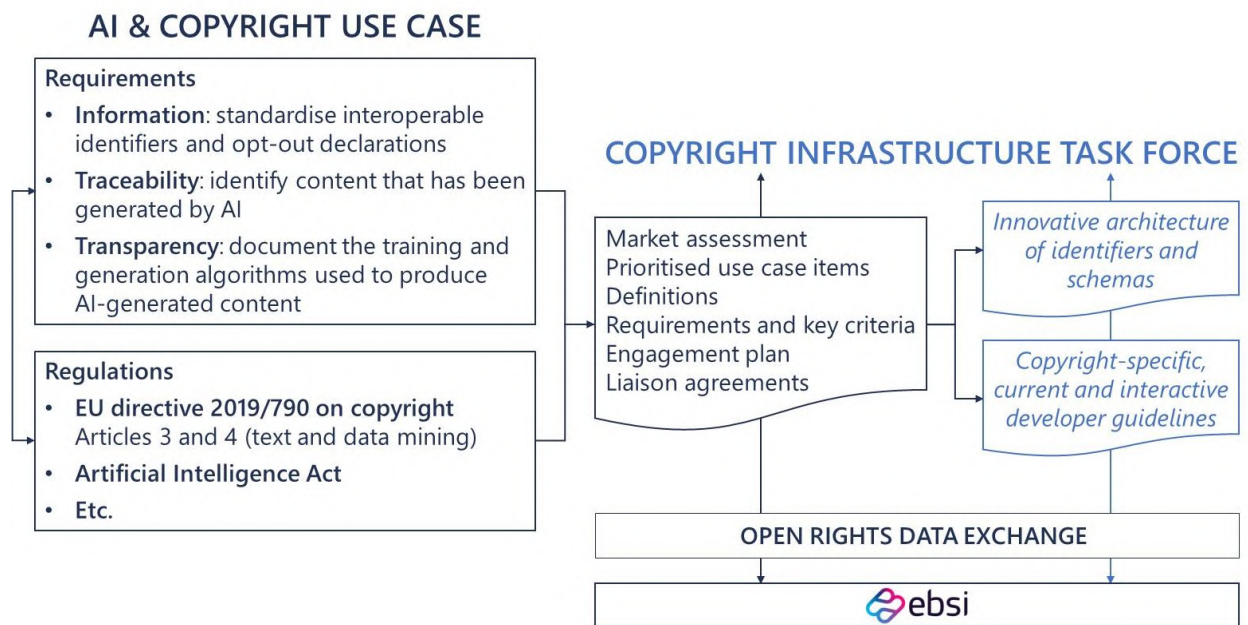


Figure 6: The AI & Copyright Use Case in relation with the CI Task Force and EBSI

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 the EDIC Europeum 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”:

- What is what, and who can tell,
- Who is who, and who is accredited to claim work authorship or right ownership,
- What may one do with what, and how can an author or rightsholder opt-out of text and data mining, and
- Where does that content come from.

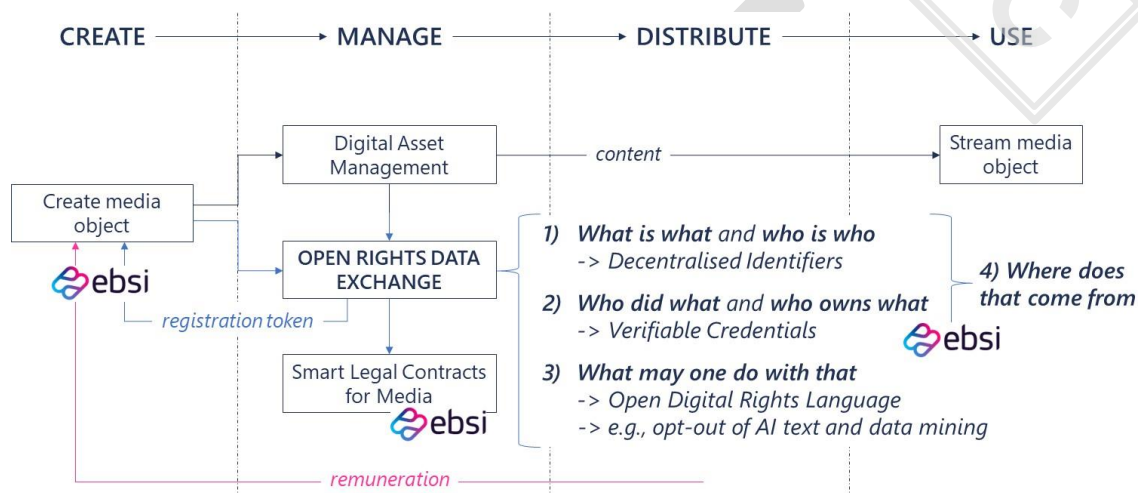


Figure 7: The questions answered by the Open Rights Data Exchange

For example, the opt-out declaration requires to 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.

