



# Study on Copyright and New Technologies: Copyright Data Management and Artificial Intelligence

Stakeholder Workshop on behalf of DG Connect  
24 June, 9-12:30 h CEST

Dr. Florian Berger (Technopolis), Philippe Rixhon (Phillipe Rixhon Associates), Dr. Alfred Radauer (FH Krems)



# Agenda

Time	Agenda Item
9:10-9:20	Welcome by the European Commission (Marco Giorello, Rodolphe Wouters) and the study team
9:20-10:05	<b>Session 1 – Impacts</b> <b>“The impacts of poor rights metadata on the creative industries”</b> Presentation of core results by Florian Berger and Alfred Radauer (20 min), Q&A (25 min)
10:05-11:25	<b>Session 2 – Industry Perspectives</b> Three commentaries on the study findings (25 minutes including a 5-minute Q&A)
11:30-12:20	<b>Session 3 – Options</b> <b>“Towards a performing copyright data framework”</b> Presentation of options by Philippe Rixhon (25 min), followed by Q&A (25 min)
12:20-12:30	<ul style="list-style-type: none"><li>• Wrap-up and outlook on next study steps (Philippe Rixhon)</li><li>• Farewell by the Commission (Rodolphe Wouters)</li></ul>



# **“The impacts of poor rights metadata on the creative industries”**

Stakeholder Workshop, 24 June 2021

Dr. Florian Berger and Dr. Alfred Radauer



## **Three main research questions were addressed in the empirical part of the study.**

- What issues/pain points/challenges can be identified with regard to rights metadata in different creative industries?
- How large are those challenges regarding rights metadata? What can be empirically shown?
- What are the economic impacts of those challenges?



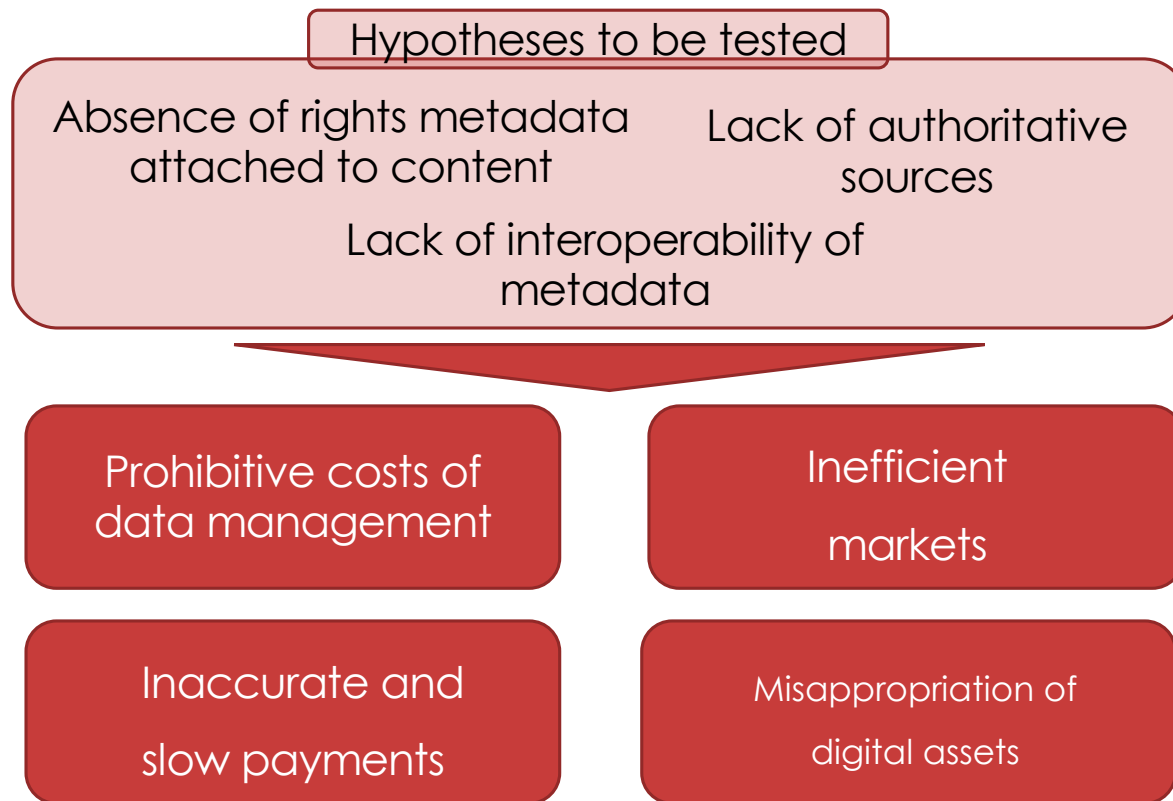


**Using a differentiated set of methodological approaches helped to bring together evidence from different angles.**



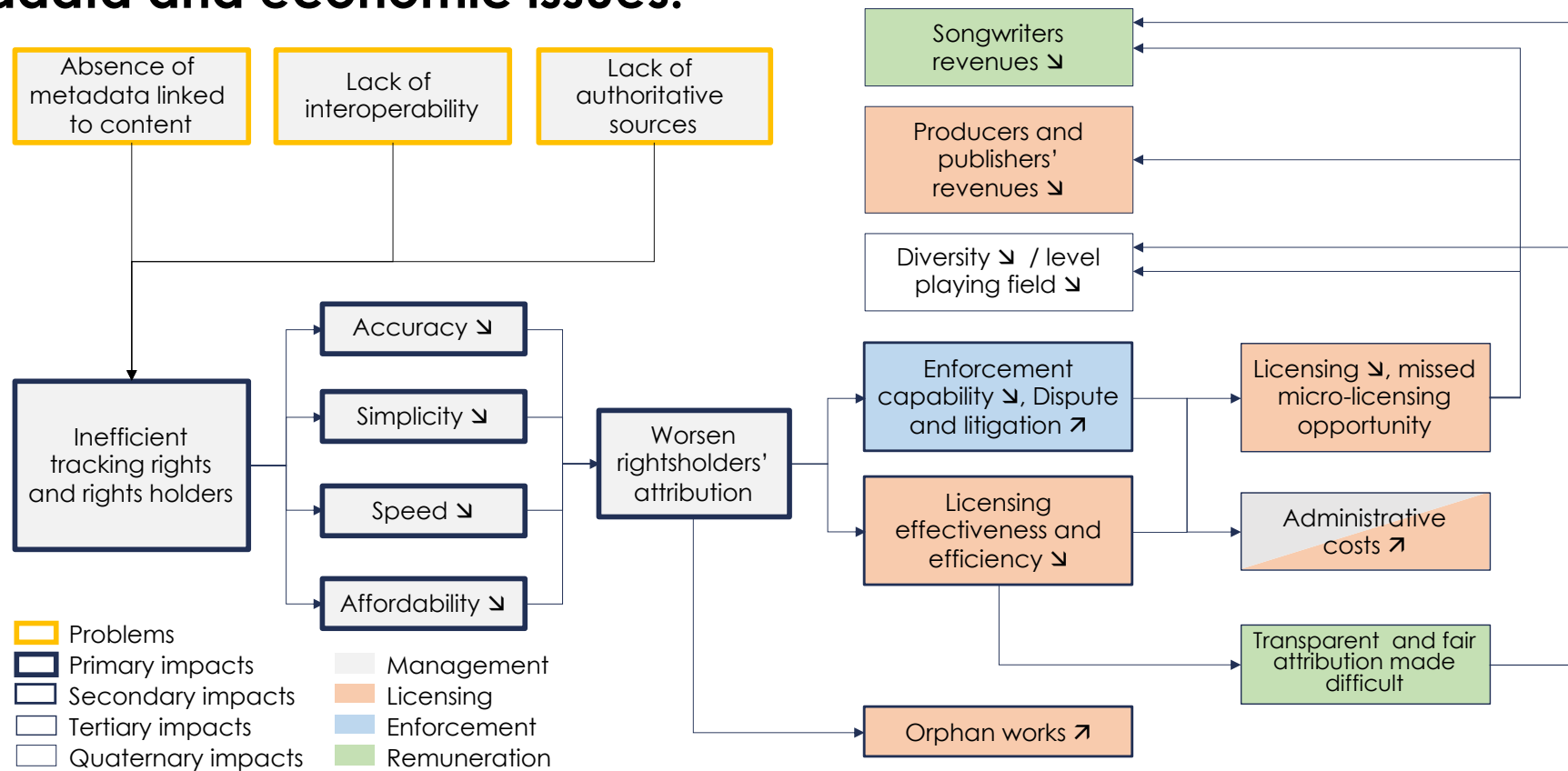


## Ex-ante working hypothesis (!) on challenges and (potential) impacts





## Impact modelling was used to disentangle interrelations between metadata and economic issues.





## Some “nuggets” of empirical evidence – music industry

In top 10% of reported tracks in the statistics of NMP composer/author data is missing for more than 1/3

For 10-50% of works, metadata are missing (survey data)

20 to 25% of music streaming revenue owed to songwriters cannot be correctly allocated

Rights metadata are missing in the music sector “to a small to medium extent (statement from survey respondents)

Anywhere from 20-50 percent of music payments don't make it to their rightful owner (interviewee)

Admin costs due to imperfect metadata could be at least **€ 50 m per year** for the EU recorded music industry

- Overall, an estimate of the works for which rights metadata are missing could be between 20 and 50%
- Additional costs due to metadata issues can be roughly estimated to be (conservatively) € 50 m per year for the European recorded music industry (per year)
- Metadata challenges could cost the industry ~ 10-50% of licensing volume

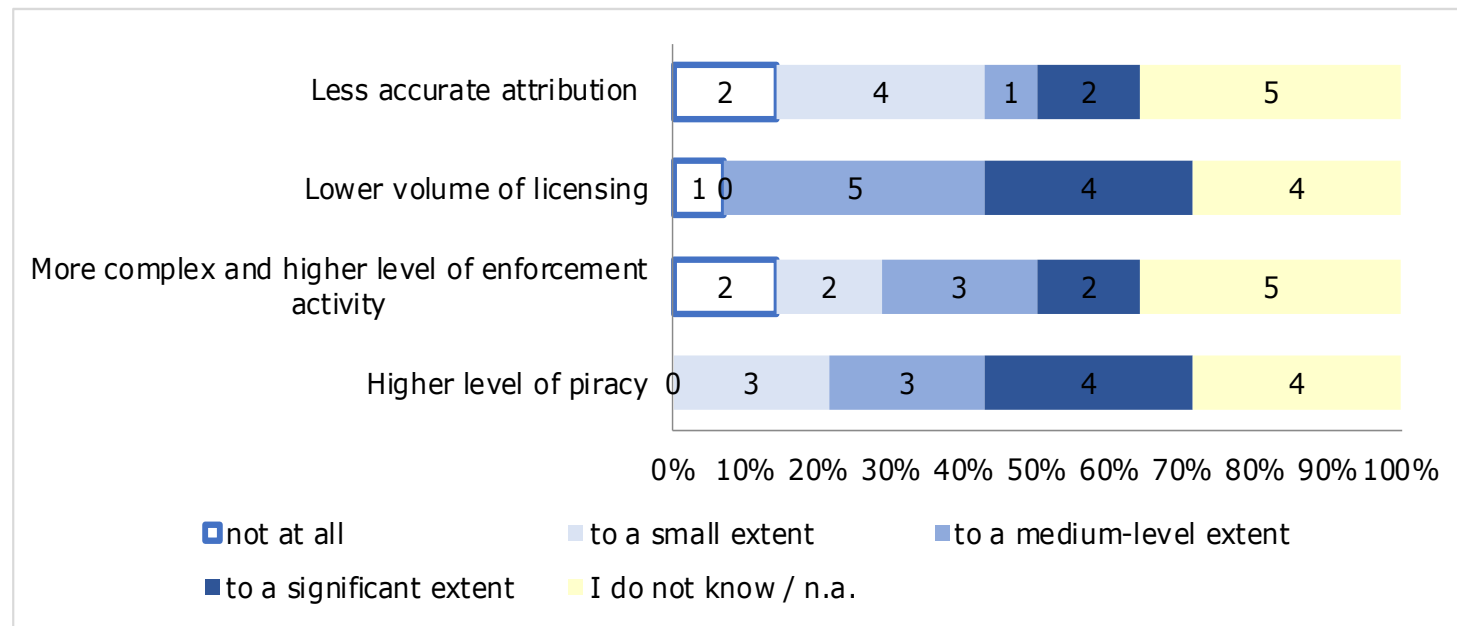


## Nuggets of empirical evidence – publishing industry

- ↗ Very diverse range of answers: From not „a problem at all“ to „considerable issue“, reflecting
  - ↗ different levels of knowledge and awareness of the issue(s) of metadata and copyright infrastructure
  - ↗ different market segments and licensing practices in the publishing industry
- ↗ Seemingly most important aspect is a need to enrich available metadata which is not bad at overall book/trade metadata level with more metadata on assets within the books
- ↗ Further aspects/problem areas
  - ↗ Question of authoritativeness
  - ↗ Images in books
  - ↗ Meta-data to specifically describe rights situation
  - ↗ Calls for more open standards



