



## UvA-DARE (Digital Academic Repository)

### Why a Copyright Exception Should Go Beyond TDM

Caspers, M.

**Publication date**

2016

**Document Version**

Final published version

**License**

CC BY

[Link to publication](#)

**Citation for published version (APA):**

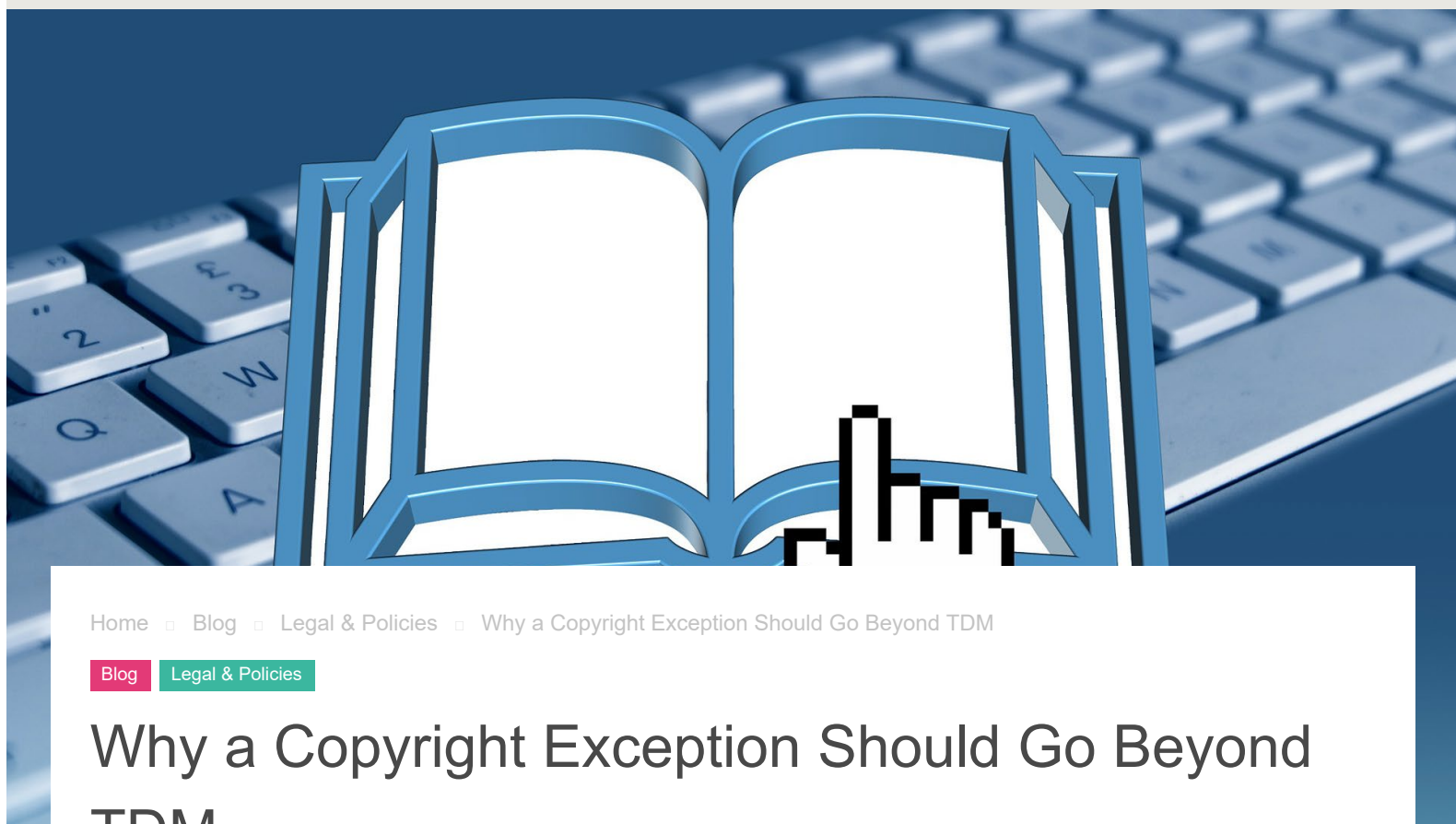
Caspers, M. (Author). (2016). Why a Copyright Exception Should Go Beyond TDM. Web publication or website, FutureTDM. <http://www.futuretdm.eu/blog/legal-policies/copyright-exception-tdm/>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



