



EUROPE'S TDM EXCEPTION FOR RESEARCH: *Will it be undermined by technical blocking from publishers?*

A mandatory exception for Text and Data Mining (TDM) for research, included in Europe's new Directive on Copyright in the Digital Single Market (DSM), risks being undermined by the Technical Protection Measures (TPMs) of publishers.

Submissions to a survey on content blocking — carried out by LIBER's Copyright & Legal Matters Working Group and the Libraries and Archives Copyright Alliance (LACA) — illustrate the barriers faced by researchers trying to perform TDM.

Key Takeaways

Submissions to the survey showed how:

- **Researchers are blocked from accessing many types of content.** Journal articles were the most common type of content mentioned (44%). Websites, eBooks, databases and newspapers were also cited.
- **Content blocking takes, on average, nearly a month to resolve.** Respondents said it took between 24 hours and 2-1/2 months to resolve the content blocking issue. The mean time was 24 days. One fifth were only partially able to resolve the issue and 11% said it was never resolved.
- **Sanctions impact whole communities, not just individual researchers.** Actions taken by publishers included 1) suspension of campus-wide access to paid for electronic subscriptions 2) threats to cut off access to content unless TDM was stopped 3) technically limiting downloads to one document only 4) a request for additional payments and 5) the introduction of CAPTCHA technology to frustrate TDM.

Research At Risk

LIBER is concerned about the potential for TPMs to impede research. We feel lock-outs are too frequent, take too long to resolve and — even when a single case is addressed — leave access for universities and their users uncertain. As one university noted in our survey:

The service has been recovered but it is still unclear how many documents are allowed to be displayed in a search. The licence still does not allow text and data mining. More and more, text and data mining is becoming an important activity for researchers. It should be guaranteed that TDM can be carried out normally and without it affecting the whole university and its other users.

Lock-Outs Must Be Resolved Quickly

In the transposition of the DSM Directive, **LIBER calls on all Member States to ensure that lock-outs will be resolved within a maximum period of 72 hours once reported to the appropriate governmental body.**



It is unacceptable that a university paying potentially hundreds of thousands of euros in annual licence fees for access to a single electronic resource risks waiting months to be reconnected, while ordinary consumers who are locked out of digital content services (e.g. iTunes, Netflix etc) are normally reconnected in hours.

If publishers and governments do not ensure that TDM for research purposes can be carried out easily and fully, they risk an escalation of tactics similar to an arms race: driving researchers to other means of accessing content. Some of our survey respondents said they reacted to TDM blocking by using SciHub to mine for content or by using software to circumvent CAPTCHA technology.

Whilst we understand the need to verify extremely atypical IP address behaviours, the relationship between universities and publishers could also be negatively impacted if publishers continue to insist on time-consuming actions before reinstating access. In our survey, the following examples of publisher conditions to reinstate access were reported:

- **Insisting that the university track down the individual who triggered the suspension to inform them of the publisher's licence terms**, which only permitted an 'insubstantial' amount of content to be downloaded (rendering TDM impossible);
- **Arranging a call between the university, publisher and the consortia who undertook the licensing;**
- **Insisting that the library publish warning messages on the university library website saying the publisher does not allow TDM**, and demanding that technical 'brakes' be installed on the university's reverse proxy to try to stop TDM.

Share Your Experience

If you or your organisation have been blocked from accessing a publisher's servers for reasons you believe are related to TDM, **fill out the survey**. It can be answered anonymously and will remain open indefinitely.

Knowing your experience will help us push for swift and transparent mechanisms to resolve blocking that prevents TDM, during consultations around the implementation of the DSM Directive. In line with Article 7 of the Directive, existing measures to allow circumvention of TPMs need to be significantly improved. Blocking by publishers should be resolved in a matter of hours. If not, upon being reported to government, access should be resumed within a **maximum period of 72 hours**.

We will publish future survey results on the LIBER and LACA websites, and share them with subscribers to LIBER's **Copyright & Legal Matters Mailing List**. In this way, everyone advocating for library and research friendly copyright legislation in Europe can access examples of what is occurring so that the situation can be improved.

Data Collected

Countries Represented: Austria, France, Germany, Spain, United Kingdom, United States.

Methods of Blocking: IP blocking (50%); CAPTCHA introduced (12.5%); database access cut off (12.5%); downloading limited to single files (12.5%); payment requested (12.5%).

Steps Taken to Resolve the Issue: Requested by publisher to tell users not to perform TDM (25%); required by publisher to install technical brakes (12.5%); email exchange with content owner (37.5%); use of a manual-bypass for Captcha (12.5%); SciHub (12.5%)

Time Required to Regain Access: 24 hours (22%); 1-2 weeks (56%); one month (11%); two months or more (11%).