



UNIVERSITEIT VAN AMSTERDAM  
Instituut voor Informatierecht



ECS Conference 2024, Goethe University Frankfurt

# Metadata Mainstreaming in EU Copyright Law

Prof. Martin Senftleben, 24 May 2024

# Copyright data issues

## Insufficient metadata quality

missing

inaccurate

non-interoperable

## Missed opportunities

visibility, accessibility

licensing opportunities

recommender systems

text and data mining



**Lack of  
data?**





# Previous work at WIPO (CRR/GE/2/21/Report)

WIPO

## Survey on Voluntary Copyright Registration Systems

Final Report

prepared by

Dr. Stef van Gompel

Dr. Saule Massalina



Council of the  
European Union

# Priority at EU level

Brussels, 20 December 2019  
(OR. en)

15016/19

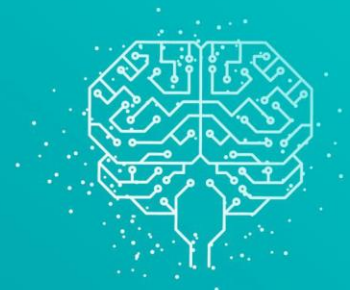
PI 156

## NOTE

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



February 2022



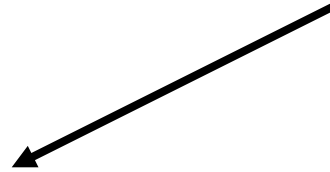
## Study on copyright and new technologies:

Copyright data management  
and artificial intelligence



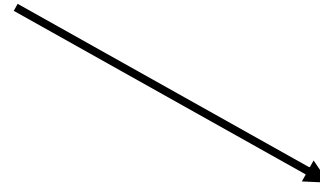
**Substantial  
research  
already  
carried out...**

# Central recommendations



## Raising awareness

= not only industry but  
also individual authors,  
performers **and users**



## Integrating initiatives

= establishing comprehensive  
infrastructure for all types of  
works



# Additional regulatory support? (JIPITEC 2022, 67)

Ensuring the Visibility and Accessibility of European Creative Content on the World Market

## Ensuring the Visibility and Accessibility of European Creative Content on the World Market

The Need for Copyright Data Improvement in the Light of New Technologies and the Opportunity Arising from Article 17 of the CDSM Directive

by **Martin Senftleben, Thomas Margoni, Daniel Antal, Balázs Bodó, Stef van Gompel, Christian Handke, Martin Kretschmer, Joost Poort, João Quintais and Sebastian Schwemer**



## Central recommendations



### **Raising awareness**

= not only industry but  
also individual authors,  
performers

### **Integrating initiatives**

= establishing comprehensive  
infrastructure for all types of  
works

### **Metadata mainstreaming through law?**

= exploring potential contributions  
of the legal system

A photograph of the University of Amsterdam campus. In the background, a modern building with a dark facade and many windows is visible, along with a tall white tower that has the university's logo on top. In the foreground, there are brick buildings, green trees, and a street with green umbrellas and people. A white rectangular box is overlaid on the image, containing the title text.

# Towards Metadata Mainstreaming



## Article 17(4)(b) CDSMD

work notification for the purpose of content moderation

OCSSPs not liable if they demonstrate that they have:

‘made, in accordance with high industry standards of professional diligence, best efforts to ensure the unavailability of specific works and other subject matter for which the rightholders have provided the service providers with the relevant and necessary information;...’

rightholder: right of communication to the public (making available)

work

territory

## Article 4(3) CDSMD

opt-out mechanism with regard to text and data mining (AI training):

‘The exception or limitation provided for in paragraph 1 **shall apply on condition that** the use of works and other subject matter referred to in that paragraph **has not been expressly reserved** by their rightholders in an appropriate manner, such as **machine-readable means** in the case of content made publicly available online.’