[Home](#) [Blog](#) [Legal & Policies](#) [Why a Copyright Exception Should Go Beyond TDM](#)

[Blog](#) [Legal & Policies](#)

# Why a Copyright Exception Should Go Beyond TDM

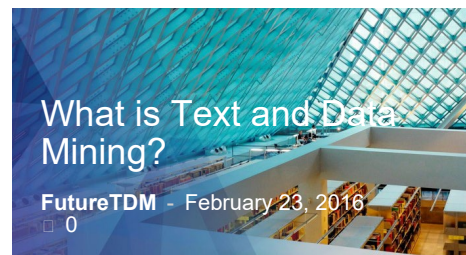
By **Marco Caspers** - May 2, 2016  2863  0

| [Share on Facebook](#)  | [Tweet on Twitter](#)

SEARCH

It is a basic principle of copyright law that ideas are not protected. What copyright does protect is the original expression of those ideas. It allows other people to use the information incorporated in a certain expression to re-use it and repackage it in an own, original way. It is this idea-expression dichotomy that seeks to prevent ideas from being monopolised by authors of works, while simultaneously promoting the free dissemination of knowledge through society. From this principle, one would expect that every act carried out with or on a copyright work

LATEST POSTS



of which the purpose is only to extract the knowledge underlying it, would not be relevant from a copyright perspective and therefore not fall under any exclusive right of the author.

As reasonable as that sounds, it is not the case. At least, not as a general rule. It is true that, generally, the mere act of reading, watching, or listening by a natural person does not infringe any copyright. For example, once you acquire a book, no one is entitled to prohibit you to read it - not under copyright law - regardless of the purpose you are reading it for. However, under the current copyright acquis communautaire, the same is not necessarily true for knowledge disseminated in digital form, such as the results of text and data mining (TDM) activities. And the reason is simple: for a computer to 'read' a work, the making of temporary copies - e.g. into the RAM - is indispensable to this activity. Moreover, in order for a user to retrieve the works prior to the analysis, it is often necessary to make reproductions thereof. For example, the miner may have to copy a text corpus from a server to its own terminal equipment to carry out the analysis. All said reproductions are covered by the reproduction right as provided under Article 2 of the Copyright Directive (2001/29/EC) and may therefore be prohibited by the rights owners.

Of course, this scenario looks an awful lot like a temporary reproduction that Article 5(1) of the same directive should exempt from the right owner's control. However, the criteria of this provision are either too restrictive or uncertain to be consistent with the principle that knowledge is not protected. Machine reading will generally not constitute "a transmission in a network between third parties by an intermediary" and it is rather debatable what must be understood by "to enable a lawful use". From Recital 33 of the Copyright Directive, it seems that the legislator's main concerns related to the lawful use of the work by the end user, in particular browsing and caching on the Internet.



However, in the case of machine reading as we refer to, there is no use of a work by the user, since the latter is in no stage confronted with the expression of work itself. Moreover, the requirement of transient copies not having any "independent economic significance" is irrelevant in the

context of the idea-expression dichotomy: ideas are not protected, regardless of the purpose for which they are used.

The European legislator seems to have done a better job as regards the idea-expression dichotomy in relation to the copyright protection of computer programs. Although computer programs are protected under copyright law, Article 5(3) of the Software Directive (2009/24/EC) permits the so-called 'black box analysis' of that software. It entitles the person who has the right to use the program to "observe, study or test the functioning of the program in order to determine the ideas and principles which underlie any element of the program if he does so while performing any of the acts of loading, displaying, running, transmitting or storing the program which he is entitled to do". Such line of reasoning could be extended to copyright works in general, where the lawful access to works should imply that reproductions are allowed to be made for the sole purpose of determining the ideas behind a work.

