

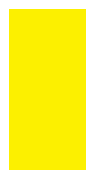
**OPEN
_FUTURE**

DEFINING BEST PRACTICES FOR OPTING OUT OF ML TRAINING

**OPEN FUTURE
POLICY BRIEF #5**

Authors: Paul Keller & dr Zuzanna Warso

29 SEPTEMBER 2023



EXECUTIVE SUMMARY

This Open Future policy brief examines the technical implementation of the EU law provision allowing authors and other right holders to opt out of having their works used as training data for (generative) machine learning (ML) systems.

With the adoption of the Copyright in the Digital Single Market (CDSM) Directive in 2019, the European Union has established a regulatory framework for the use of copyrighted works as commercial data for machine learning: Articles 3 and 4 of the CDSM Directive introduced copyright exceptions for text and data mining (TDM), which authorize the types of reproductions made in the context of training ML models on publicly available copyrighted works. Together, the two articles provide a clear legal framework: academic and cultural institutions may freely use lawfully available works for ML training (Art. 3), while others may only do so if the rights holders have not reserved their use (Art. 4).

This EU framework is unique in the world because it respects the rights of creators to exclude their works from ML training data. This addresses concerns voiced by creators and other rightholders about the impact of machine learning on the creative process and their income while providing legal clarity to ML researchers and developers regarding the use of publicly available information to train their models.

However, it is currently unclear how opt-outs from ML training based on the machine-readable reservation of rights provided for in Article 4 will work in practice, as there are currently no generally recognized standards or protocols for the machine-readable expression of the reservation.

While there are several potential solutions that allow creators and other rightholders to communicate their rights reservations in a machine-readable format, it is unclear whether and how opt-outs expressed through these tools will be respected by ML model developers. As a result, there is significant uncertainty for creators and rights holders about the practical benefits of investing in working with any of these tools.

This ongoing lack of clarity on how the opt-out from Article 4 of the CDSM Directive can be used creates a risk that the balanced regulatory approach adopted by the EU in 2019 will not work in practice, which would likely lead to a reopening of substantive copyright legislation during the next mandate.

In this situation, there is a growing recognition among stakeholders of the need to identify best practices for the communication of opt-outs under Article 4 of the CDSM. Such best practices need to address both the supply side (providing certainty to creators and rights holders on how to express opt-outs) and the demand side (incentivizing entities developing ML models to respect opt-outs).

In this brief, we argue that this requires the intervention of an actor with sufficient credibility to provide guidance on how to express machine-readable rights reservations. In the current

constellation, the entity best placed to take on this role is the European Commission, which is responsible for ensuring the implementation of the provisions of the CDSM Directive.

In the short term, the Commission should publicly identify data sources, protocols and standards that allow authors and rights holders to express a machine-readable rights reservation in accordance with Article 4(3) CDSM, that are freely available and whose functionality is publicly documented. Such an intervention would provide guidance to creators and other rightholders seeking means to opt out of ML training, and it would provide more certainty to ML developers seeking to understand what constitutes best efforts to comply with their obligations under Article 4(3) of the CDSM Directive.

Over time, this approach should be superseded by the emergence of a robust standard maintained independently of any direct stakeholders. It will be critical that both the standard and the technical and organizational infrastructures that support it are managed as a public good that is trusted by all relevant stakeholders.

THE ISSUE IN BRIEF

This brief takes a closer look at a key element of the EU legal framework governing the use of copyrighted works to train (generative) machine learning (ML)¹ systems: the rights reservation mechanism in Article 4(3) of the 2019 Directive on Copyright in the Digital Single Market. It seeks to understand how this mechanism can best be implemented to allow creators and other rightholders to opt out of the use of their works for the training of ML systems and suggests a number of recommendations on how to make the rights reservation mechanism more useful both for creators and rightholders and for developers of ML systems.

The content of this brief has been informed in part by a stakeholder seminar organized by Open Future, which brought together policymakers, academics, technologists, civil society organizations, and representatives of creators and rightholders to discuss best practices for opting out of ML training.² It reflects Open Future's understanding of the field in mid-September 2023.

BACKGROUND AND ASSUMPTIONS

With the adoption of the Copyright in the Digital Single Market (CDSM) Directive in 2019, the European Union has harmonized the rules that apply to the use of copyrighted works for so-called text and data mining (TDM), thus providing itself with a regulatory framework for dealing with the use of copyrighted works for training machine learning systems.

Articles 3 and 4 of the CDSM Directive introduce a number of copyright exceptions for text and data mining. And while this terminology may not immediately conjure up associations with machine learning and artificial intelligence, it clearly applies to the copyright-relevant acts that occur when copyrighted works are used to train ML models, as further explained below.