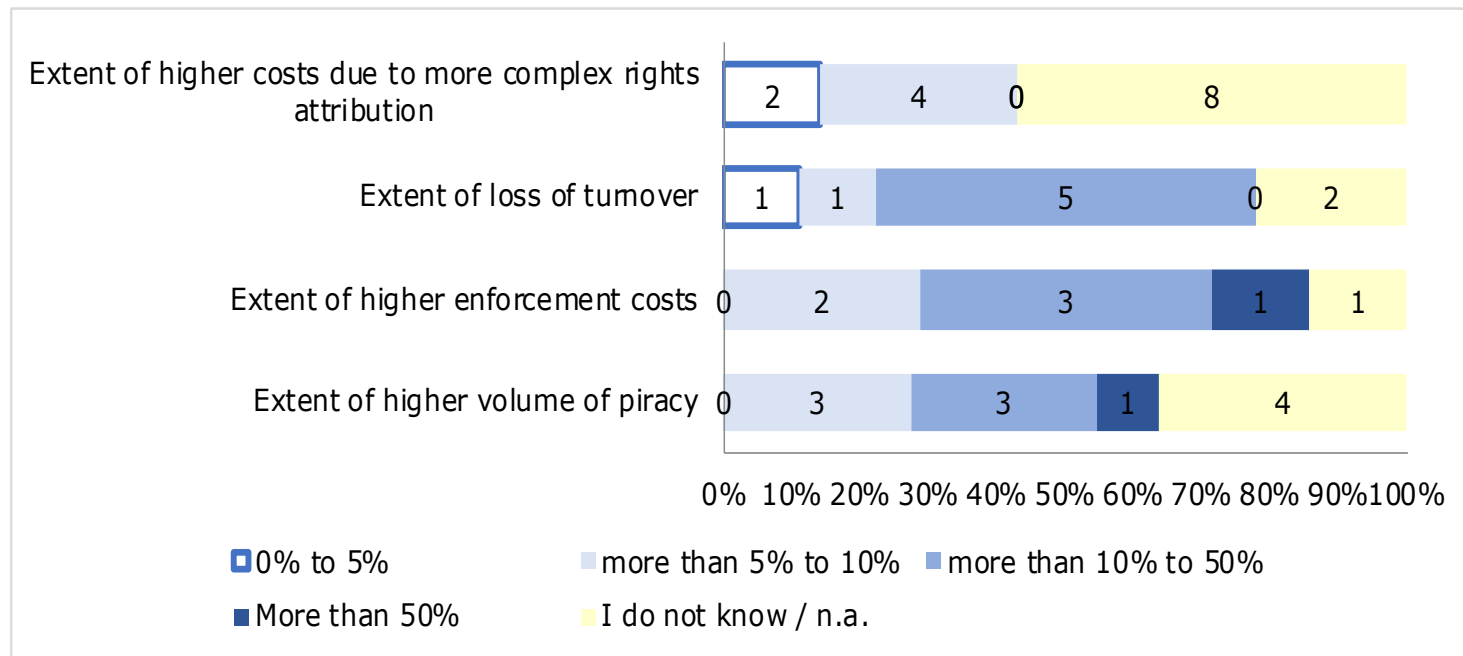
## Approximation of impacts only possible to a limited extent – but effects on licensing activities and piracy mentioned often.



Source: Survey among stakeholders in the publishing sector, April/May 2021, participants were asked to what extent a lack of copyright metadata could lead to different impacts.



## Approximation of impacts only possible to a limited extent – but mentionable effects on on turnover and costs reported.



Source: Survey among stakeholders in the publishing sector, April/May 2021, participants were asked to what extent a lack of copyright metadata could lead to different impacts.



## Nuggets of empirical evidence – Film and TV production and broadcasting



Delgado et al (2006). Metadata and Rights Interoperability for Content Interchange between TV Programs Producers, Proceedings of the 10th International Conference on Electronic Publishing pp. 267-278.)

Ma et al (2016). The Dual Impact of Movie Piracy on Box-Office Revenue: Cannibalization and Promotion, Carnegie Mellon University





## Overall observations, summary and wrap-up

- ↗ High diversity between industries with respect to rights metadata awareness, existing industry initiatives, standards and priorities
- ↗ However, significant challenges in all of the three industries seem to persist
- ↗ Precise assessment is difficult even for industry experts, but the combination of different insights suggest room for improvement



## **Q&A - Session**

Discussion, Questions, Comments

**technopolis**  
group

PHILIPPE  
RIXHON  
ASSOCIATES

 **UCLouvain**

**DALDEWOLF**

 **FH KREMS**  
UNIVERSITY OF APPLIED SCIENCES



# Thank you.

Abidjan • Amsterdam • Berlin • Bogotá • Brighton • Brussels • Frankfurt/Main •  
London • Paris • Stockholm • Tallinn • Vienna



# **Copyright Data and New Technologies**

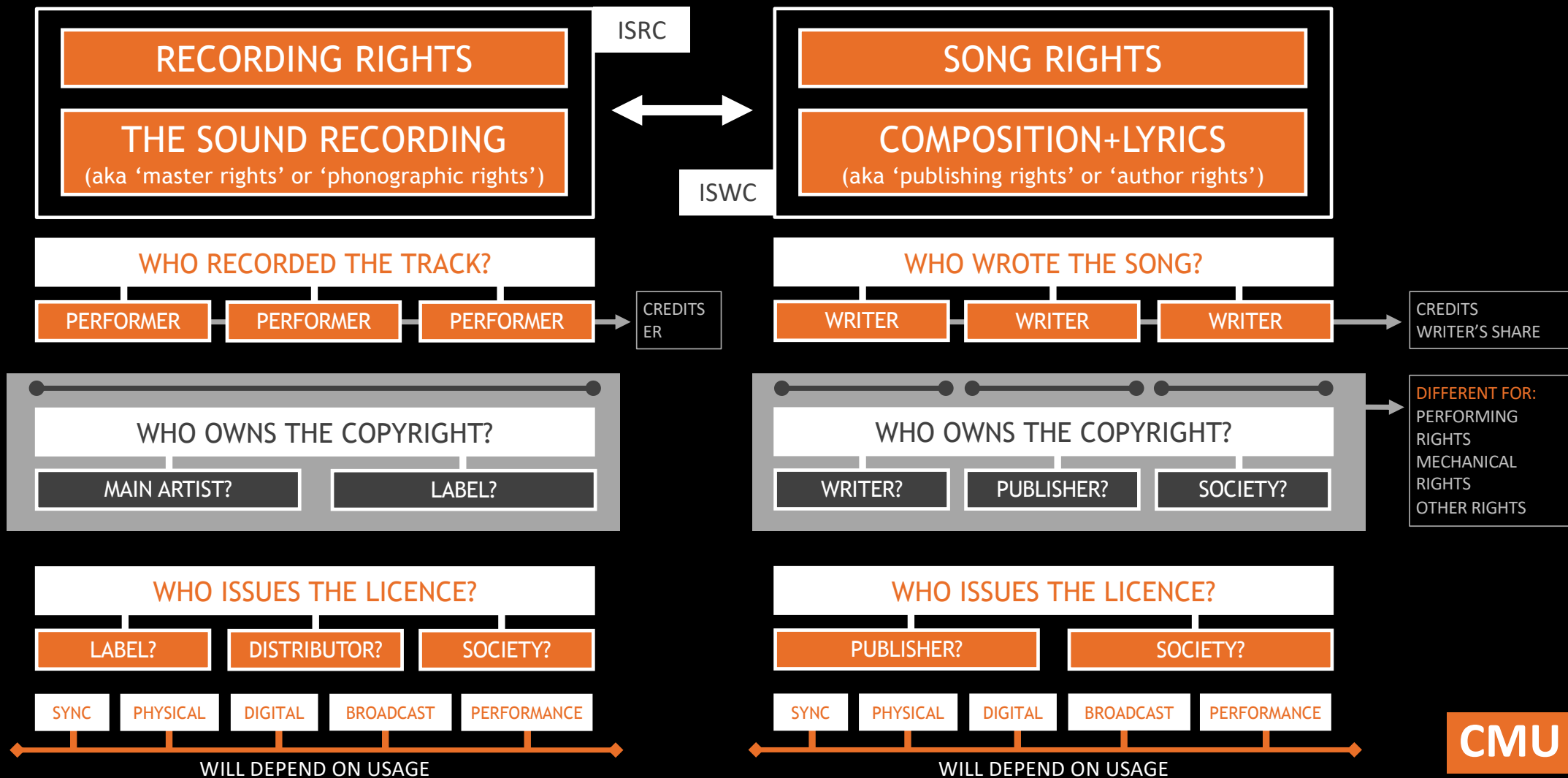
## **Towards a collective copyright infrastructure**

### **Perspectives from the music industry**

Stakeholder Workshop, 24 June 2021

Chris Cooke and Turo Pekari

# MUSIC RIGHTS DATA: WHAT WE NEED TO KNOW...



# MUSIC RIGHTS DATA: WHY DOES THIS MATTER?

WHO DOES  
THE DEAL?

WHO DO  
WE PAY?

WHO DO  
WE CREDIT?

WHAT DO WE  
RECOMMEND?

PEOPLE  
GET  
SUED

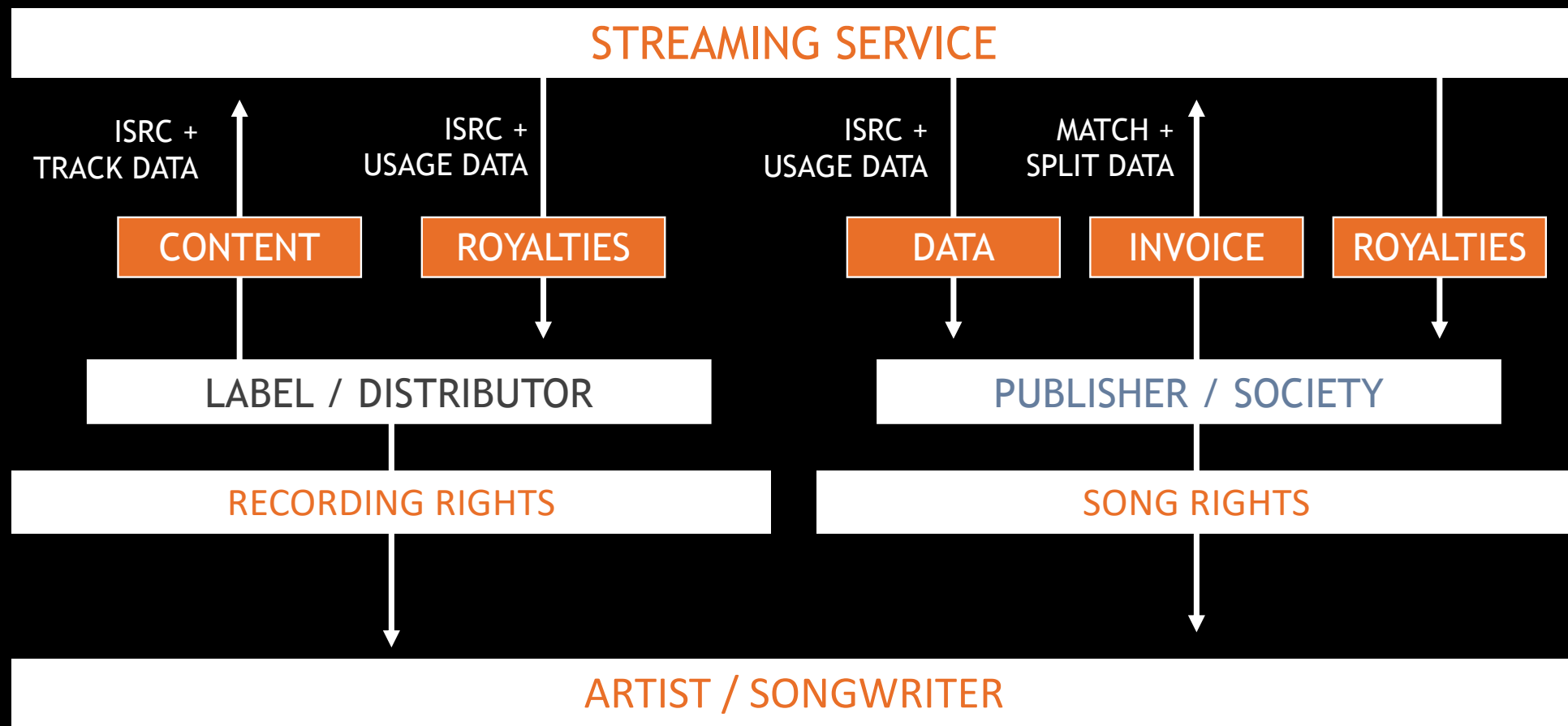
CREATORS  
DON'T GET  
PAID

MORAL  
RIGHTS  
INFRINGEMENT

INFERIOR  
USER-  
EXPERIENCE

WHAT HAPPENS WITHOUT THE DATA?

