

**Appeal for action on violations of the Berne Convention
by the application to copying of creative works for AI development
of the TDM exception in Articles 3 and 4 of the 2019 EU Directive on Copyright**

(July 2023)

As creators of copyrighted works – including text, illustrations, photographs, lyrics, music, audio, and video – our economic and moral rights have been gravely harmed by the copying and ingestion of our work, without permission or payment, to compile generative artificial intelligence (“AI”) language models – through the process disingenuously and anthropomorphically referred to by AI developers as “training” – that reuse our works forever after to "generate" derivative works as their output.

Much of the copying of our works for generative AI, including “scraping” of Web pages and compilation of “datasets” for use in generative AI, has been carried out from, and/or by entities in, the European Union, claiming to rely on the exceptions to copyright for “text and data mining” (TDM) in Articles 3 and 4 of the Directive on Copyright in the Digital Single Market (“DSM Directive”) enacted by the European Union in 2019.¹

But allowing these exceptions to be applied to copying for ingestion and reuse by generative AI systems constitutes a significant violation of the obligations of EU member states as parties to the Berne Convention² and the WIPO Copyright Treaty.

We urge the European Union to promptly cure this violation of the Berne Convention and provide effective redress for the violations which have already occurred.

And we urge the United States government to use all available means to bring the European Union into compliance with the Berne Convention, as incorporated in the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS), in connection with the application of Articles 3 and 4 to generative AI.

Article 3 of the DSM Directive creates an exception to copyright for “reproductions and extractions made by research organisations and cultural heritage institutions in order to carry out, for the purposes of scientific research, text and data mining of works or other subject matter to which they have lawful access.”

But this does not satisfy the three-part test in Article 9.2 of the Berne Convention, which allows for exceptions to the exclusive right of reproduction only (1) in certain special cases, (2) provided that such reproduction does not conflict with a normal exploitation of the work and (3)

1. Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC, <<https://eur-lex.europa.eu/eli/dir/2019/790/oj>>
2. Berne Convention for the Protection of Literary and Artistic Works, <<https://www.wipo.int/wipolex/en/treaties/textdetails/12214>>.

when this does not unreasonably prejudice the legitimate interests of the author.

Copying of our works for AI development pursuant to this exception directly conflicts with the normal exploitation of our work through licensing for AI usage and unreasonably prejudices our legitimate interests as authors, in violation of Article 9 of the Berne Convention – especially when, as in this case, copies and derivative works are being made by some “research organisations” for the commercial use and benefit of for-profit sponsors and affiliates of those nominally not-for-profit “front” entities.

Allowing AI developers to copy works to build software that generates unauthorized derivative works that compete with the original works has the potential to conflict with the normal exploitation of those original works. And those derivative works can, and already do, include works – fake news, fascist propaganda, phishing spam, defamation, etc. – prejudicial to the interests of the authors from whose original works they are derived.

Article 4 of the 2019 EU Directive on Copyright creates an even broader exception for “reproductions and extractions of lawfully accessible works and other subject matter for the purposes of text and data mining,” even for commercial use by for-profit entities.

Because Article 4 would, unless limited, violate the Berne Convention, it purports to apply “on condition that the use of works and other subject matter referred to in that paragraph has not been expressly reserved by their rightholders in an appropriate manner, such as machine-readable means in the case of content made publicly available online.”

It is essential to recognize that this purported condition is, as a practical matter, completely meaningless. It neither mitigates nor cures the violation of the Berne Convention. Even if an author or other rightsholder was aware of Article 4, and desired to reserve their TDM rights, it has, to date, been *entirely impossible* to do so. Article 4 allows commercial use of our work by for-profit entities, without permission or payment, with no actual possibility to opt out.

Neither the European Commission nor the EU Intellectual Property Office has issued guidance regarding any “appropriate manner” in which rightsholders can reserve their TDM rights. Nor has any standard or machine-readable means to do so been adopted.

An explicitly non-normative TDM Reservation Protocol (TDMrep) for Web content has been proposed to, but not adopted by, the World Wide Web Consortium (W3C).³ But even if this or a similar proposal were to be adopted as a standard, it would not suffice to bring Article 4, as applied to copying for AI development, into compliance with the Berne Convention.

The opt-out requirement in Article 4 constitutes a formality prohibited by Article 5 of the Berne Convention, and an otherwise-impermissible exception cannot be “saved” by making it conditional on a prohibited formality.