Some conditions may be necessary to ensure that the interests of right owners are not unreasonably affected. Fortunately, the Software Directive provides some inspiration on this point as well. The exception for the decompilation of software under Article 6 is coupled with some criteria, from which necessary conditions can be formulated.

Accordingly, TDM and any form of extraction of knowledge from works by machines should be exempted when:

- those acts are performed by the licensee or by another person having a right to use a copy of a work, or on their behalf by a person authorised to do so (cf. Article 6(1)(a));
- those acts are confined to the parts of the original work which are necessary in order to extract the knowledge, insofar as this can be determined (cf. Article 6(1)(c)); and if
- no original expression of the works is preserved in the output of those acts, unless this is permitted by law or authorisation by the author.

However, a work may not be:

- used for goals other than to extract the knowledge from the work (Article 6(2)(a));
- given to others, except when necessary to extract the knowledge from the work and other will not use the work for other purposes than described under the first point (Article 6(2)(b); and

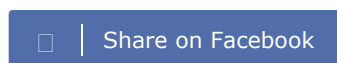
As we have learned from the landscape of optional exceptions under the Copyright Directive, which are implemented in various ways, a general TDM or ‘knowledge extraction’ exception should be made mandatory to ensure its effectiveness. For that purpose, it may also be necessary to overcome that users are prevented in other ways from extracting knowledge from works. For example, through the application of TPMs to the works, users may be technically excluded from acts that they are permitted to carry out under copyright law. Article 6(4) of the directive currently already obliges Member States to take appropriate measures to ensure that users may benefit from the enumerated copyright exceptions, in the absence of voluntary measures by copyright holders. Nonetheless, it appears from a survey conducted by FutureTDM among legal correspondents in different Member States that, according to national implementations, users must often go through administrative procedures in order to obtain necessary access. One way to ensure the benefits of the copyright exception as discussed in this blog post would be to only prohibit the distribution of means to circumvent TPMs, and not the act of circumventing itself, as under the TPM regime of the Software Directive (see Article 9(1)(c)).

Copyright law reform is not an often recurring activity in the European legislating process. New or revised provisions should therefore be well-thought-out and future proof. Correspondingly, a TDM exception should also take into account the future ways to extract knowledge from works in ways that are irrelevant from a copyright perspective. Merely focussing on TDM is only looking at the trends in the past years, but does not take into consideration similar activities in the future. The above proposed exception would do just that by providing the legislator an

instrument to streamline the copyright regime with its basic principle: the free use of ideas and knowledge. As radical as it may sound, all the ingredients are already in the European copyright framework.

***This post is authored by Marco Caspers (UvA) & Lucie Guibault (UvA).***

Image Credit: CC0 | [Pixabay](#)



Previous article

European Data Forum 2016

Next article

TDM for Scientific Publications:  
Venues and Ideas



**Marco Caspers**

<https://www.uva.nl/en/home>



NO COMMENTS

## LEAVE A REPLY

[Log in to leave a comment](#)

or login with:



The Future of Text and Data Mining

The platform is a result of the FutureTDM project, seeking to improve Text and Data Mining uptake in the EU.

[Project Website](#)



Horizon 2020  
European Union Funding  
for Research & Innovation

Horizon 2020 research and innovation programme under grant agreement No 665940.

[EC Europe](#)



The FutureTDM platform is in continuous improvement. Follow us to find out about our latest developments.

[Twitter Channel](#)



The [Open Knowledge Library](#) of this platform contains the main outcomes of the FutureTDM project licensed under a [Creative Commons Attribution 4.0 International License](#) (CC-BY) to be adapted with appropriate credit.