# MUSIC RIGHTS DATA: STREAMING ROYALTIES



# Towards a collective copyright infrastructure

## Data awareness

- Education and knowledge about metadata issues
- Support and for utilisation of metadata tools
- Support for development of industry testbeds

## Tools and practices

Standards

Formats

APIs

Data exchange platforms

## Culture

- Collaborative mindset
- Culture of data sharing
- Culture of trust
- From competition to co-opetition



**Thank you!**

**Questions or comments  
from the audience?**



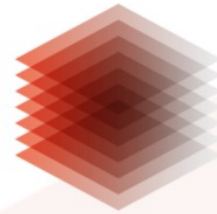
# **Copyright Data and New Technologies Established and emerging standards Perspectives from the publishing industry**

Stakeholder Workshop, 24 June 2021

Lambert Heller, Paola Mazzuchi and Piero Attanasio

---

LEIBNIZ INFORMATION CENTRE  
FOR SCIENCE AND TECHNOLOGY  
UNIVERSITY LIBRARY



TIB

# **Copyright and New Technologies: A perspective from cultural heritage and research infrastructure**

Lambert Heller  
Study on Copyright and New Technologies:  
DG CNECT Stakeholder Workshop of 24 June 2021



# DOI, DataCite, Open Research Knowledge Graph: Persistent Identifiers have come a long way

DataCite Search

Works People Data Centers Members Support Sign in

Data from: Costs of CRISPR-Cas mediated resistance in *Streptococcus thermophilus*

Pedro F. Vale, Guillaume Lafforgue, François Gatchitch, Rozenn Gardan, Sylvain Moineau & Sylvain Gandon

DataPackage published 2015 via Dryad Digital Repository

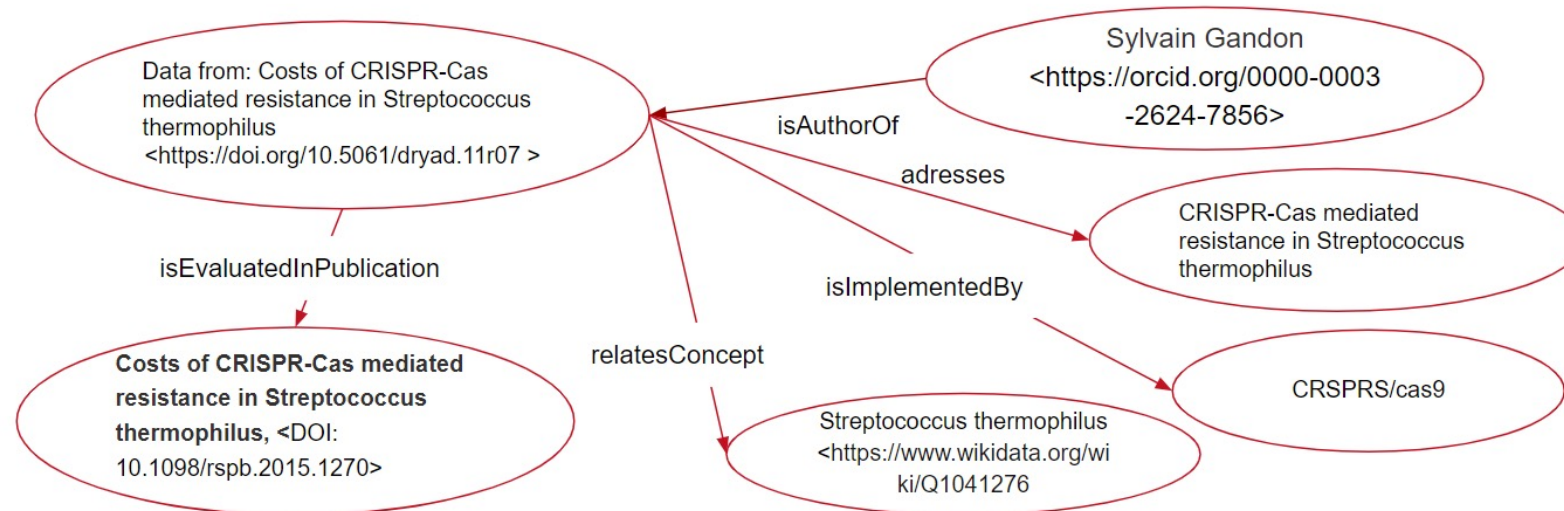
Has part 1

Is referenced by 1

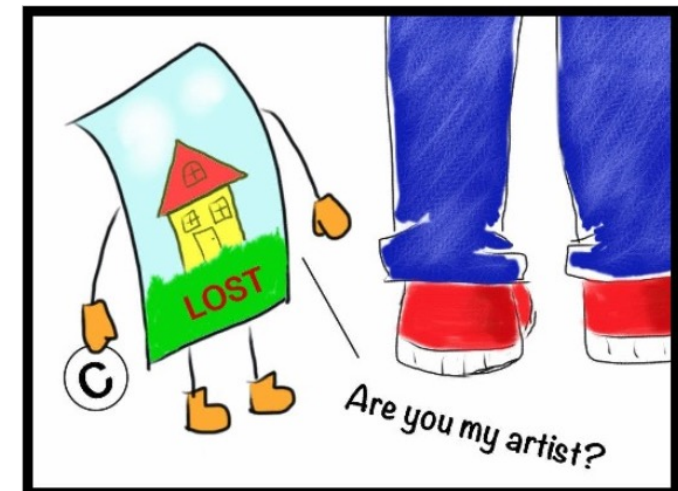
Is supplement to 1

<https://doi.org/10.5061/dryad.11r07> Cite

Author	Sylvain Gandon; Pedro F. Vale
Research Problem	CRISPR-Cas mediated resistance in <i>Streptococcus thermophilus</i>
Methods	CRISPR/cas9
related Concepts	<i>Streptococcus thermophilus</i> , immunity
related Publication	DOI: 10.1098/rspb.2015.1270



Copyright is evolving, complex and disputed - we need open, permissionless means to identify, claim and transfer content



## Further reading / watching

- TIB's director Sören Auer recent ERC grant project “Open Research Knowledge Graph” <https://www.orkg.org/>
- My talk “The path to Self-sovereign identity (SSI) in research” <https://youtu.be/0XLzylhYBF8>
- ConDIDI, my most recent EC-funded SSI project <https://projects.tib.eu/condidi/>
- Some more context on me and my R&D group at TIB <https://tib.eu/osl> respectively <https://tib.eu/Lambo>



LEIBNIZ INFORMATION CENTRE  
FOR SCIENCE AND TECHNOLOGY  
UNIVERSITY LIBRARY



## MORE INFORMATION

**tib.eu**

### Contact

Lambert Heller

T +49 511 762-5348, [lambert.heller@tib.eu](mailto:lambert.heller@tib.eu),  @Lambo



Creative Commons Attribution 3.0 Germany  
<https://creativecommons.org/licenses/by/3.0/de/deed.en>

*Leibniz*  
Leibniz  
Association

## Principles for the copyright infrastructure

- 1) Rights data management is an independent function from rights management;
- 2) It should be based on standards to guarantee interoperability;
- 3) It should remain distributed...
- 4) ... and voluntary;
- 5) Solutions should be neutral to content genre and to business models  
*commercial vs. not-for-profit; direct vs. collective licenses; subscriptions vs. pay per use, etc.*



# Innovation needs

Developments must be driven by market needs

*but*

Market forces alone are not sufficient

Support innovation in

Enabling technologies (standards, resolution mechanisms, tools, protocols...)

Licensing and information services coherent with the principles

**Thank you!**

**Questions or comments  
from the audience?**



# **Copyright Data and New Technologies**

## **AI and blockchain, two keys to a level playing field and fair remuneration, perspectives from the film industry**

Stakeholder Workshop, 24 June 2021

Maria Tanjala, Sami Arpa and Sten-Kristian Saluveer

# Largo.ai

NEXT GENERATION  
STORYTELLING

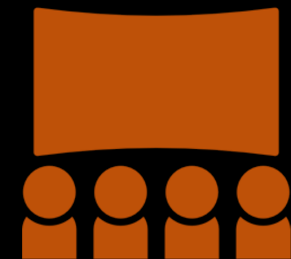
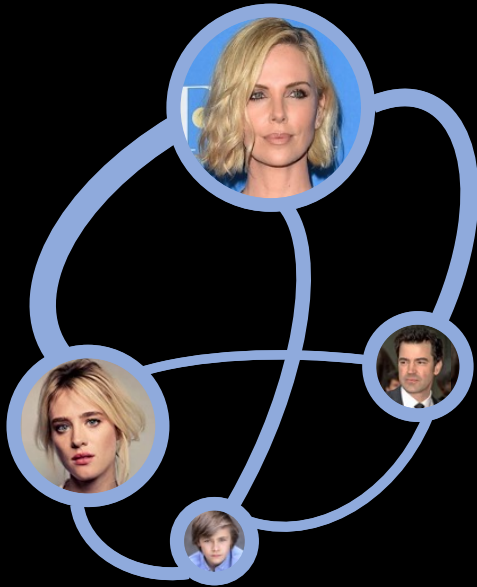
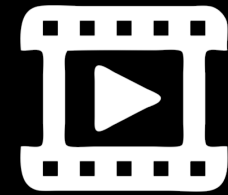
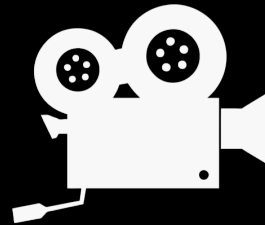
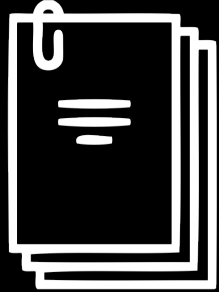


Internationale  
Filmfestspiele  
Berlin

EFM  
Start-ups  
2020



# A FILM PRODUCTION INCLUDES MANY IMPORTANT DECISIONS



WHAT IS THE DECISION MECHANISM IN THE  
INDUSTRY FOR ALL THOSE DECISIONS?

- Gut Feeling

**PROBLEM:**

Netflix and Amazon use **data-driven intelligence** instead of GUT  
FEELING for production decisions and they are disrupting the industry.

The rest of industry has limited access to such technologies.

This is where Largo comes in as a 3rd party provider.

