

EDRLab 2020

Creating a TDM Reservation Protocol

Text and Data Mining

- ❖ “ ‘text and data mining’ means any automated analytical technique aimed at analysing text and data in digital form in order to **generate information** which includes but is not limited to patterns, trends and correlations”

EU Copyright Directive

- ❖ **Copyright exception for scientific research** by research organizations and institutions (*across the Union*).
- ❖ **Exception** or limitation to the rights of (*EU?*) rightsholders **on lawfully accessible content** (*stored in EU?*), for reproductions and extractions for the purposes of TDM.
- ❖ BUT ... on the condition that the use of such content for TDM has not been **expressly reserved** by their rights holders in an appropriate manner, **such as machine-readable means** in the case of content made publicly available online.
- ❖ = OPT-OUT by rights holders from the copyright exception

Deadline

- ❖ Member States have until June 2021 to implement into national law the provisions of the EU Directive.

Project

- ❖ **Be open**, try to join together the publishing and TDM sectors and find a simple solution to this simple problem.
- ❖ The Publishing sector started by expressing:
 - ❖ Requirements (what the solution should offer)
 - ❖ Use cases (examples of usage of the solution)
- ❖ Then TDM sector + tech arm of publishing sector find a non-binding solution to recommend to publishers of all sorts.

Requirements

- ❖ The solution shall:
 - ❖ Specify how a rightsholder can indicate that TDM rights are reserved on a **Web page** he controls.
 - ❖ Specify how a rightsholder can indicate that TDM rights are reserved on **each individual Web resource** he controls, at least when such Web resources are present **in a Web page**.
 - ❖ Specify how the rightsholder can **indicate whether a licence for TDM rights is available** (in case TDM rights are reserved).
 - ❖ Specify **how a TDM agent can apply** for a TDM license.
 - ❖ Specify a **machine-readable format** for TDM licenses.

Use cases

- ❖ Use Case 1: a newspaper provides information with TDM opt-out
- ❖ Use Case 2: a publisher provides free extracts of ebooks with TDM opt-out
- ❖ Use Case 3: an academic publishing platform provides HTML and PDF articles with TDM opt-out and possibility to get a license
- ❖ Use Case 4: a subscription service provides EPUB files with TDM opt-out

Existing initiatives - toolbox

- ❖ ACAP (flexible usage rights, based on robots.txt, no law to enforce if, refused by web search engines)
- ❖ RightsML
- ❖ RightFind® XML for Mining Solution (Copyright Clearance Center)
- ❖ Assert your rights in online content (EPC / Copyright Hub)
- ❖ Article Sharing Framework (STM Association)
- ❖ Crossref TDM service

Now?

- ❖ Reach a broad support from the EU publishing sector (around the FEP)
 - ❖ Work document available on demand.
- ❖ Have the international TDM sector onboard (as much as possible)
- ❖ Design a simple technical solution, to be taken into account during implementations into EU national laws + recommended to all EU publishers.