

Position paper: Copyright Reform to Facilitate Research and Innovation

Text and data mining (TDM) is “any automated analytical technique aiming to analyse text and data in digital form in order to generate information such as patterns, trends and correlations.” There is huge potential for text and data mining—in terms of scientific advancement and discovery, civic engagement, and economic activity and innovation within the Digital Single Market.

The European Commission recognizes that researchers encounter legal uncertainty about whether—and how—they may engage in text and data mining, and are concerned that publishers’ contractual agreements may exclude TDM activities. In addition, the Commission observes that the optional nature of existing exceptions could negatively impact the functioning of the internal market.

To rectify this situation the Commission proposes changes to existing rules “to ensure that researchers can carry out text and data mining of content they have lawful access to in full legal certainty, including across borders.”

What is proposed in the Directive?

In the [Proposal for a Directive on Copyright in the Digital Single Market](#)¹, the European Commission proposes to introduce a mandatory exception for reproductions and extractions made by research organisations in order to carry out text and data mining of works to which they have lawful access for the purposes of scientific research. Those entities that will be considered “research organisations” are universities, research institutes, or any other organisation whose primary goal is to conduct scientific research on a not-for-profit basis or pursuant to a public interest mission. Contractual provisions that attempt to curtail activities available under the text and data mining exception will not be enforceable.

The problems with the proposal

1. Limitation of Beneficiaries

The proposed exception would be available only to research organisations that operate on a not-for-profit basis or pursuant to a public interest mission as recognised by a Member State. The practical effect of this limitation means that the private sector will be excluded from the benefits of the exception. In addition, the proposal restricts the ability to undertake TDM for important stakeholder groups such as journalists, citizen scientists, social enterprises, civil society organisations and cultural heritage organisations, all of whom stand to benefit from automated data analysis.

¹ <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016PC0593>

The Commission's proposal to create a TDM exception available only to non-profit research organisations will create a situation where text and data mining outside of the academic sector would be limited to data sources that are available for licensing. By extension, it is unclear how companies like data startups would be able to operate if they wish to conduct text and data mining on a corpus that is not available for licensing—for example, the internet (or large subsets thereof). The Commission's proposal ignores this significant challenge, which further restricts Europe's competitive position in relation to research and scientific discovery.

2. Limitation of Approved Purposes

The proposal limits the scope of the TDM activity to “purposes of scientific research.” In addition to the limitation on the beneficiaries of the exception described above, this constraint would decrease the potential impact of novel TDM uses, such as for journalism-related investigations, market research, or other types of activities not strictly considered “scientific research”. The term is not defined in the proposal.

Improper narrowing of TDM to protect publishers' revenue streams

The narrowing of the beneficiaries and approved purposes for TDM rests on the unproven assumption that there would be more harm to rightsholders who currently wish to issue TDM licenses to commercial users (primarily in the biotechnology and life sciences sectors) than good that would be created by broadening the exception to cover all users and purposes.

In its [impact assessment](#)² the Commission explicitly recognizes that under a broad exception where rightsholders would not be able to issue TDM-specific licenses, they would remain in a position to generate revenues from TDM by pricing it into existing access licenses, and by providing access to specific services that facilitate TDM. Even if we accept that this would still result in a small negative impact for rightsholders (due to costs of adapting their licensing practices), a broad exception is still desirable.

This is primarily because the negative impact on rightsholders can be expected to be offset by the positive impact for public interest stakeholder groups (journalists, civil society, cultural heritage institutions) and startups that would result from such an option.

Recommendations

In the light of the above we believe that the Directive needs to be amended to ensure that they achieve the goal of facilitating research and innovation across all parts of society:


- The TDM exception contained in article 3 should be amended **to allow anyone to undertake text and data mining**. This means removing the limitation on research organisations as the sole beneficiaries of the proposed exception.

² <https://ec.europa.eu/digital-single-market/en/news/impact-assessment-modernisation-eu-copyright-rules>

- The TDM exception contained in article 3 should be amended **to allow text and data mining for any purpose**. This means removing the limitation on scientific research as the only purpose allowed for under the proposed exception.

Want to know more?

If you would like to know more about how copyright reform can facilitate research and innovation, please contact COMMUNIA at info@communia-association.org or visit our website <http://www.communia-association.org/>.

 This publication is in the Public Domain.