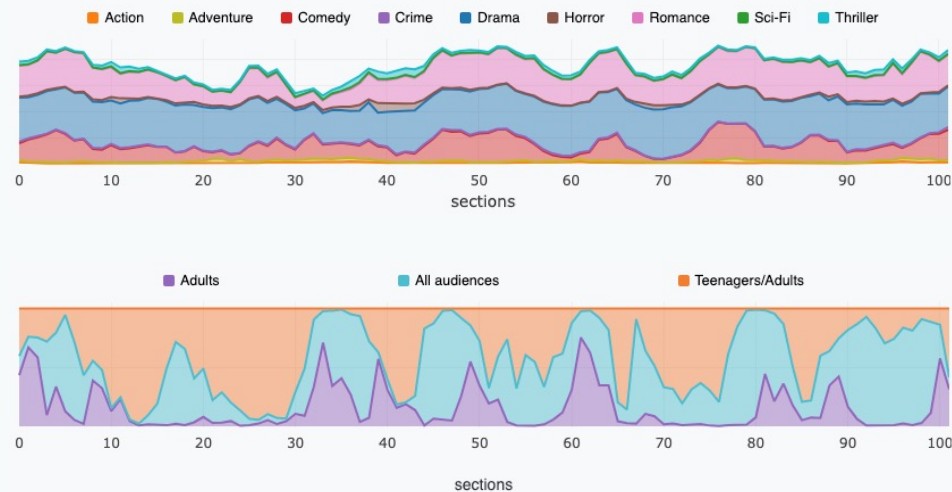


**Combine your gut feeling and expertise with artificial intelligence**

SaaS platform: instant results with the movie script and video input providing story insights and market potential for producers, distributors, screenwriters, and financiers.

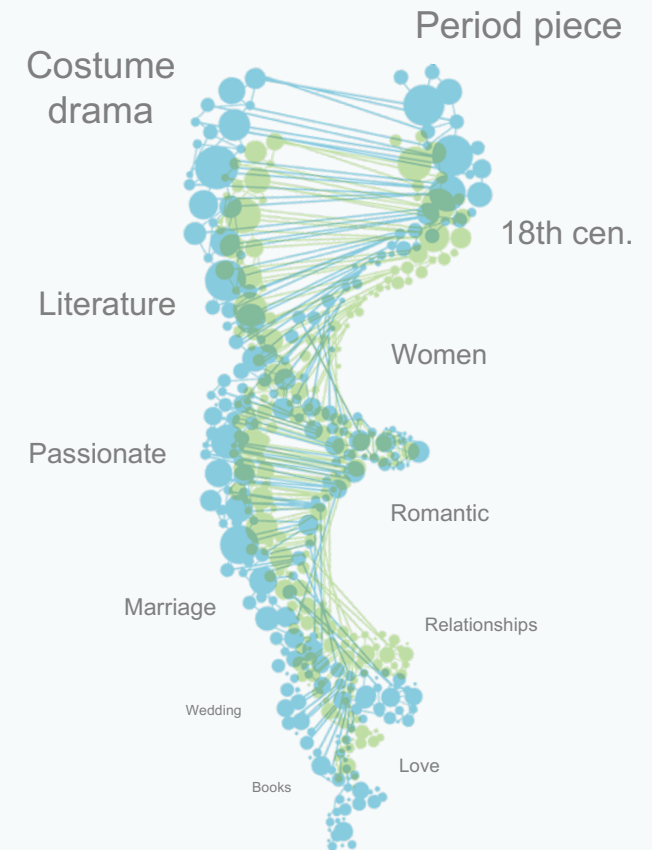


# INNOVATION



Our patented neural network system finds the main ingredients of a content from a script or a video, which enables to provide:

- Content insights
- Casting propositions
- Market forecast



PATENT – PCT PUBLISHED  
WO 2020/222027 A1



# LEARNING FROM A RICH DATA

**400,000**

Movies/series metadata

**1,800,000**

Talents analyzed

**49,000**

Movie scripts/dialogues analyzed

**25,000**

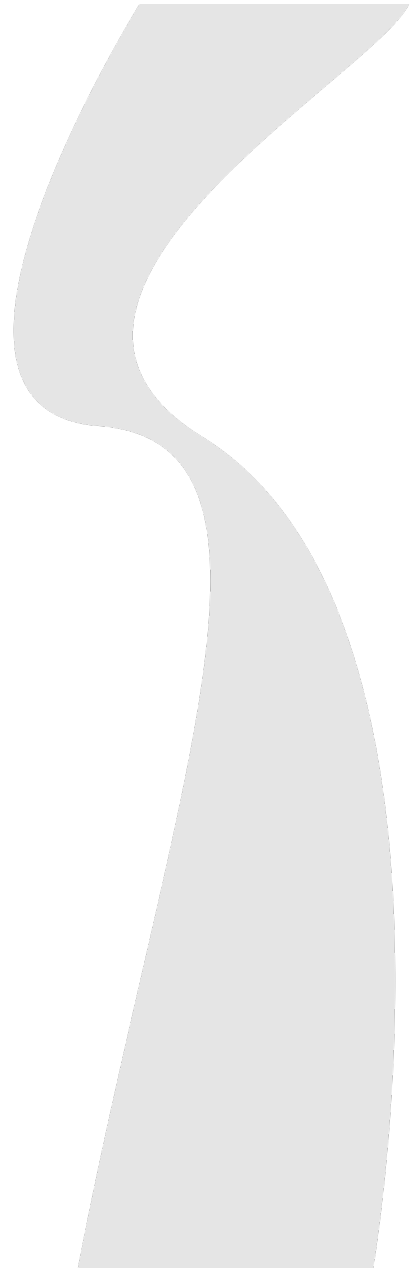
Movie trailers analyzed

**%80+**

Accuracies in financial forecasts

**5 min.**

Required analysis time for a project



# RECENT CASE STUDY WITH LARGO.AI

**22**

European  
producers

**80**

European  
movies



August 2020

January 2021

6 MONTHS ACCESS TO LARGO.AI TOOLS

**58%**

Improvement on financial  
forecasts

**60%**

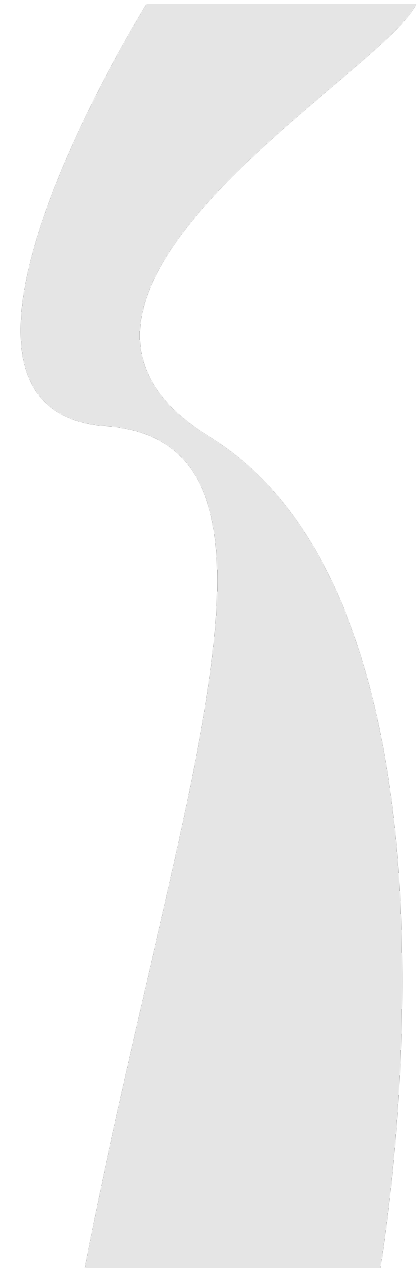
Continue using AI tools

# AI: A DEMOCRATISATION TOOL FOR THE MOVIE INDUSTRY

Making the learnings from data accessible to all stakeholders in the industry in an affordable way

Maximizing the value of the content / better deals with the distributors and platforms

Creating opportunities for the new talents





EFM Workshop: <https://bit.ly/3f9jL6p>

**Sami Arpa**  
sami.arpa@largofilms.ch

Supported by

**EPFL**

**SICTIC**

**X FIT**  
fondation  
pour l'innovation  
technologique

**WISSENSCHAFT.  
BEWEGEN**  
GEBERT RUF STIFTUNG

 **Innosuisse**

 **canton de  
vaud**

# Directive (EU) 2019/790 - On copyright and related rights in the Digital Single Market

## Fair remuneration in exploitation contracts of authors and performers

Article 18: Principle of appropriate and proportionate remuneration

Article 19: Transparency obligation

Member States shall ensure that authors receive on a regular basis, at least once a year, and taking into account the specificities of each sector, up to date, relevant and comprehensive information on the exploitation of their works from the parties to whom they have licensed or transferred their rights.

---

Philippe Rixhon: “The lack of reliable authoritative data has a negative impact on rightsholders who may miss revenue streams linked to their unidentified content.”

We are  FilmChain

[filmchain.co](https://filmchain.co)

We empower participants in film and TV industry with **access to money, reports and data.**

Backed by



Innovate UK

Imperial College  
London



Barclays Accelerator  
 POWERED BY techstars

# Problems

- Film & TV stakeholders - producers, financiers, sales agents, talent - don't know how much they are owed and when; participation management is complex.
- Payments are delayed for years or never reach the right owners.
- Companies struggle with manual, opaque, error-prone systems.
- Participants are paid in cheques and can't verify their entitlements.

## FilmChain fintech solution

An impeccable fintech solution with verification mechanisms must replace the suboptimal, manual, legacy processes.