Besides TRACE4EU, the European Commission [Next Generation Internet initiative](#) (NGI) is co-funding [NGI Trublo](#) and its project [Liccium App](#), which helps creators to inseparably bind public and verifiable claims to their content, and [NGI TrustChain](#) and its project [Creator Credentials](#), that can help identify creators and rightsholders and create trust in claims and attributions. Integrating the NGI prototypes with the TRACE4EU development seems to be a promising starting point to technically develop a standardised opt-out mechanism.

Training of AI with copyright protected content – Provision on text and data mining (TDM) including opt-out mechanism

Articles 3 and 4 of the DSM directive provide for an exception to the exclusive right of the author of a work and other subject matter for purposes of text and data mining (TDM). TDM is allowed for scientific research by research organisations and cultural heritage institutions and for any other purpose including commercial purpose by any user, subject to the so called “three step test”¹⁷. Other uses than those for scientific research based on art 3 may be withdrawn through mechanism i.e., the opt-out. According to Article 4 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. Especially since the flush of generative AI services in the digital single market it is clear that rightsholders feel the need to reserve their rights to prohibit the application of the exception.

Following the provisional outcome of the Artificial Intelligence Act in December 2023, requiring general purpose AI systems and AI models they are based on to comply with EU Copyright legislation, the opt-out mechanism under Article 4 needs a standard to be realised. However, considering that most AI models are developed outside European Union, the opt-out scheme is new and therefore unfamiliar in a global context. There are no harmonised rules on TDM on international level and the issue on the use of works for training of AI is open. Alternatives are either 1) no exception at all which means that the main rule on the need for consent applies when works are reproduced for training purposes or 2) an exception exists but either solely for scientific research, or also for other (commercial) uses but without an opt-out mechanism. The policies related to training of AI are not yet fixed in all jurisdictions. Several court cases are pending that could clarify interpretation of copyright law. In the European Union provisions may vary from one Member State to the next regardless of the aforementioned directive. Specifying the details of the use case (piloting the ORDF), legal experts of the CI Task Force and the Commission will analyse the key issues and recommend wordings in early 2024. These wordings are needed to be subsequently implemented by engineers. The progress is reported at the Council Working Party on Copyright during the Belgian Presidency. To be efficient and complete the ORDF needs to cater for various models in a global context.

¹⁷ According to Art 5 of the Directive on Copyright in the Information Society (2001/29/EC) exceptions and limitations can “only be applied in certain special cases which do not conflict with a normal exploitation of the work or other subject-matter and do not unreasonably prejudice the legitimate interests of the rightsholder”.

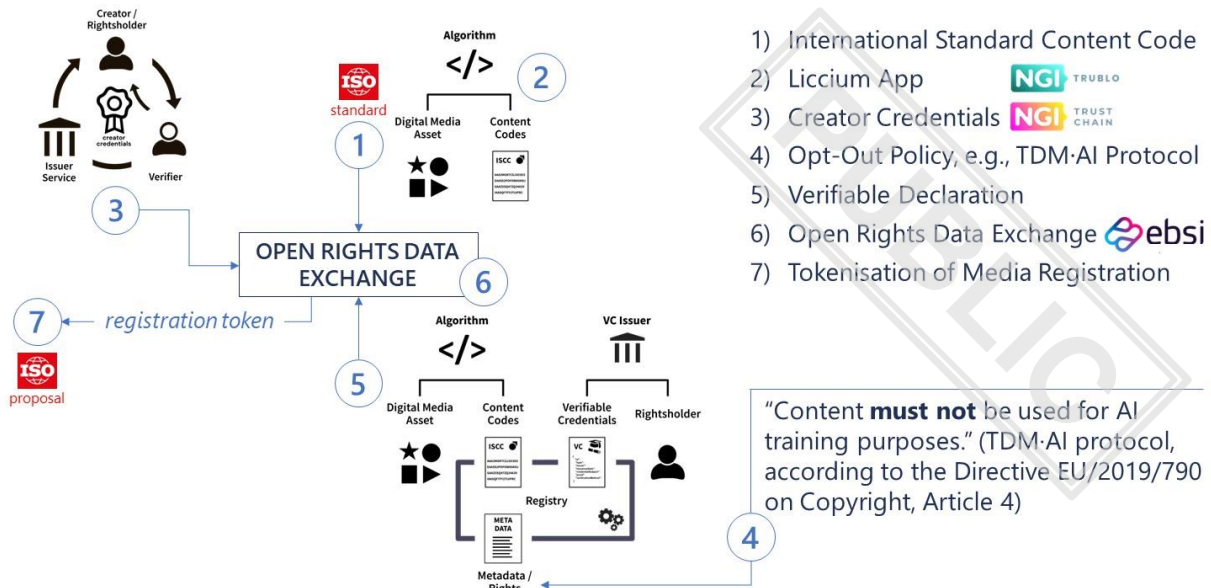


Figure 8: Integration of European prototypes to build a standard opt-out application