3. TDM Reservation Protocol (TDMRep), Final Community Group Report, <<https://www.w3.org/2022/tdmrep/>>.

The proposed TDM Reservation Protocol is technically complex and burdensome. It would depend on the inclusion of new codes in HTTP headers, HTML source code, or a new “tdmrep.json” file in the “.well-known” directory of the Web site root. But:

1. Authors (except a very few of the most technically sophisticated self-publishers of Web content) typically do not control any of these headers, codes, or files; and
2. Authors have no authority retroactively to modify contracts licensing their work for use on the Web to require Web publishers to include these headers, codes, or files.

In addition, the proposed TDM Reservation Protocol would apply only to future “scraping” of Web content, and would have no effect on works scraped before its proposal or adoption. It would provide no means for authors of works published in print formats to reserve their rights. And even if a new opt-out or TDM rights reservation protocol for printed works were to be defined, it would be impossible retroactively to include any new opt-out or statement of reservation of rights in works that have already been published and distributed in print.

Time is of the essence for corrective action by the European Union and United States governments.

We understand the reluctance of the EU and its member states to revisit the issues addressed in the 2019 Directive on Copyright. But compilation of generative AI language models were not what was discussed when TDM and Articles 3 and 4 were being considered.

We also appreciate the hesitancy of the US government to act prematurely. But the EU has already acted prematurely. AI companies – including US-based ones – are now relying on the exceptions in Articles 3 and 4 of the DSM Directive to carry out copying for AI development that would otherwise be clearly infringing. US entities are already able to outsource these otherwise-infringing activities to EU affiliates, subsidiaries, or proxies. And for-profit entities are already outsourcing these activities to nominally-nonprofit affiliates, subsidiaries, or proxies.

By allowing the TDM exceptions in Articles 3 and 4 of the DSM Directive to be applied to copying for AI development, the EU has *already* won the race to the bottom to create the world’s most favorable jurisdiction of convenience for copyright-infringing AI, before the US or any other jurisdiction can consider the issue or adopt appropriate rules for AI development and before there is any chance for rightsholders to negotiate terms for permission, remuneration, attribution, and/or objection to prejudicial uses of our works for AI development. Unless action is taken to cure these violations of the Berne Convention in the EU, it will be mostly irrelevant what legislation might be enacted in the US to require fair compensation for AI development.

EU-based entities have *already* copied works included on perhaps billions of Web pages and in an unknown number of scanned books and other printed works for AI development.

AI companies have *already* derived billions of dollars in their valuations from this infringing copying. Without the works copied, “ingested”, and available to be regurgitated,

generative AI would generate only worthless garbage. Much of the value of these companies is attributable to this infringing copying of human creators' works.

We urge the EU and its member states to promptly amend the DSM Directive or promulgate regulations or guidance to clarify definitively that the exceptions in Articles 3 and 4 do not apply to copying for purposes of developing generative AI software. Still, no purely forward-looking change in the text or interpretation of Articles 3 and 4 will redress the violations of authors' rights that have already occurred. Respect for authors' rights will also require redress for the infringements to date.

We appeal for prompt and effective action to bring EU member states into compliance with the Berne Convention as it applies to copying of our works for AI development and use.



National Writers Union (NWU)
<https://nwu.org>



American Society for Collective Rights Licensing (ASCRL)
<https://ascrl.org/>



Artists Rights Society (ARS)
<https://arsny.com>



American Photographic Artists (APA)
<https://apanational.org/>



Romance Writers of America (RWA)
<https://www.rwa.org>



National Press Photographers Association (NPPA)
<https://nppa.org>



American Society of Media Photographers (ASMP)
<https://www.asmp.org>



Sisters in Crime (SinC)
<https://www.sistersincrime.org>



Horror Writers Association (HWA)
<https://horror.org>



Graphic Artists Guild (GAG)
<https://graphicartistsguild.org>



Authors Guild
<https://authorsguild.org>



Dramatists Guild of America
<https://www.dramatistsguild.com>



Society of Composers & Lyricists (SCL)
<https://thescl.com>



Science Fiction & Fantasy Writers Association (SFWA)
<https://www.sfwaworld.org>



Novelists, Inc. (NINC)
<https://ninc.com>



Songwriters Guild of America (SGA)
<https://www.songwritersguild.com>



Music Creators North America (MCNA)
<https://www.musiccreatorsna.org>



Garden Communicators International (GCI)
<https://gardencomm.org>



Concept Art Association (CAA)
<https://www.conceptartassociation.com>



Society of American Travel Writers (SATW)
<https://satw.org>



Textbook & Academic Authors Association (TAA)
<https://www.taaonline.net>



Alliance For Women Film Composers (AWFC)
<https://theawfc.com>



Composers Diversity Collective (CDC)
<https://www.composersdiversitycollective.org>



Game Audio Network Guild (G.A.N.G.)
<https://www.audiogang.org>