Offers instant access to money, reports and data



Unifies the management of entitlements and rights



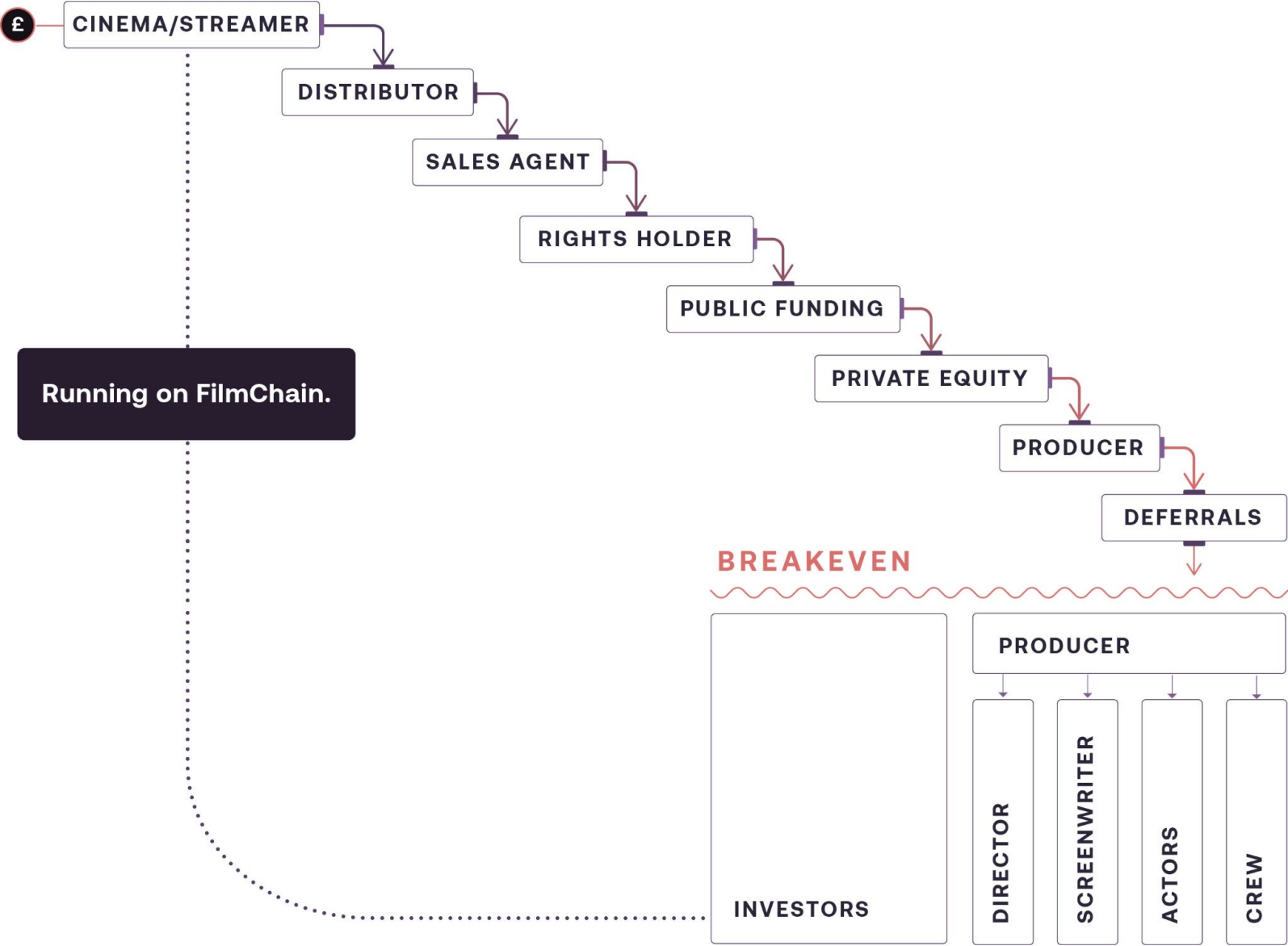
It's transparent and auditable



Offers analytics & performance insights



FilmChain  
money flow

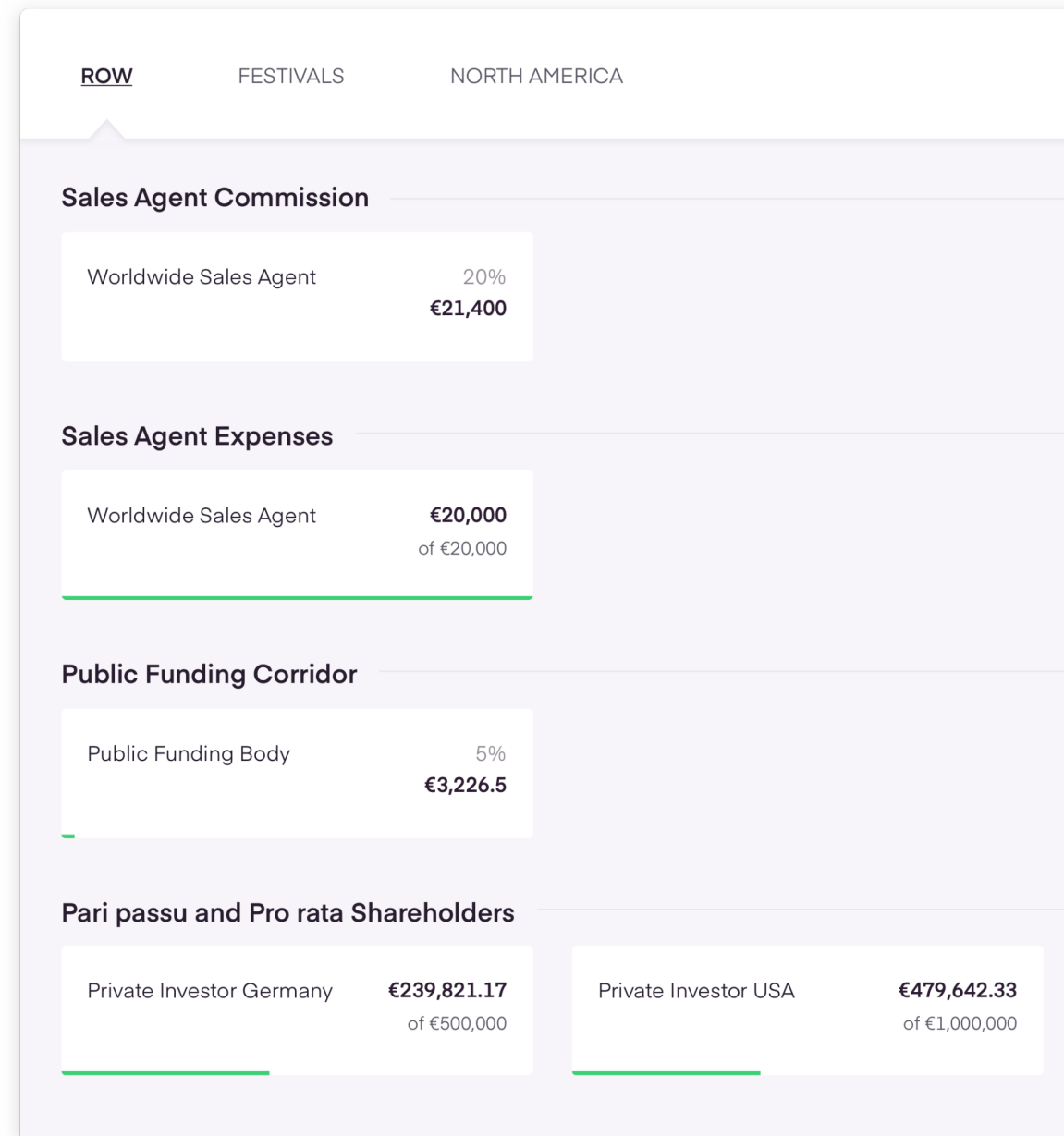


# Product - Digital CAM reimagined

— Recoupment schedule - digitised royalties dashboard for countless of beneficiaries.

— Transactions - the distribution of every incoming transaction

— Wallet - the individual user's visualisation of all payment information and withdrawals.

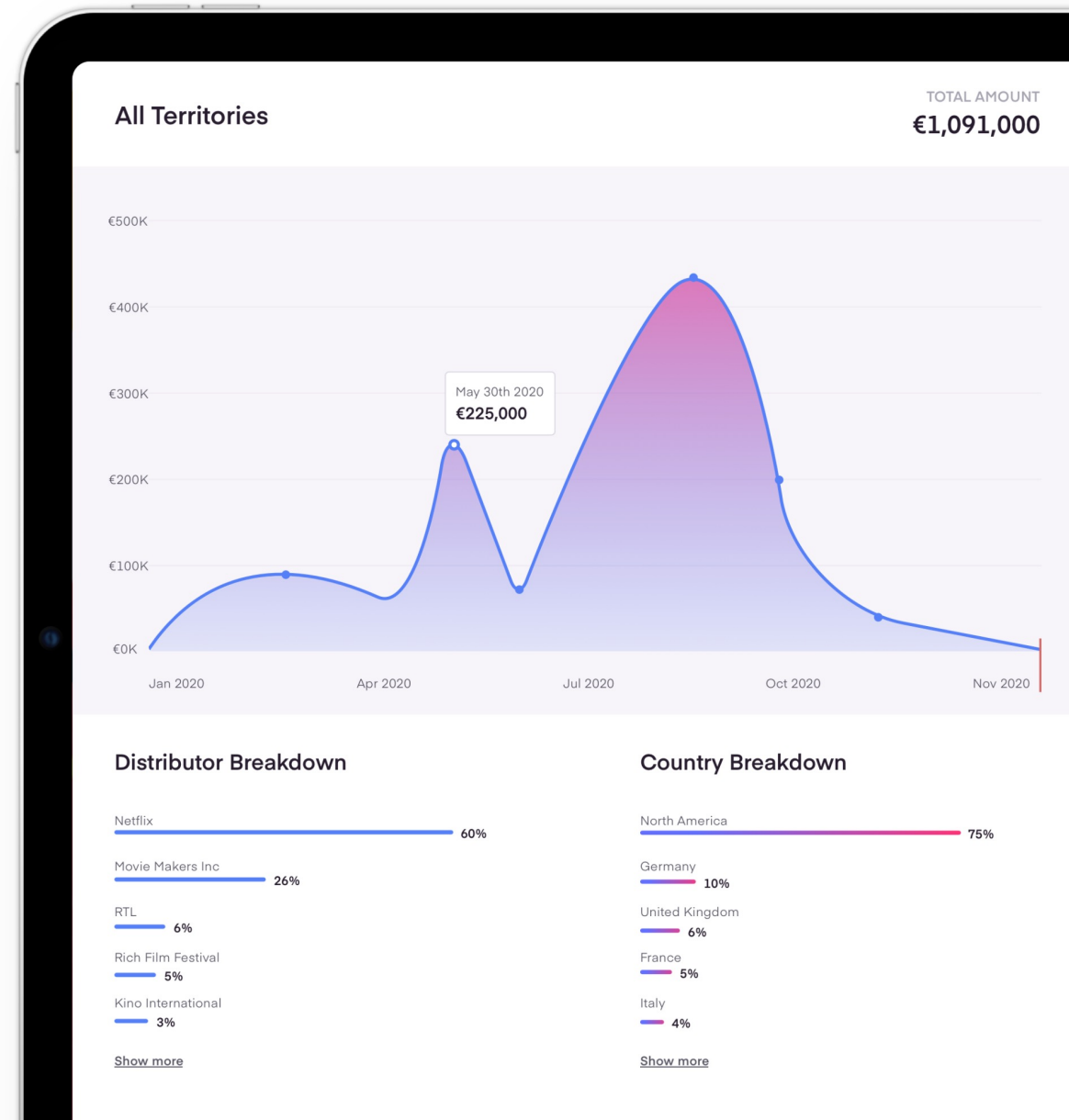


# Analytics and Rights Tracker - data on demand

— Rights Tracker - add, track and manage new sales

— Analytics - rich insights and data visualisation.

Financial predictions to be released in Q4.



# What we need

- **Enforcement** of the Copyright directive on fair remuneration and transparency obligations across all sectors of exhibition, distribution, marketing and exploitation of content.
- **Support** with developing data processing infrastructures such as FilmChain and other companies that already invest financial capital and know-how in technology to achieve better dissemination of money and data.
- **Funding** for media tech and createch startups.

Mission: get the  
money made by  
the films we love to  
the right people.

CONTACT US

maria@bigcouch.co

[filmchain.co](http://filmchain.co)

Thank you!



Startup of the  
Year

London Business  
Awards 2020



Zinemaldia  
Startup Challenge

San Sebastian Film  
Festival 2020



Equals Money  
Award

Makers &  
Shakers 2020



Best Woman-led Investment In  
Innovation

UKBAA 2019



Blockchain Founder  
of the Year

UK TFA 2018

**Thank you!**

**Questions or comments  
from the audience?**

**technopolis**  
group

**PHILIPPE  
RIXHON  
ASSOCIATES**

 **UCLouvain**

**DALDEWOLF**

 **FH KREMS**  
UNIVERSITY OF APPLIED SCIENCES



## **Copyright Data and New Technologies Towards a performing copyright data framework**

Stakeholder Workshop, 24 June 2021

Philippe Rixhon

## Metadata; data about data

### Rights management information defined in *acquis communautaire*

- **Identification metadata**  
*What is what, who is who*
- **Rights metadata**  
*Who did what, who owns what, what can we do with that*

### Additional metadata defined by industry practices

- **Descriptive metadata** *for search and enjoyment*
- **Usage metadata** *for a fair trade*
- **Administrative metadata** *for the trust in data*

In the digital world, every content is data. .jpeg, .mp3, .mp4, .pdf, .epub, you name them, all of them are data. And all data about data is metadata.

Two types of metadata are mentioned in the *acquis communautaire* –

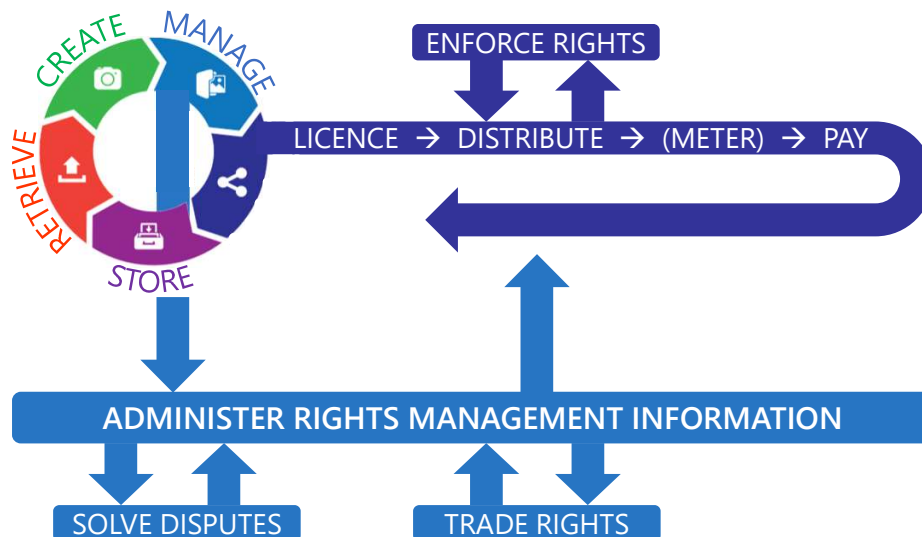
- Identification metadata, the smallest set of data which uniquely identifies works, related subject matters, and stakeholders, and
- Rights metadata that informs about *who did what, who owns what, and what can we do with that*

There are three other types of metadata –

- Descriptive metadata, helpful to search for a particular entity, to enrich its enjoyment, or for general interest. *Is it jazz or hip-hop, landscape or portrait?*
- Usage metadata, essential to monitor and communicate usage, monetisation, and remuneration of content, and
- Administrative metadata, necessary to assess the reliability and trustworthiness of the other metadata



## Purposes of rights management information



METADATA | LAISSEZ-FAIRE | OPEN COPYRIGHT DATA FRAMEWORK | WAY FORWARD

2/20

A look at the generic purposes of rights management information in the creative industries. The digital content supply chain goes through 5 phases: *create*, *manage*, *distribute*, *store*, and *retrieve*.

