Kluwer Copyright Blog

The UK government moves forward with a text and data mining exception for all purposes

Alina Trapova (University College London) and João Pedro Quintais (Institute for Information Law (IViR)) · Wednesday, August 24th, 2022

Photo by Kevin Ku on Unsplash

As **previously reported**, between October 2021 and January 2022 the UK Intellectual Property Office held a **public consultation** on the intersection between artificial intelligence (AI) and intellectual property laws (more specifically, copyright and patents). The outcome of the consultation is supposed to inform the government with respect to a potential legislative reform of the **UK Copyright Designs and Patents Act 1988 (CDPA)**. The consultation closed in the beginning of January 2022. On 28 June 2022, the government published both **the responses** by the 61 participants (among which trade associations in the publishing, technology, and music industries, academics, libraries, and cross-sector bodies) together with its **own position** as far as legislative changes and next steps are concerned.

The consultation looked at three specific areas:

- The computer-generated works (CGW) provision – **section 9(3) CDPA**
- The text and data mining (TDM) exception – **section 29A CDPA**
- Ownership for patent inventions

With respect to the CGW provision the government has decided to make no changes to the law. The consultation's ambition to collect evidence on the usefulness of section 9(3) was not successfully achieved. Considering the absence of evidence that the protection provided by this provision is harmful and in light of the early development stages of AI when it comes to creative output, no proper evaluation of the provision was possible. Thus, the government cautiously

decided to keep monitoring developments in the field and reserved its right to **"amend, replace or remove protection in future if the evidence supports it".**

The more interesting development relates to the TDM exception in section 29A CDPA. Presently in the UK, TDM is permitted with respect to the use of copyright protected works without the rightholders' permission *only for non-commercial scientific research*. This is now set to change.

The consultation outcome – rightholders vs users

A typical TDM process involves downloading works, pre-processing and formatting these works, analysing them, building new datasets and storing the source data for verification purposes. While some of these stages, such as the analysis stage, may arguably be covered by the temporary copies exception (**section 28A CDPA**), most of the other stages would typically involve an infringement the right of reproduction absent the permission of rightsholders.

In this respect, in order to carry out TDM for commercial purposes, which lie beyond the scope of the present UK exception, the consultation sought evidence on the available licensing models. However, no such quantitative evidence was provided by the respondents. Users reported mixed experiences with licensing. They **indicated** that at times they were "priced out of licensing for research projects which have restricted funding" and other times licensing was "costly and unworkable when mining content from large numbers of individual rightsholders." When it came to the introduction of a new TDM exception (or more accurately, expanding the scope of the existing one), unsurprisingly, rightholders were against, arguing that this would interfere with their exclusive right to exploit their works. Conversely, users were very much in favour of an exception as it would remove the need for licensing.

All-purpose TDM

Following the consultation, the UK government has now decided to introduce a new copyright and database right exception which allows TDM for any purpose, i.e. including commercial uses. Licensing will no longer be an issue and rightholders will not be able to opt-out or contract out of the exception. The government believes that this approach would significantly benefit a wide range of stakeholders – from researchers, AI developers, small businesses, through cultural heritage institutions, journalists, all the way to engaged citizens. Emphasis is placed also on research outcomes as a result of this reform which would serve the wider public interest, e.g. supporting research and innovation in public health. The creative sector is also expected to reap some benefits since TDM and AI techniques are regularly used to create new works, which would no longer require the unpacking of complex rightholder claims at the input stage. An overall advantage of this solution according to the UK government is that of easing the obtaining of permissions. Indeed, in certain fields even if one would be willing to seek a license to use a work, this will not always possible due to the potentially high fees as well as the high number of rightholders to deal with.

The narrower EU exceptions

It is especially interesting to contrast the UK proposed approach with current EU law. In the recent Copyright in the Digital Single Market Directive (**CDSMD**), TDM is defined as "any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations." Articles 3 and 4 then contain two TDM-related mandatory exceptions.

Article 3 provides an exception for acts of TDM for the purposes of scientific research – covering both natural and human sciences – by research organisations and cultural heritage institutions, regarding works/subject matter to which they have lawful access, and subject to a number of additional conditions. This provision has to interact with the optional exception covering uses for scientific research purposes in Article 5(3)(a) InfoSoc Directive, which applies to certain TDM activities.

Article 4 sets forth an exception for reproductions and extractions of lawfully accessed works/subject matter for the purposes of TDM. This is meant to add legal certainty for those acts that may not meet the conditions of the temporary and transient copy exception in Article 5(1) InfoSoc Directive. The new exception is subject to reservation by rightholders, including

through "machine-readable means in the case of content made publicly available online", for

instance through the use of metadata and terms and conditions of a website or a service. Such

reservation shall not affect the application of the TDM exception for scientific purposes in Article 3.

The table below provides a summary of the scope of the EU TDM exceptions.