The CDSM Directive defines text and data mining as “any automated analytical technique aimed at analyzing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations.” This definition covers current approaches to machine learning that rely heavily on correlations between observed features of the works used as training data.

The exception in Article 3 of the Directive allows text and data mining for the purposes of scientific research by research organizations and cultural heritage institutions, as long as they have lawful access to the works to be mined:

Article 3(1): Member States shall provide for an exception [...] for reproductions and extractions made by research organisations and cultural heritage institutions in order to

¹ We use the terms Machine Learning instead of Artificial Intelligence, since we consider it a more accurate description of the technology discussed in this brief.

² The seminar was held under [the Chatham House rule](#), which is why we are not providing a list of participants.

carry out, for the purposes of scientific research, text and data mining of works or other subject matter to which they have lawful access.

The exception in Article 4 extends this by allowing anyone to use lawfully accessible works for text and data mining:

Article 4(1): Member States shall provide for an exception or limitation [...] for reproductions and extractions of lawfully accessible works and other subject matter for the purposes of text and data mining.

It does, however, impose an important condition that allows authors and other rightholders to opt out of this exception:

*Article 4(3): The exception or limitation provided for in paragraph 1 shall 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.*

Taken together, Articles 3 and 4 provide a clear legal framework for the use of copyrighted works as input data for ML training in the EU: researchers in academic research institutions and cultural heritage institutions are free to use all lawfully accessible works to train ML models for the purpose of their research. Everyone else – including commercial ML developers – can only use works that are both lawfully accessible and whose rightholders have not explicitly reserved their use for text and data mining.

The result is a differentiated legal framework that privileges uses of works in the public interest (the Article 3 exception), but allows those rightholders who actively manage their works to control how their works can be used for other purposes (the Article 4 exception). At the same time, this opt-out approach ensures that the vast majority of copyrighted material that is not actively managed by its creators or other rightholders can be used by anyone to train ML models.

In the current international context, the EU legal framework, which explicitly recognizes that creators and other rightholders have the right to exclude their works from use as training data for ML models, occupies a unique position: Most other jurisdictions lack similar protections for creators, either because they do not explicitly regulate the use of copyrighted works for ML training (e.g., United States) or because they contain blanket exceptions that permit such uses by anyone for any purpose (e.g., Japan).

The protection for creators provided by the EU framework is important in light of growing concerns among creators and other rightholders about the impact of generative ML models on the creative process, the market, and therefore their remuneration. The remainder of this paper will examine how creators and other rightholders can make use of the rights reservation system set out in Article 4(3) of the CDSM Directive in practice.

Assumptions Underlying this Document

As outlined above, we assume that Articles 3 and 4 of the CDSM Directive provide the relevant legal framework for the use of copyrighted works in ML training in the European Union.³ This means we assume that all reproductions of copyrighted works that take place as part of the training of ML models (compilation of training datasets, loading of training data into memory for the purpose of training) are covered by these exceptions – provided that they meet the requirements set out in the provision (lawful access, respect for opt-out).

We also assume that trained models do not contain copies of the works on which they were trained, nor can the models themselves be considered derivative works of the works on which they were trained.⁴ In this context, it is important to note that there is some evidence that ML models can "memorize" works contained in their training data,⁵ but this memorization seems to be limited to edge cases that are outside the intended behavior of these models.

The discussion in the rest of this paper is based on the above assumptions. It also focuses on ML training for general-purpose models used outside of scientific research and, as such, is concerned with uses that fall within the scope of the general TDM exception in Article 4 of the CDSM Directive.

A Pivotal Moment

We are at a very critical juncture. While it can be argued that with the TDM exemptions in the CDSM Directive, the EU legislator has indeed succeeded in providing a legal framework ahead of a technological breakthrough – the CDSM Directive was adopted in the summer of 2019 (1), while Generative ML models had their breakthrough moment in the summer of 2022 (2) – it is also clear that, at the time of this analysis, there is not a single standardized and widely used opt-out mechanism, although various mechanisms have become available, for the purpose of Article 4 of the CDSM Directive (4).

³ This aligns with the position of the European Commission expressed in this [answer to a parliamentary question given by Mr. Breton on behalf of the European Commission](#) in March 2023.

⁴ See for example [Drexl, Hilty et al. \(2019\)](#), who note that "Training data is **not** stored within the model during the training process, and once the training process is completed, the model is fully usable independently of the data.

⁵ See [Nicholas Carlini et al. \(2023\)](#) and [Kent K. Chang et al. \(2023\)](#).