The phase *manage* generates the data which is necessary for the phase *distribute*.

The phase *distribute* is divided into licence the content, distribute the digital asset, eventually meter the consumption of the content, and pay the royalties. No data, no money, poor data, poor payment.

We also rely on metadata when distributed content needs protection and detectability to enforce the rights of authors or owners.

Two people can claim full ownership of the same content, or two companies can argue about the availability of a digital asset. Disputes must be solved. Their resolution requires trustworthy metadata.

Finally, rights can be traded. That, in turn, entails accurate and transparent information.

## Rights data management explicitly or implicitly specified in the *acquis communautaire*

- **Identify** works (and other subject matters), rights, and stakeholders
- **Link** works and rights, works and stakeholders, rights and stakeholders
- **Monitor** usages of works and remunerations of rights
- **Fulfil** human-and-machine-readable contracts
- **Protect** personal data and privacy as well as business confidentiality

A data analysis of the *acquis communautaire* related to copyright leads to a generic specification of rights data management.

First, the creative ecosystem must be able to identify content, rights, and stakeholders, not only rightsholders but also any stakeholder who must be identified to manage regulatory exceptions or limitations.

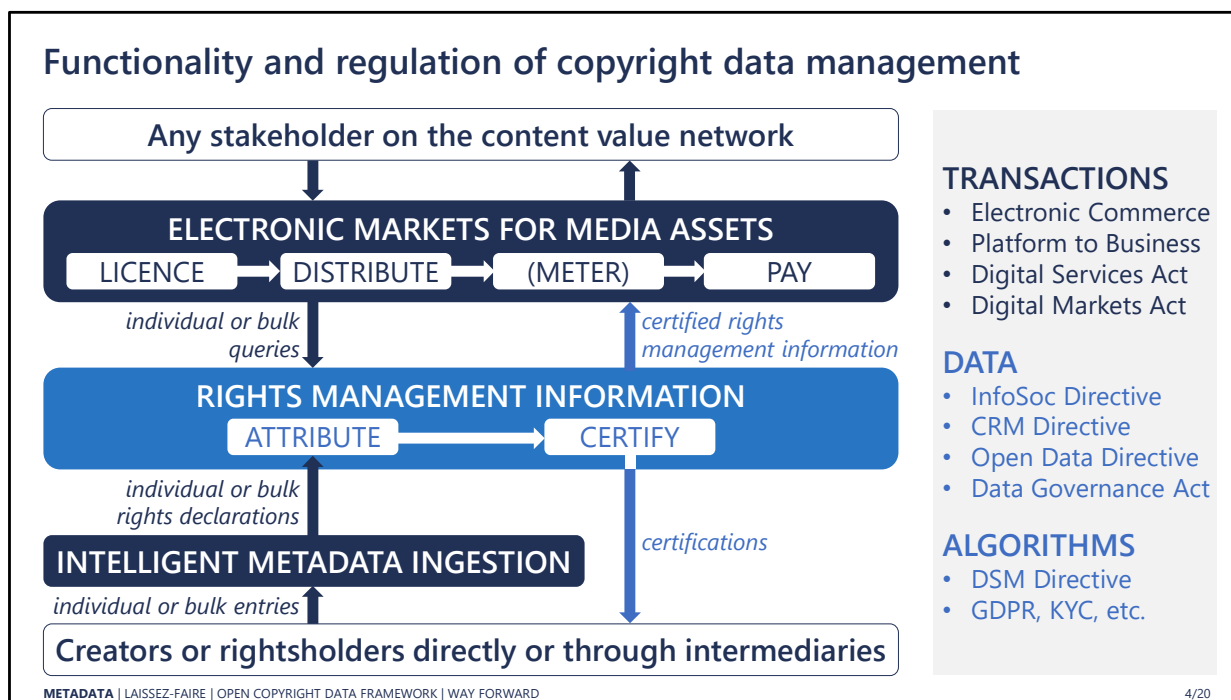
Then, one must manage the relationships between these three types of entity; works, rights, and stakeholders.

Then, one must monitor the flow of content from creators to consumers and the counterflow of remuneration from consumers to creators.

This is only possible through the fulfilment of human-and-machine-readable licensing agreements, human-readable to ensure the transparency and enforceability of agreements, machine-readable to cope with the exponential volume of transactions.

And all of that while respecting personal data, privacy, and business confidentiality.

And all of that on the whole content value network, which is extremely complex, dynamic, and fluid.



A quick look at the generic functionality of copyright data management.

The information must be stored somewhere. It is entered into databases that you can also call registries. The metadata should be ingested through intelligent interfaces, not systems that requires users to be at the same time IT scientists and IP experts, because most rightsholders are not. This function attributes a work to an author. *Maria wrote that song. John took that picture.*

Then, one way or the other, the rights information is certified. *It is true that Maria wrote that song, you can trust me. It is true that John took that picture, you can trust me.*

On the other side of the databases, we have a multitude of stakeholders. They meet, one way or the other, on what I call B2B or B2C Electronic Markets for Media Assets.

To licence and distribute content, they need data from the bases.

Interestingly, these different layers correspond roughly to different regulations. Distinguishing among layers helps distribute roles and liabilities, master regulatory complexities, trigger business opportunities, build APIs, and secure interoperability.

## Issues around rights metadata

### ROOT CAUSES

- Lack of rights awareness
- Lack of understanding of metadata flows and purposes
- Lack of cooperation

### PROBLEMS IDENTIFIED AT THE COUNCIL

- Missing or erroneous metadata
- Lack of interoperability of metadata formats
- Lack of authoritative sources

### INDUSTRY PAIN POINTS

- Prohibitive costs of data management
- Inefficient markets
- Inaccurate and slow payments
- Misappropriation of digital assets

The working group on copyright at the European Council identified three main problems. Metadata can be missing or erroneous, and, although progress has been made in various ways in different sectors, data interoperability and source authority are not yet what they should be.

These problems have root causes. There is a lack of rights awareness at both ends of the media supply chain. Creators are insufficiently aware of their rights. Consumers are insufficiently aware of the rights of others. In the middle of the chain, where people deal with rights on a daily basis, one can observe an insufficient understanding of metadata flows and purposes. Last but not least, one must admit that cooperation is not yet adequate. Dialogues are good to define problems, but cooperation is needed to solve them.

The data problems are creating issues all over the content industries. There are, of course, differences in strengths and priorities. In general, handling all the data necessary for or generated by the distribution of digital assets costs a fortune. The remunerated exchange of content value is still a headache. Think about articles 15 or 17 of the directive on copyright. Payments are slow and often inaccurate, mostly due to the cumbersome treatment of metadata. And piracy is still a big issue.

## Prospective consequences of *laissez faire*

- Minimal **compliance** of large stakeholders
- **Struggle** of creators and small & medium media enterprises
- **Demise** of the Digital Single Market
- **Risk** for content diversity, European culture and identity

What will happen if we let it go?

The large stakeholders, at any place on the value network, are resourceful. They deploy proprietary, commercial data systems to comply *a minima* with the *acquis communautaire* and its 27 different implementations.

European creators and small & medium enterprises do not have such human and financial resources. They struggle to cope with the specified rights data management. They depend on systems mostly designed for the analogue world and adapted half-heartedly for the digital era. Basically, if we let it go, the problems will remain.

A true, actionable digital single market will not emerge without a genuine single market for data, in other words, without interoperable metadata from authoritative sources.

Last but not least, there is a real danger of cultural uniformization.

## The copyright data framework

- ... **is** a structure **underlying** the exchange of rights management information built upon a set of **foundational standards** and technologies supporting a **distributed network** of rights declarations, attributions, certifications, and queries.
- Standards for identifiers, metadata formats, and data exchanges emerged from the further developments of **European initiatives**.
- A framework is **minimally prescriptive** but **maximally supportive and inclusive**. It allows **many solutions** to be used while enabling **numerous ways** in which one can **cooperate** in originating, enriching, governing, and distributing trusted information, helping streamline current processes and trigger innovative businesses.

METADATA | LAISSEZ-FAIRE | OPEN COPYRIGHT DATA FRAMEWORK | WAY FORWARD

7/20

The copyright data framework exists. It is the structure underlying the exchange of rights management information. An *under-lying* structure is an *infra*-structure. This existing data framework is built upon a set of foundational standards and technologies. It supports the distributed network of rights declarations, attributions, certifications, and queries in the digital era.

Many standards for identifiers, metadata formats, and data exchanges emerged from the further developments of European initiatives such as <indecs>. <indecs> developed a framework of metadata standards to support eCommerce based on intellectual property with the support of the European Commission. EDItEUR and mEDRA are two other relevant European projects among others.

A framework is minimally prescriptive but maximally supportive and inclusive. It allows many solutions to be used. It enables numerous ways in which individuals and organisations can cooperate in creating, enriching, governing, and distributing trusted information. It helps streamline current processes and trigger innovative businesses. It must be neutral to business models. It must benefit everybody in the ecosystem.

## Opening the copyright data framework

... requires to make it **incorruptible**, **trustworthy**, and **independent** at hand of –

- **Open Content-Dependent Identifiers**
- **Digital Identity Wallets**
- **Content Binding Protocols**
- **Content Certification Protocols**

*Network protocols are sets of conventions that dictate how to format, transmit and receive data so that network devices can communicate, regardless of the differences in their underlying designs or standards.*

As I said, the copyright data framework, the *infra*-structure, exists. Can we improve it for the benefit of the creative industries? Yes, if we open it up.

The framework organises the exchange of data among the various stakeholders of the content network. Its governance is backed by identifiers and protocols.

Opening the copyright data framework requires to make it incorruptible, trustworthy, and independent. Two types of identifiers, and two types of protocols can help.

Open content-dependent identifiers depend neither on registration authorities nor on commercial companies. They bridge the gap between high-level content identifiers, for example ISRC, ISWC, ISBN or ISAN, and their connections to digital manifestations of content.

Digital identity wallets will be useable either to identify users or to prove certain personal attributes, typically to access public or private digital services. No need any more to have an account.

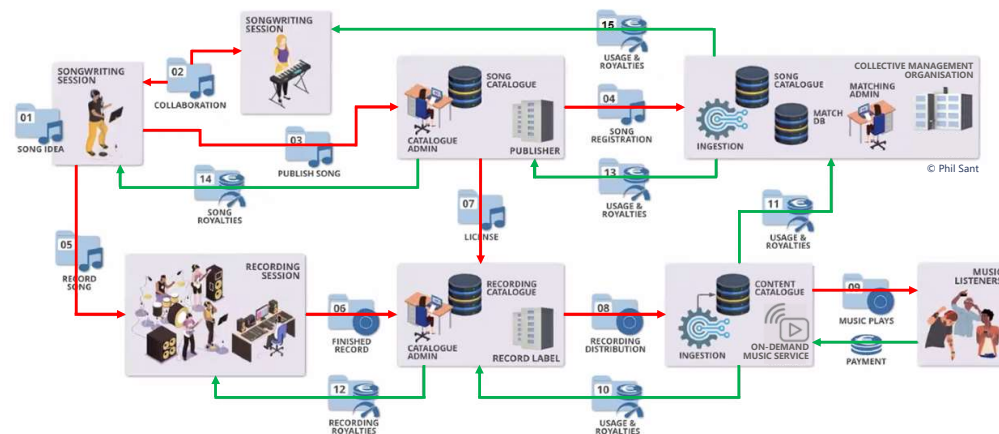
A content binding protocol bind immutably open content-dependent identifiers of works, or related matters, with identities of creators or rightsholders. At that binding, one can connect dynamic rights metadata via permissioned links. This immutable binding mechanism is comparable to bookkeeping. You may not delete a booking but must add a new booking to the previous ones to show what happened with a

value or to correct an erroneous booking. That mechanism allows the auditor to check the accountant's ledgers. A similar mechanism is used for the registration of nominal shares in a company, or for land or company registries.