Provision ?		
Scope & Conditions ?	**Article 3**	**Article 4**
Acts	Text and data mining	Reproductions and extractions
Rights	Database Dir.: articles 5(a) and 7(1) InfoSoc Dir.: article 2 CDSMD.: article 15(1)	Database Dir.: articles 5(a) and 7(1) InfoSoc Dir.: article 2 Software Dir.: articles 4(1)(a) and (b) CDSMD.: article 15(1)
Purpose of use	Scientific research	Text and data mining
Beneficiaries	Research organisations; Cultural heritage institutions	All (unrestricted)
Lawful access	Yes	Yes
Other conditions for beneficiary	Works must be stored with "appropriate level of security"	–
Contractual Derogation	No	Yes

Other Reservations or Exclusions	No	Yes: via “machine-readable means in the case of content made publicly available online”
Other Conditions	Three-step test; Partial protection of technical measures under 6(4) InfoSoc Dir.	Three-step test; Partial protection of technical measures under 6(4) InfoSoc Dir.
Connection to other exceptions in *acquis*	InfoSoc Dir.: article 5(3)(a)	InfoSoc Dir.: article 5(1)

As multiple critics have noted, both TDM exceptions are restrictive may exclude many important applications in this domain, especially as regards the development of AI technologies (see e.g. **here**, **here** and **here**).

Looking forward

Against this restrictive EU background, then, the UK’s proposal for an “all-purpose no opt-out TDM” exception is commendable. Indeed, such an exception would likely be superior to its restrictive EU counterparts in facilitating research and investment on new technologies in a post-Brexit world. Naturally, the devil is in the details and it is necessary to assess the actual wording of the new exception. In particular, should the exception be as broad as proposed, there will undoubtedly be questions from rightsholders representatives about compatibility with international copyright law. On this point, however, there are strong arguments in favour of a broad TDM exception, with some authors suggesting that there might not even be a “**need for a TDM exception for the act of extracting informational value from protected works**” and others stating that “**TDM falls outside the scope of international copyright harmonization altogether**” and therefore “**international copyright law does not limit the freedom of national policymakers to devise appropriate domestic solutions to reconcile copyright protection with the right to research that underlies TDM activities**”.

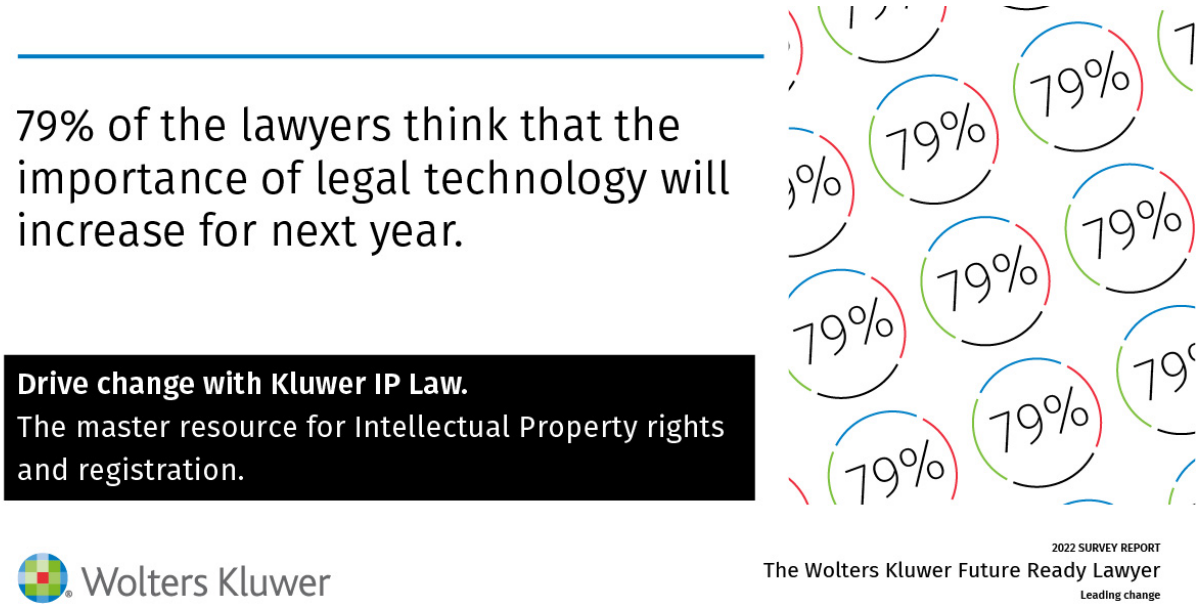
To make sure you do not miss out on regular updates from the Kluwer Copyright Blog, please subscribe here.

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.



This entry was posted on Wednesday, August 24th, 2022 at 11:35 am and is filed under Artificial Intelligence (AI), CDSM Directive, European Union, Exceptions and Limitations, Legislative process, Text and Data Mining (TDM), United Kingdom
You can follow any responses to this entry through the Comments (RSS) feed. You can leave a response, or trackback from your own site.