



To: Alcallforviews@ipo.gov.uk

From: [REDACTED]

Section A: Questions Arising from this Consultation

Copyright Clearance Center (CCC) welcomes the opportunity to respond to questions raised in the UK Consultation on Artificial Intelligence and Intellectual Property.

About CCC:

CCC is the primary collective management organization (CMO) for text in the United States. We offer global corporate licensing on a fully voluntary, non-exclusive basis, plus academic licensing services primarily within the US. Our mission is to advance copyright, accelerate knowledge, and power innovation. CCC manages copyright licensing and content delivery on behalf of -- among others -- (1) publishers who manage copyrights themselves and (2) non-publishers (such as pharmaceutical companies) who manage both their own copyrights and third-party materials. In addition, CCC services include copyright education, library staffing, library and publisher software development, APIs, and data/metadata services.

With more than forty years of expertise in copyright and information management, CCC designs and delivers innovative information solutions that power decision-making by helping people integrate and navigate data sources and content assets. We collaborate with our customers to accelerate discovery and progress by shortening the distance between data, information, and insight.

CCC is a founding member of the International Federation of Reproduction Rights Organizations (IFRRO) and works closely with our UK counterpart Copyright Licensing Agency (CLA).

We offer no opinions herein with respect to computer generated works or patents.

UK IPO Questions and CCC Responses:

Copyright – text and data mining (TDM)

6. If you license works for TDM, or purchase such licences, can you provide information on the costs and benefits of these? For example, availability, price-point, whether additional services are included or available, number and types of works covered by the licence etc.

CCC Response:

CCC's TDM Offering- Launched approximately seven years ago after three years of development and millions of US dollars in investment, CCC offers a globally available licence for TDM, aimed primarily at the needs of corporate researchers at companies of any size. We have found that the most

significant challenges around TDM are of a technical, not legal, nature. Text is unstructured data to a machine and requires significant investment to make it more usable.

CCC's TDM offering includes a copyright licence designed initially for life sciences and pharmaceutical research. In addition, we aggregate, store, tag, enrich and standardize publisher metadata, and make the content available for download in XML format for our clients. We have more than 50 participating publishers, offer more than 14 million articles available in full text for mining and more than 20 million abstracts. Customers can mine full text of subscribed, unsubscribed and open access content, with business rules based in part on the user's subscription holdings. Clients include UK based companies. Pricing scales upwards and downwards based on a variety of factors.

In terms of the UK market, our TDM licence compliments licences offered by (1) CLA, especially for mining of news content, and (2) individual publishers.

Open Access (OA) and Licensing- In considering what kind of licences support TDM, the IPO should also consider the ever-growing business of open access publishing as part of the current TDM infrastructure. "Open access (OA) refers to freely available, digital, online information. Open access scholarly literature is free of charge and often carries less restrictive copyright and licensing barriers than traditionally published works, for both the users and the authors."¹ According to a recent study, paid-for OA represents 7% of the value of market value of the journal publishing market.²

Open Access is a fast-growing part of the scholarly publishing ecosystem. CCC engages with open access by supporting more than 30 publishers in managing open access payment workflows with over 1000 institutions globally. We also help users access and discover OA articles -- including making these articles available for TDM under open licences -- while providing thought leadership about OA to the publishing and user communities.

The ability to text mine OA content is a key component of the value proposition offered by OA publishers. Indeed, many institutions and funding bodies make the ability to text mine a requirement of the payment of OA fees for articles written by their authors. Admittedly, the right to mine OA articles is one of many motivating factors which has driven growth of the OA market. Thus,

¹ <https://guides.library.cornell.edu/openaccess>

² https://www.stm-assoc.org/2021_10_19_STM_Global_Brief_2021_Economics_and_Market_Size.pdf at Page 12 ("According to Delta Think, in 2019, the paid Open Access market was worth just over \$760m, up 13% from 2018. They also estimate that it grew to around \$850m in 2020. Just over 30% of all scholarly articles are published as paid-for Open Access, accounting for just over 7% of the total journal publishing market value.")

while we do not believe it possible to separate the value of TDM rights from the other OA value propositions, we simply note that TDM rights are embedded in OA business models, OA content is available for TDM on a voluntary basis, and there is value in the TDM rights.

7. Is there a specific approach the government should adopt in relation to licensing?

CCC Response:

As noted in the *Consultation outcome: Government response to call for views on artificial intelligence and intellectual property* updated 23 March 2021 : “Most copyright owners felt that a voluntary licensing model would best balance their need to be remunerated for the use of their works against the need for access to copyright works by the AI sector.” This theme was echoed in the Government’s response: “Review the ways in which copyright owners license works for use with AI, and consult on measures to make this easier, including improved licensing or copyright exceptions, to support innovation and research.”