Rights metadata are assertions about *who did what, who owns what, and what can we do with that*. A content certification protocol can create trust in such assertions and in content authenticity to ensure the accountability of entities, even if they must or prefer to remain pseudonymous.



## Opportunity 1: addressing inaccurate & slow payments to songwriters



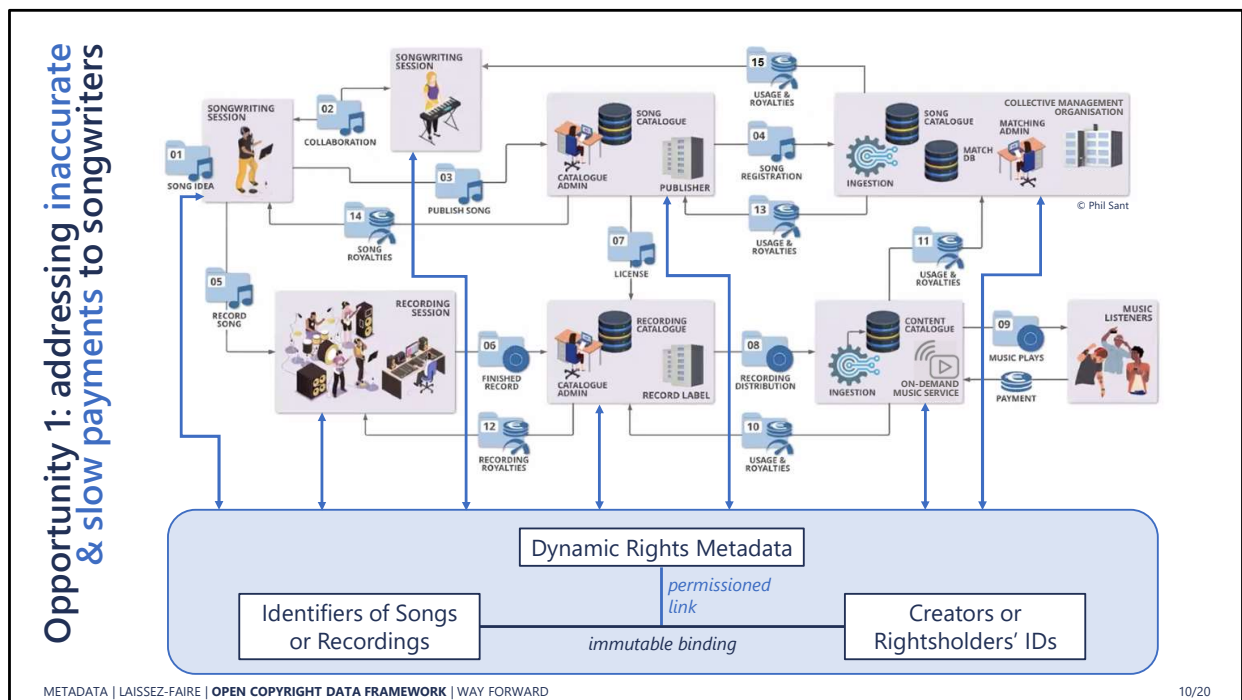
Distributed Network of Rights Management Information  
based on an Open Copyright Data Framework

METADATA | LAISSEZ-FAIRE | OPEN COPYRIGHT DATA FRAMEWORK | WAY FORWARD

9/20

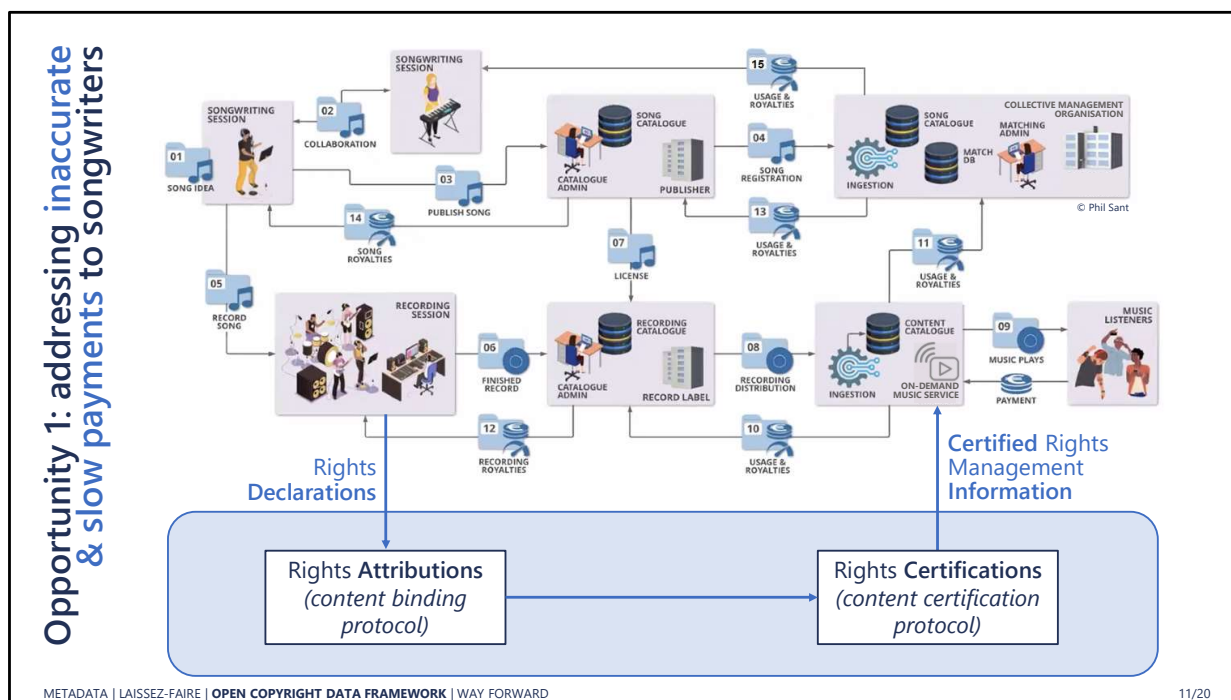
Imagine that we would have a distributed network of rights management information based on an incorruptible, trustworthy, and independent copyright data framework. What could happen?

A look at music streaming. On this diagram, music flows red from the left to the right, from the songwriters' brains to the listeners' ears. Remuneration should flow green from the right to the left, from the subscribers' pockets, or advertisers' budgets, to the artists' bank accounts. At each step from the left to the right, music – that is data – carries rights management information – that is metadata. At each step from the right to the left, metadata carries money. These data flows are complex – much more complex than this diagram suggests. Accordingly, they are slow, costly, and prone to inaccuracies.



In the case of music streaming, one could use a content binding protocol to bind immutably content-dependent identifiers of songs or recordings with identities of songwriters, performers or rightsholders. One could do it as close as possible to the moment of creation. One could attach, through permitted links, sets of dynamic rights metadata, which could be securely augmented, step by step, when the song or its recording would move from the left to the right.

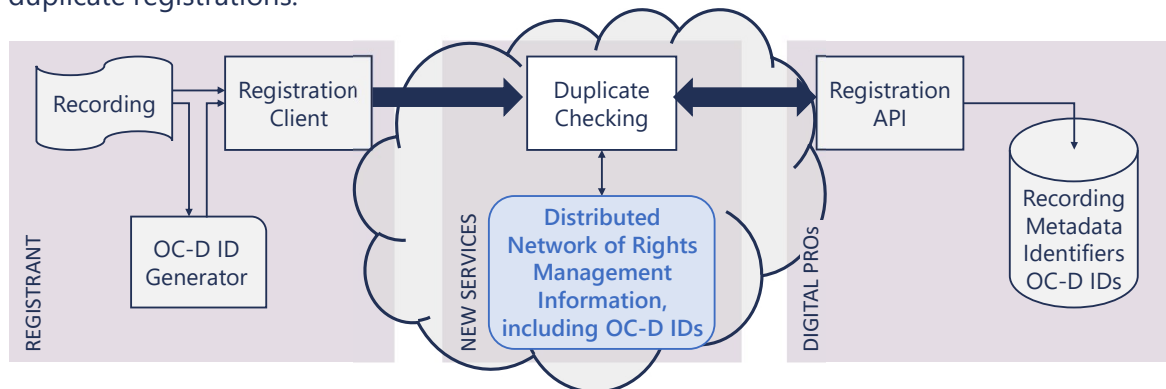
An open copyright data framework would allow a service bus to replace the monodirectional red and green arrows by bidirectional blue connectors. What engineers call a “bus” provides interaction services for software applications via an event-driven and standards-based messaging engine. It is built with middleware technologies. It is geared to isolate the links between a service and a data transport channel. The MovieLabs Digital Distribution Framework in the film industry, the Digital Data Exchange and Cis-Net in the music industry, and the ONIX standards in the publishing industry are in a sense all examples of middleware. A content binding protocol would strengthen them.



An open data framework would fulfil a function of public interest. Creators could declare their rights. These rights could be attributed and certified. When queried, the distributed network could provide certified rights management information. That would be its remit. That network would not license, distribute, meter, or pay. A multitude of competing not-for-profit or commercial organisations would be able to rely on its certified rights metadata to license, distribute, meter or pay because they could trust it.

## Opportunity 2: addressing the prohibitive costs of duplication

Open Content-Dependent Identifiers (OC-D IDs) can be used to check a registration against the **Distributed Network of Rights Management Information**, flag, and de-duplicate duplicate registrations.



The pre-screening allows deduplication against multiple data sources. It can be used to diffuse identities created at PROs, e.g., ISRCs or local codes.

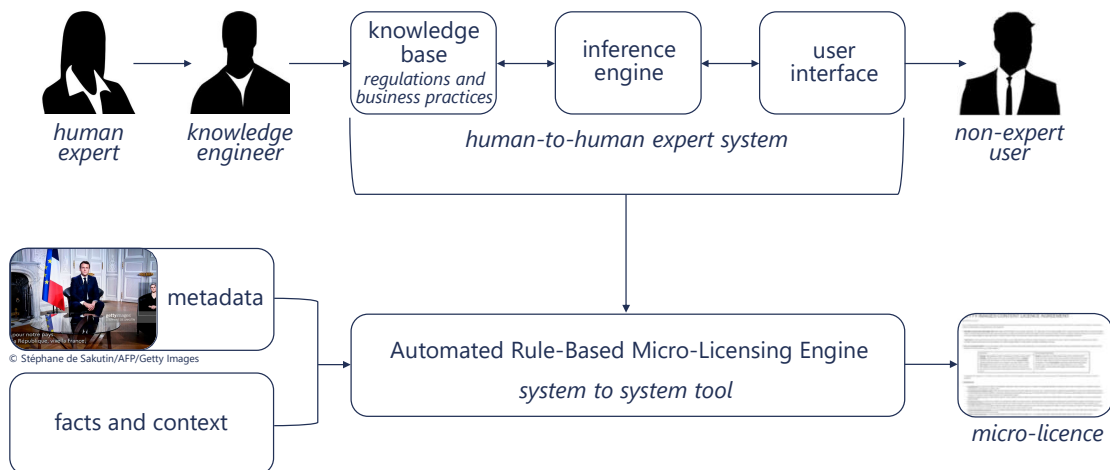
METADATA | LAISSEZ-FAIRE | OPEN COPYRIGHT DATA FRAMEWORK | WAY FORWARD

12/20

Duplication of rights management information within or across databases is a source of errors. On a distributed network based on an open copyright data framework, services could emerge to prevent duplications, or to deduplicate or match existing records.

### Opportunity 3: addressing inefficient markets through micro-licensing

Micro-licensing can help press publishers exploit their newly granted neighbouring right.



