

Kluwer Copyright Blog

An EU exception for Text and Data Mining in the pipeline

Irini Stamatoudi (Hellenic Copyright Organization) · Wednesday, May 11th, 2016

Our era is the era of ‘Big Data’. The proliferation of data that we have experienced in recent years is unprecedented, and its volume only grows.

Some numbers: we create 2.5 quintillion bytes of data every day and 90% of the data in today’s world has been created in the last two years. The volume of data is doubling every three years (Please also see [Big Data at the Speed of Business](#), New York, IBM, 2014). Facebook users alone share over 30 million pieces of content per month, and Twitter has 350 million tweets per day. The global academic and research community generates over 1.5 million new scholarly articles annually; and there are an estimated 50 million academic articles in circulation since 2010.

Although researchers have access to more data, the growth in volume of data has outstripped their ability to access, read and analyse it effectively. The data has become too much to handle. It is in this context that text and data mining (TDM) has become important. In fact, it is considered to be a tool to facilitate research at all levels and bring to light new aspects of and insights into existing data by analysing, combining, isolating or synthesising information with the aid of an automated tool.

Discussion on TDM has been taking place in the EU for some time now and it is very likely that the legislative proposal (revising the InfoSoc Directive) that will be tabled in summer 2016 will also include an exception to this end. It seems that the proposal will be made on the assumption that TDM involves *restricted acts* in relation to *copyright protected works* and cannot be accommodated by existing exceptions or limitations to them.

With regard to the *restricted acts* TDM presupposes in many cases (depending on the tool used and the outcome intended) the processing, extraction and copying of the information mined.

With regard to the *protected works*, according to the CJEU’s case law (*Infopaq*) even an excerpt of 11 words may be protected to the extent that it constitutes its author’s own intellectual creation.

The exceptions that could be applicable in the case at issue are: a) temporary acts of reproduction (art. 5(1) InfoSoc Directive); b) scientific research; c) normal use of a database; and d) the extraction of insubstantial parts of a database. Yet, each one of them presents limitations that render it inappropriate to cover TDM.

For example, in relation to temporary copies one could argue that on most occasions permanent (rather than temporary) copies are made (at some stage of the TDM process) and therefore the

exception for temporary copies cannot apply. On top of this, TDM has an independent economic significance (though the copies relating to each step of it may not) and therefore constitutes a separate act of exploitation for which the author's authorisation is needed.

In relation to scientific research, not all types of TDM qualify as 'scientific' research; and even those that do will not fall within the relevant exception if they are for commercial purposes, i.e. direct or indirect economic or commercial advantage. In addition, it is not feasible in the circumstances to indicate the source and the author since there are usually vast numbers of works being processed during the mining process.

It seems that the new legislative text will take all these issues into account.

For more information on Text and Data Mining, please see I. Stamatoudi, "Text and Data Mining" in I. Stamatoudi (ed.), *New Developments in EU and International Copyright Law*, Kluwer Law International, 2016 or go to www.kluwerIPLaw.com.

To make sure you do not miss out on regular updates from the Kluwer Copyright Blog, please subscribe [here](#).

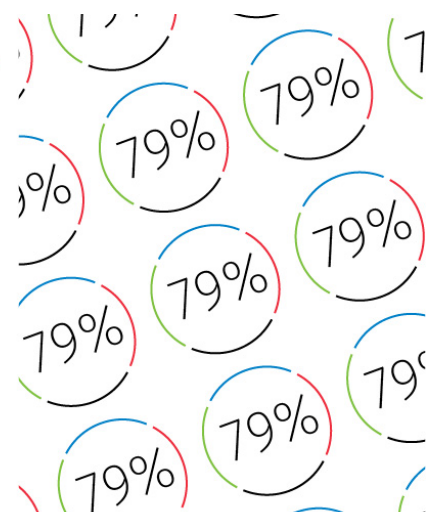
Kluwer IP Law

The **2022 Future Ready Lawyer survey** showed that 79% of lawyers think that the importance of legal technology will increase for next year. With Kluwer IP Law you can navigate the increasingly global practice of IP law with specialized, local and cross-border information and tools from every preferred location. Are you, as an IP professional, ready for the future?

Learn how **Kluwer IP Law** can support you.

79% of the lawyers think that the importance of legal technology will increase for next year.

Drive change with Kluwer IP Law.
The master resource for Intellectual Property rights and registration.



2022 SURVEY REPORT
The Wolters Kluwer Future Ready Lawyer
Leading change

This entry was posted on Wednesday, May 11th, 2016 at 9:59 am and is filed under [European Union, Exceptions and Limitations, Text and Data Mining \(TDM\)](#)

You can follow any responses to this entry through the [Comments \(RSS\)](#) feed. You can leave a response, or [trackback](#) from your own site.