**rightholder: reproduction/extraction**

**work?**

**territory?**



## Article 8(4) CDSMD

opt-out mechanism with regard to making available of out-of-commerce works by cultural heritage institutions:

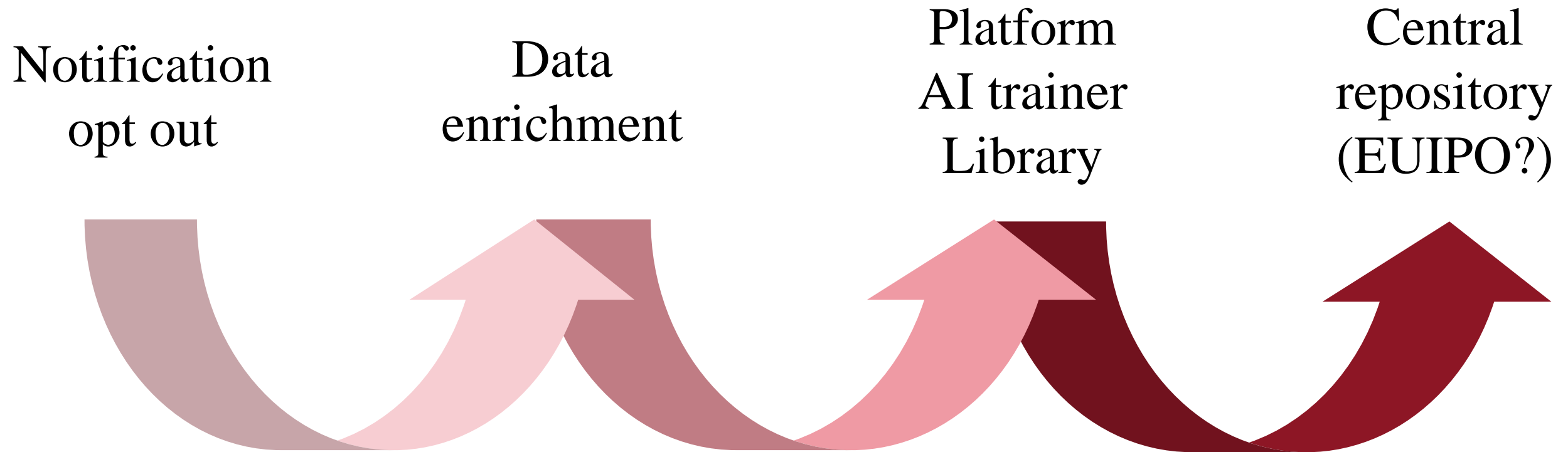
‘...that all **rightholders may, at any time, easily and effectively, exclude their works** or other subject matter from the licensing mechanism set out in paragraph 1 or from the application of the exception or limitation provided for in paragraph 2, either in general or in specific cases...’

**rightholder: reproduction/communication to the public**

**work**

**territory**

# Implementation in practice



# Which information and which standard?

## Authorship, ownership

authors

performers

rightholders

type of rights

territories

## Description

title

genre

style

semantics

contents

# Regulatory intervention?



## Incentive scheme

= already exists in the  
form of notification/ opt-  
out possibility

## Extra data enrichment

= own interest because of  
additional visibility and  
licensing opportunities?

## Further measures?

= stakeholder dialogue, best practice  
guidelines, legislative intervention?





For further reading: <https://ssrn.com/abstract=4817796>

SSRN Copyright Data Improvement fo X

← → ↺

https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=4817796

📄 🔍 ⭐

🔒 👤 📌 ☰

## Copyright Data Improvement for AI Licensing – The Role of Content Moderation and Text and Data Mining Rules

19 Pages • Posted: 7 May 2024

[Martin Senfleben](#)  
Institute for Information Law (IViR), University of Amsterdam; University of Amsterdam

Date Written: May 4, 2024

### Abstract

To enable European authors, performers and creative industries to benefit from licensing opportunities in the field of new technologies, such as AI training, it is important to establish a comprehensive **metadata** infrastructure that ensures the visibility and accessibility of European work repertoires in digital and algorithmic environments, and facilitates rights clearance. Recognising the need for **metadata** improvement, various European initiatives aim to increase awareness among artists and rightholders, and build bridges between existing **metadata** collections and infrastructures. One central factor in the equation, however, has remained underexplored to this day: copyright norms may serve as legal vehicles to encourage rightholders to constantly provide updated **metadata** in standardised form. By strategically using copyright provisions as

### Paper statistics

DOWNLOADS	ABSTRACT VIEWS	RANK
228	464	248,888

33 [References](#)

### PlumX Metrics



### Related eJournals

Information Systems Legislation & Regulations eJournal





UNIVERSITEIT VAN AMSTERDAM  
Instituut voor Informatierecht



The end. Thank you!

**Questions? Comments?**

Contact: [m.r.f.senftleben@uva.nl](mailto:m.r.f.senftleben@uva.nl)

Prof. Martin Senftleben, 24 May 2024