Legal reform takes time and is a blunt instrument for solving market challenges. Legal reform typically encounters resistance. Laws are also very hard to enact and harder to change once enacted. Markets have often evolved new solutions and challenges by the time new laws are in place. By contrast, licensing is able to meet client needs faster and better than legal reform. It is quick to adapt to changing market conditions and by its very nature must solve users’ challenges. As the operator of global, voluntary collective licenses in this sector we agree with the copyright owners. We also note that during the roundtable event held on 7 December 2021, the UK Government expressed concerns that had been raised by small and medium enterprises about license availability and cost. CCC and CLA’s TDM licenses are available to SMEs in the UK at fair cost; CLA has an appropriate and proven TDM license which is used by SMEs and larger businesses in the media monitoring sector. However, we recognize that more could be done to facilitate use of STM content and tools by SMEs and that we can start to do this work *without waiting* for the end of this consultation. This does not require any change in the law. CCC and CLA propose that working with the UK Government, publishers, users and other stakeholders, we consider creating a pilot for licensing SMEs to mine STM content. This pilot could run on the individual SME level or we could arrange a national license, funded by UK government, for all SMEs meeting an agreed profile. By creating a TDM license pilot in the UK that combines CLA and CCC’s existing capabilities, we could open TDM for a broader range of entities without delay and help make the UK the pre-eminent home for innovative SMEs/start-ups using TDM.

8. Please rank the options in order of preference (most to least preferred) and explain why.

CCC Response:

Option 1 Improve licensing environment for the purposes of TDM-

Option 1 is our first choice. While we believe UK law as drafted is “fit for purpose,” not every SME is a copyright expert or wishes to investigate all licensing options. As proposed above, the UK Government, working with the CLA, CCC, users and rightsholders, can advance the interests of the UK and make the UK an increasingly important hub of SME and start-up innovation.

Option 0 Make no legal change-

Option 0 is our second choice. UK law on point is already fit for purpose on TDM, drawing a distinction between commercial and non-commercial re-use. Moreover, as the UK rules have been in place since June 2014, businesses have grown up and adapted to their contours. A change would adversely impact upon pre-existing expectations without a matching benefit.

We note here that those who argue that the UK is falling behind in AI/TDM will often point to the US as a supposed haven for AI, TDM and technology start-ups, while ignoring the fact that *TDM is most typically licensed* in the US market. As UK law exists today, its TDM exception can be broader than US fair use – depending on the actual facts of the use as set forth below.

Under US law, copying copyrighted content to train AI states a cause of action for infringement.³ Whether “fair use” excuses the infringement is fact dependent. While the commercial/non-commercial distinction in UK law is relevant to a fair use inquiry, it is merely one of four factors, the most important of which is market harm. As stated by the US Supreme Court in *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 590 (1994) “[the fourth factor] requires courts to consider not only the extent of market harm caused by the particular actions of the alleged infringer, but also ‘whether unrestricted and widespread conduct of the sort engaged in by the defendant ... would result in a substantially adverse impact on the potential market’ for the original.”

Thus in the case of *Fox News Network, LLC v. TVEyes, Inc.*, 883 F.3d 169 (2d Cir. 2018), the Second Circuit Court of Appeals rejected a fair use defense in a case of transformative compiling of recorded broadcasts into text searchable database that allowed search and viewing of short excerpts. By contrast, the Second Circuit had previously considered the text mining of scanned books for non-commercial social science research in *Authors Guild v. Google*, 721 F.3d 132 (2d Circuit 2015) and held that copies made for that purpose would likely fall under fair use.

³ *Thomson Reuters Enter. Ctr. GmbH v. ROSS Intelligence Inc.*, C.A. No. 20-613-LPS (D. De. March 29, 2021) (downloading and copying of Westlaw database for the purpose of training AI) (slip opinion attached).

There are no US cases about copyright infringement in the context of AI or TDM of journals or newspapers for commercial or non-commercial purposes, and any blanket conclusions as to how a US court would rule with respect to fair use on copying for AI/TDM of news, journals, or born digital content is not only speculative, but misunderstands the fact-dependent investigation in the fair use defense.⁴ In the US, we have established TDM markets for, *inter alia*, scientific journals, news and other digital content. For those types of works, even copying for non-commercial TDM could, in certain circumstances, be infringing non-fair use in the US.

Option 3: Adopt a TDM exception for any use, with a rights holder opt-out-

Option three is our third choice. Although Option three tracks the Article 4 of the EU's Digital Single Market Copyright Directive (DSM), we have yet to see the full impact of Article 4 or the results of any legal challenges. DSM Article 4 may well exceed the permissible limits of the Berne Convention Article 5(2) by introducing a formality in the form of requiring copyright holders to opt-out in order to reserve rights.