METADATA | LAISSEZ-FAIRE | OPEN COPYRIGHT DATA FRAMEWORK | WAY FORWARD

13/20

In the case of this specific image of President Macron, everything is already on the Internet: content, rights metadata, marketplace, and licence agreement. Humans can deal with this licensing procedure, but only from time to time. A fair, real-time, and transparent exploitation of article 15 would require much more powerful instruments.

Automated rule-based micro-licensing can help.

Therefore, one needs expert systems. Briefly, a knowledge engineer interviews a human expert and – assisted by machine learning – builds a knowledge base, containing copyright regulations and business practices of creative industries. This knowledge base is interpreted by an inference engine, that communicates with a non-expert user through an appropriate user interface. This is a human-to-human system.

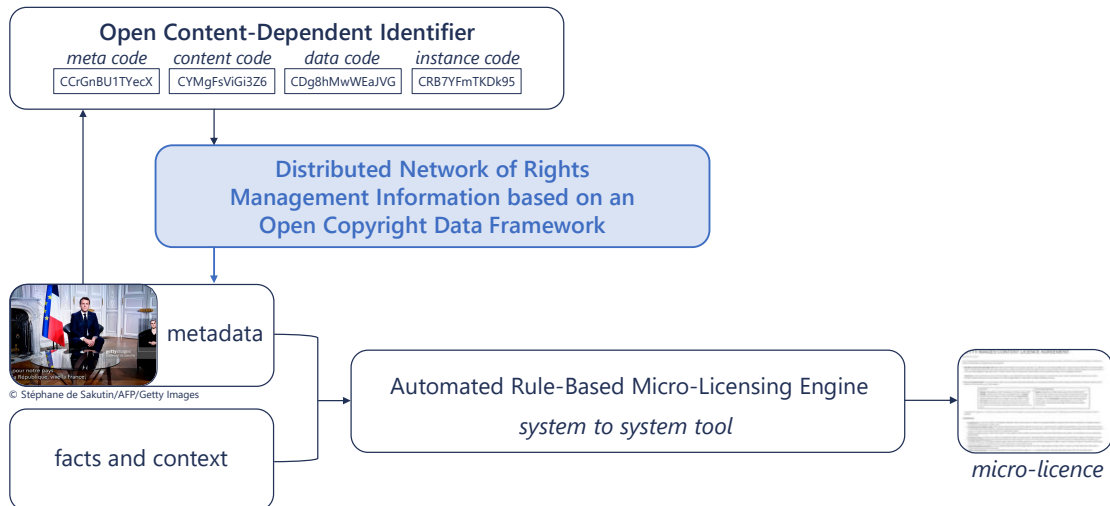
Once tested and stable, this expert system is compiled into an automated rule-based micro-licensing engine. That is a system-to-system tool.

On one side, a system inputs the selected image, its metadata, and facts and context around the required licensing.

On the other side, the micro-licensing engine produces a machine-and-human-readable micro-licence. This licensing process can be simple, accurate, fast, transparent, and affordable. At least, if we have the necessary metadata.

### Opportunity 3: addressing inefficient markets through micro-licensing

Micro-licensing can help press publishers exploit their newly granted neighbouring right.



METADATA | LAISSEZ-FAIRE | OPEN COPYRIGHT DATA FRAMEWORK | WAY FORWARD

14/20

So, what if metadata is not attached to President Macron's image, and there is an open copyright data framework?

We can analyse the DNA of the image, typically by using an open content-based identification tool such as the International Standard Content Code. It will produce a series of cryptographic hashes – from abstract and persistent characteristics on the left to concrete and volatile attributes on the right.

With that DNA, we can then query the distributed network to retrieve certified rights metadata concerning the image, relying typically on content binding and certification protocols.

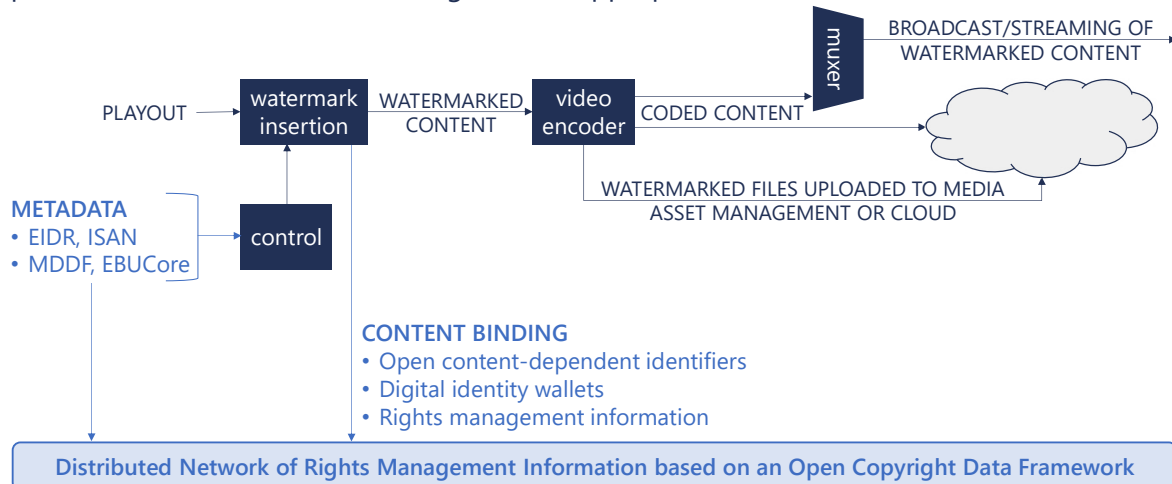
Then, we can reconcile image and metadata.

And finally, fire the micro-licensing engine.

Traditional media companies estimate that they have lost billions of Euros in advertising revenues to the online platforms. One would need only a fraction of that to build the systems I just outlined. Then, the press publishers could insist on transparency, leverage their knowledge of consumption data, and restore a level-playing field to price their licences at an appropriate level.

## Opportunity 4: addressing misappropriation of TV programmes

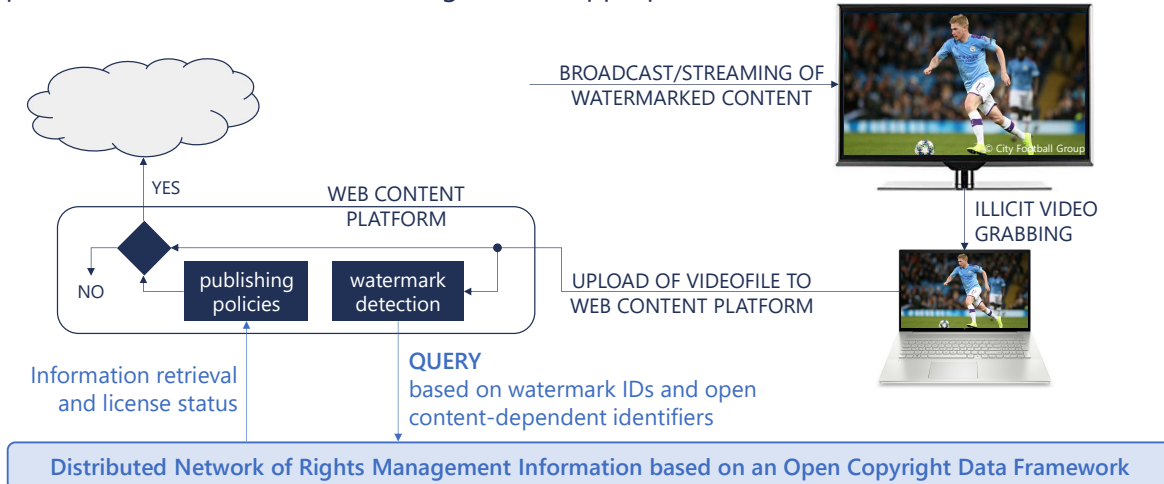
The combination of digital watermarks and **open content-dependent identifiers** can help protect content and TV channels against misappropriation.



I could continue with more opportunities arising from the existence of an open copyright data framework. Again, I suggest opening up the existing framework, not building a new one. Opening up means strengthening it, making it more trustworthy, and independent.

## Opportunity 4: addressing misappropriation of TV programmes

The combination of digital watermarks and **open content-dependent identifiers** can help protect content and TV channels against misappropriation.



METADATA | LAISSEZ-FAIRE | OPEN COPYRIGHT DATA FRAMEWORK | WAY FORWARD

16/20

During the interviews conducted for this study, I had the opportunity to discuss how such an open copyright data framework could dramatically enhance the current watermarking systems used to detect illicit video grabbing, and enforce the rights of commercial broadcasters. It does not require a big stretch of imagination to expand this to the application of article 17. You will tell me that large online platforms and a few commercial service providers are already building solutions. I can only repeat that rights data management should be neutral to business models. An open framework creates a space for many solutions. It must, and can, benefit everybody in the content ecosystem, also the online platforms and the commercial service providers.



## Characteristics of an open copyright data framework

- Compatible with the Berne Convention as the **rights declarations are voluntary**.
- Supports multiple interoperable **sectoral or territorial policies**.
- Contributes to the development of a genuine **single market for data** in compliance with the **GDPR** and respect of **business confidentiality**.
- Applies the **Once-Only Principle** to reduce administrative burden by asking to provide some information only once.
- Helps clarify and assign **responsibilities** and **liabilities**.

A few words on the characteristics of an open copyright data framework.

It is compatible with the Berne Convention as the rights declarations remain voluntary.

It can support a series of interoperable sectoral or territorial policies. It does not need one governance that fits all.

It can contribute to the development of a genuine single market for data, in compliance with the GDPR, and respect of business confidentiality.

It would apply the proven *Once-Only Principle* to reduce administrative burdens by asking individuals and organisations to provide standard information only once.

It can help clarify and assign responsibilities and liabilities. This means not only help realising the potential of the *acquis communautaire* but also paving the way for the implementation of ongoing regulatory initiatives such as the digital services act, digital markets act, or data governance act.

## Adoption of an open copyright data framework

- Adopted if it effectively supports applications that address the pain points of **all stakeholders** more efficiently.
- Requires rights **awareness** and **understanding** of metadata flows and purposes to gain traction.
- Should be fast-tracked through a close **cooperation** with the stewards of the existing copyright data framework.

Then, a few words on the adoption of an open copyright data framework.

It will be adopted if, and only if, it effectively supports applications – providing them with the right information at the right time, so much so that these applications can address industry pain points more efficiently – better, faster, and at a lesser expense.

The open copyright data framework will require rights awareness and understanding of metadata flows and purposes to gain traction.

The opening can be fast-tracked through a close cooperation with the stewards of the existing copyright framework; most of them are not-for-profit organisations maintaining, developing, and promoting foundational standards and technologies.

## Emergence of an open copyright data framework

- Copyright-related legislation is a State prerogative.
- The EU can foster the emergence of an open framework through *soft regulation*; i.e., co-regulations, indirectly referenced voluntary technical standards, recommendations, open method of coordination, education, information, and economic instruments.
- A *soft* approach seems *sufficient*. It appears *justified* in view of the digital and media agendas of the European Union.
- There are various *EU funding mechanisms* to support the different instruments of a *soft* regulation.

Finally, a few words on the emergence of an open copyright data framework.

Copyright legislation is a State prerogative.

The European Institutions can foster the opening of the existing data framework through *soft* regulation, namely a mix of co-regulations, indirectly referenced voluntary technical standards, recommendations, open method of coordination, education, information, and economic instruments. This is a very powerful toolbox whose use would respect the national prerogatives and not require any new legislation.

I personally think that this *soft* approach would be sufficient. To me, it appears also justified in view of the digital and media agendas of the European Union.

And, by the way, there are various EU funding mechanisms to support the different tools of such a *soft* regulation.

Can an open copyright data framework help release the digital potential of European creative industries?

What governance is needed to trust rights management information, and what could be the role of the Commission and public authorities?



[framework@rixhon.net](mailto:framework@rixhon.net), by Friday 2 July, thank you!

The European Institutions are wondering what can help release the digital potential of European creative industries. Do you think that an open copyright data framework could contribute?

Opening up the existing copyright data framework would require a governance, or a set of governances, endorsed by an overwhelming majority of stakeholders. Do you have any idea about that?

Thank you!