

TDM Reservation Protocol

Working Pad

Initial Summary of the issue	0
Practical examples	1
Free extracts on a bookseller's website	1
Newspaper and news agency content	1
Academic content	2
Conveying TDM opt-out information between publishers and e-distributors	3
Useful Reports	3

Initial Summary of the issue

In addition to their significance in the context of scientific research, text and data mining techniques (TDM) are widely used both by private and public entities to analyse large amounts of data in different areas of life and for various purposes, including for government services, complex business decisions and the development of new applications or technologies.

The <u>EU Directive 2019/790</u>, better known as the EU Copyright Directive, specifies in its Article 4 an exception or limitation to the rights of rightsholders on lawfully accessible content, for reproductions and extractions for the purposes of TDM. This exception or limitation applies 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.

Note 1: this exception leaves intact the mandatory exception for text and data mining for the purposes of scientific research by research organizations and institutions provided by Art 3 of the EU Copyright Directive, where rights cannot be reserved. It therefore concerns TDM for any purpose other than scientific research, including commercial purposes.

Note 2: the TDM exception applies to TDM usage in the European Union in relation to content from European and foreign right-holders. Outside of the EU, where the EU Copyright legislation does not apply, the said exception does not apply: exclusive rights of right-holders to authorise acts of reproduction are maintained. In such cases, no TDM can be performed without the explicit authorisation of these rightsholders.

Note 3: there is a risk of a large fragmentation of national implementations of the Directive: a French report (by CSPLA) already recommends a "firewall" approach, using a product named ORRC developed by the CFC; This technical solution can manage access rights to publishers' sites by TDM robots, for a fee.

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

Members of the European Publishing Industry are therefore seeking collaboration with W3C and BSA members involved in the development of TDM platforms or services, for the specification of a simple and practical technical solution capable of expressing the rules set by the new EU Copyright Directive - Art.4. This non-binding solution would then be recommended to publishers in Europe.

This document aims at sharing useful information between members of the publishing industry, who are not technical experts, and engineering teams involved in TDM, who are not legal experts.

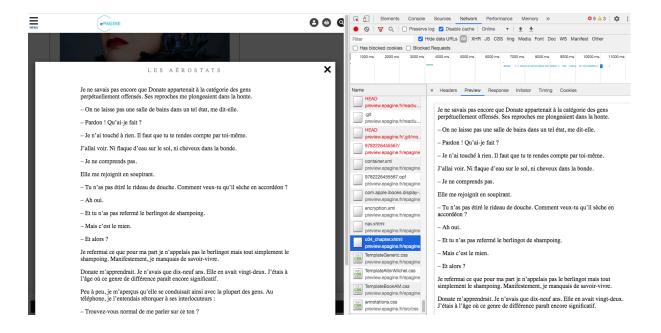
Practical examples

These real life examples are presented here to highlight situations where a technical solution for TDM opt-out is requested by content providers.

Free extracts on a bookseller's website

ePagine provides free extracts for the ebooks it sells. The html content of the extracts can easily be fetched by TDM crawlers.

https://www.epagine.fr/ebook/9782226455567-les-aerostats-amelie-nothomb/



Newspaper and news agency content

francetvInfo content is freely available. Some of its content is provided by the Agenge France Presse.

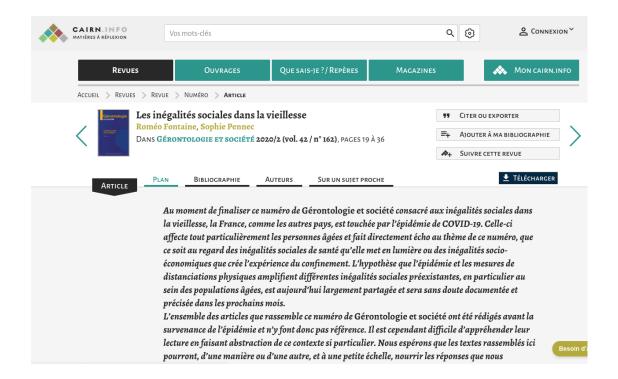
https://www.francetvinfo.fr/sante/maladie/coronavirus/infographies-covid-19-decouvrez-quelest-le-taux-d-incidence-dans-votre-commune 4151943.html



Academic content

Many Cairn.info articles are freely available.

https://www.cairn.info/revue-gerontologie-et-societe-2020-2-page-19.htm



Conveying TDM opt-out information between publishers and e-distributors

Deciding that TDM is reserved is a prerogative of content rightsholders, i.e. publishers. Therefore an additional issue for the publishing industry, which does not involve TDM agents, is to define how TDM information will be passed through publishing workflows.

ONIX metadata can be a great tool for conveiging such information between parties, when used.

Textual information contained in a contract is still a proper alternative to machine-readable information.

Useful Reports

France: report from the CSPLA about TDM exceptions, jan 2021.