The envisaged integration would rely on open standards – ISO standards and W3C recommendations. Not only does the European TDM-AI protocol compare favourably with other related initiatives, but it can even be combined with them.

	TDM-AI	C2PA	IPTC	TDMrep	Robots.txt	Comments
Schema definition	✓	✓	✓	✓	✓	Suggested or specified schema
Web-published content	✓	✓	✓	✓	✓	Can be used if content is published on websites
Opt-out embedded in asset metadata	✓	✓ (Selected file formats*)	✓ (Visual content)	(potentially EPUB, PDF)	✗	Opt-out is bound to the asset
Certificate embedded in asset metadata	✓	✓	✗	✗	✗	Verification of declarer's identity embedded within metadata.
Opt-out resilient when content is shared, metadata or watermarks are removed	✓	✗	✗	✗	✗	Opt-out persists in case of social sharing and removal of metadata or watermarks
Certificate resilient when content is shared, metadata or watermarks are removed	✓	✗	✗	✗	✗	Attribution remains variable regardless of social sharing and removal of metadata or watermarks
Verifiable timestamp	✓	✗	✗	✗	✗	Declaration includes a verifiable timestamp
Opt-out resilient to content manipulation	✓	✗	✓	✓	✓	Opt-out remains accessible when content is altered or modified
Certificate resilient to content manipulation	✓	✗	✗	✗	✗	Verifiable attribution certificate remains accessible when content is altered or modified
Comments	In development, based on upcoming ISO standard ISO/DIS 24138	New industry standard proposal	Industry standard	W3C Community Group Report	Global standard (RFC 9309)	

Figure 9: A comparison of opt-out protocols for discussion

The legal and technical requirements of the AI & Copyright Use Case could be completed by 30 June 2024, and the design of a solution by 31 December 2024.

11) Indicative budget and sources of funding

11.1) Resources required for the operations of the CI Task Force

- 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.

11.2) First estimation of how the project 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 2024-2025)
- 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 contributions¹⁸,
- 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 years 2025 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.

¹⁸ 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.

11.3) 2024 budget

Core team:

- Estonia: at least 1 "copyright" officer + 1 "data/digitisation" officer
- Finland: at least CI Task Force initiator + 1 "copyright" officer + 1 "data/digitisation" officer
- Latvia: at least 1 "copyright" officer + 1 "data/digitisation" officer
- Lithuania: at least 1 "copyright" officer + 1 "data/digitisation" officer
- Secretariat: 1 officer

i.e., ~ 10 people with a workload of ~2 days/month.

Labour costs of all other members of the core team are carried by their organisation as in-kind contributions.

Rent and office costs:	€ 5,000.00
------------------------	------------

Establishment of the CI Task Force:	€ 5,000.00
-------------------------------------	------------

Legal, accounting, and auditing services:	€ 7,800.00
---	------------

Travel costs:

- | | |
|--|-------------|
| • 4 planned trips @ 10 to 11 people @ €800.00: | € 33,600.00 |
| • 8 planned trips @ 1 to 2 people @ €800.00: | € 11,200.00 |

Promotion costs:

- | | |
|----------------------------------|-------------|
| • Logo and corporate identity: | € 5,000.00 |
| • Website, incl. members' space: | € 10,000.00 |
| • Collaterals: | € 5,000.00 |

Reserve, e.g., for the AI & Copyright Use Case:	€ 84,400.00
---	-------------

Total:	€ 167,000.00
---------------	---------------------

11.4) First estimation of the 2025 budget

Summary:

• Rent and office costs:	€ 11,000.00
• Legal, accounting, and auditing services:	€ 15,600.00
• Promotion costs:	€ 23,800.00
• Travel expenses:	€ 89,600.00
• Reserve:	€ 27,000.00

• Total:	€ 167,000.00

The first estimations of the 2024 and 2025 budgets show that the seed financial contribution from Finland and in-kind contributions from Estonia, Finland, Latvia, Lithuania, and the Secretariat are sufficient to launch the CI Task Force over a period of 2 years, until it secures research and operation grants from more national or European sources.

12) 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/>

13) 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