DSM Article 4 may also in certain cases violate the 3-Step Test of Article 9, which states:

...(2) It shall be a matter for legislation in the countries of the Union to permit the reproduction of such works in certain special cases, **provided that such reproduction does not conflict with a normal exploitation** of the work and does not unreasonably prejudice the legitimate interests of the author.... (emphasis added).

For numerous publication sectors—including without limitation news and science as discussed herein—machine applications, AI and TDM licensing are today “normal exploitations,” and an exception would violate the Berne requirements. For example, hedge funds today license news content for algorithmic trading. An exception that transfers wealth from news publishers to hedge funds will do nothing to burnish the UK's reputation as a start-up capital because these licenses are already available and in use, and might bring the UK out of Berne compliance. Moreover, TDM and AI are advancing at a rapid pace. The right to copy materials for TDM and AI uses in other sectors may evolve in the near future to constitute “normal exploitations.”

Options 2 and 4: We do not believe these options to be advisable or viable.

9. If you have experience of the EU exception with opt out for rights holders, how has this affected you?

⁴ See, e.g., *Basic Books, Inc. v. Kinko's Graphics Corp.*, 758 F. Supp. 1522 (S.D.N.Y. 1991) (Noting that the same acts could be fair use for one party while not being fair use for another: “The use of the Kinko's packets, in the hands of the students, was no doubt educational. However, the use in the hands of Kinko's employees is commercial.”)

CCC Response:

CCC works with copyright owners ranging from individual authors to the largest publishers. Our larger rightsholders have told us that they are opting out of the EU exception. This leaves smaller creators with fewer lawyers and less knowledge to fend for themselves as their materials are used by large commercial entities. As such we expect the exception with opt-out to have the unintended consequence of favouring large copyright owners and users, while disadvantaging learned society publishers, publisher SMEs, and authors.

10. *How would any of the exception options positively or negatively affect you? Please quantify this if possible.*

CCC Response:

CCC has invested over 7 years and millions of dollars in aggregating, enriching, tagging and converting journal articles into ready-to-mine XML. An exception could have an adverse impact upon our ability to sell licenses as well as a chilling effect on further innovation by us (or our publishers) in this field. Our licenses generate royalties for our participating publishers who have themselves invested in the creation, publication and curation of content. An exception would likely have a negative impact upon them as well.

Section B: Respondent information:

A: Please give your name:

[REDACTED]

B: Are you responding as an individual, business or on behalf of an organisation?

We are responding on behalf of Copyright Clearance Center (CCC).

C: If you are responding on behalf of an organisation, please give a summary of who you represent:

A summary of CCC is set forth in response to Section A. More information can be found at www.copyright.com

D: If you are an individual, are you....?

N/A

E: If you are responding on behalf of an organisation, are you....?

We are a reproduction rights organization

F: If you are responding on behalf of a business or organisation, in which sector(s) do you operate? (choose all that apply)

While we are primarily involved in sector 11 (Information and communication – Publishing, audio-visual and broadcasting), we have clients in almost every sector listed below.

- 1) *Agriculture, forestry and fishing*
- 2) *Mining and quarrying*
- 3) *Manufacturing – Pharmaceutical products*
- 4) *Manufacturing – Computer, electronic and optical products*
- 5) *Manufacturing – Electrical equipment*
- 6) *Manufacturing – Transport equipment*
- 7) *Other manufacturing*
- 8) *Construction*
- 9) *Wholesale and retail trade; repair of motor vehicles and motorcycles*
- 10) *Transportation and storage*
- 11) *Information and communication – Publishing, audio-visual and broadcasting*
- 12) *Information and communication – Telecommunication*
- 13) *Information and communication – IT and another Information Services*
- 14) *Financial and insurance activities*
- 15) *Real estate activities*
- 16) *Scientific and technical activities*
- 17) *Legal activities*
- 18) *Administrative and support service activities*
- 19) *Public administration and defence*
- 20) *Education*
- 21) *Human health and social work activities*
- 22) *Arts, entertainment and recreation*
- 23) *Other activities – please specify*

G: How many people work for your business or organisation across the UK as a whole?
Please estimate if you are unsure.

10–49

H: The Intellectual Property Office may wish to contact you to discuss your response. Would you be happy to be contacted to discuss your response?

Yes, we would be happy to be contacted.

I: If you are happy to be contacted by the Intellectual Property Office, please provide a contact email address.



J: Would you like an acknowledgement of receipt of your response? Yes/No

Yes, we would like acknowledgement of receipt