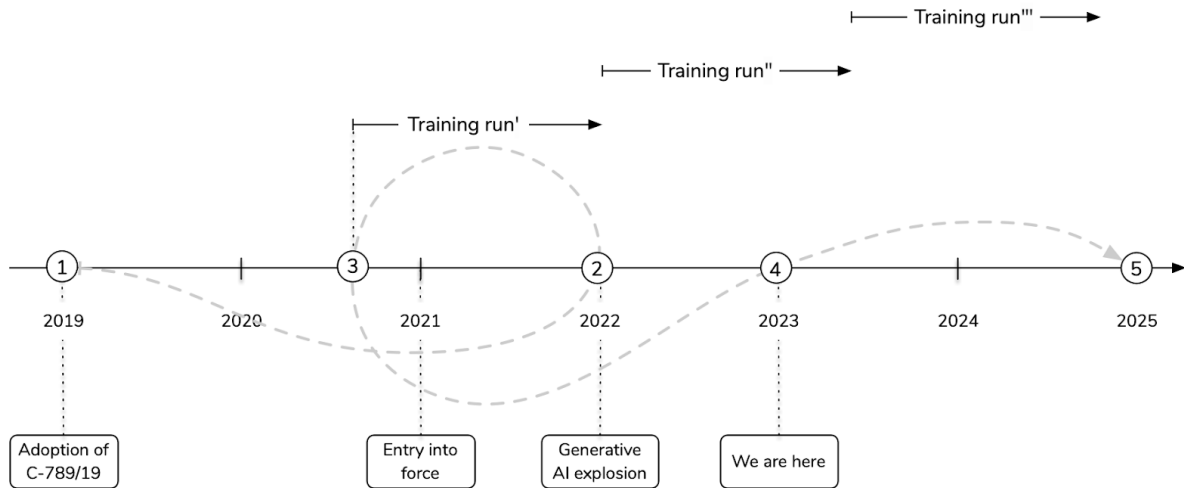


Figure 1: Development in time of understanding the relevance of TDM exceptions for ML training

This situation is exacerbated by the fact that there can be significant time delays due to the duration of the training of ML models. Although exact information is not always publicly available, it has been observed that there is a delay of up to 1.5 years between the initial selection of training data and the release of a trained model.⁶ This means the data selection for the ML models that had their breakthrough moment in the summer of 2022 had taken place in early 2021 (3) – at a time when few, if any, creators or rightholders were aware of the impact that generative ML systems would have on the creative sector or how these systems are trained.

Looking to the future, this also means that the adoption of universally followed best practices for opting out copyrighted works from ML training would only be reflected in trained models after a considerable delay (5).

This time lag between opt-out decisions and their practical effect raises concerns among creators and rightholders, particularly in the context of the general understanding that ML models, once trained, cannot (or it is extremely complex to) “unlearn”⁷ information derived from specific works that were present in their training data.



Figure 2: Simplified development pipeline for an LLM-based chatbot

⁶ This observation largely applies to the models that have been released over the past 12 months and that have been central to much of the discussion about the impact of generative AI. There are reasons – including the availability of more powerful hardware, such as the H100 series of GPUs – to believe that the delay between data selection and actual training runs may shorten considerably.

⁷ Recent research points to the possibility for ML models to “unlearn” specific concepts (but not individual works used as part of training data). See for example [Belrose, Schneider-Joseph, Ravfogel, Cotterell, Raff, and Biderman \(2023\)](#).

This situation increases the urgency of developing best practices for opting out copyrighted works from ML training that provide guidance to creators and other rightholders as well as to developers of (generative) ML models.

Continued lack of clarity on how to make use of the opt-out from Article 4 of the CDSM Directive creates the risk that the balanced regulatory approach adopted by the EU in 2019 might fail in practice, which would likely lead to a reopening of substantive copyright legislation during the next mandate. Given the length of EU legislative processes, this would prolong the status quo and, as a result, fail to provide protection for creators and other rightholders in the immediate future.

STATE OF THE ART

There are currently no generally recognized standards or protocols for the machine-readable expression of the reservation of rights provided for in Article 4 of the Directive. When the Directive was adopted, it was expected by policymakers and stakeholders that such standards would gradually emerge through industry practice. At the time of the Directive's adoption, many observers⁸ pointed to the [robots.txt](#) protocol⁹ as a possible vector for expressing the rights reservation, or saw `robots.txt` as an example of a future standard, but this has not been the case.

TDM Reservation Protocol

At the time of the generative ML explosion in the summer of 2022, there was only one publicly documented initiative to create a TDM reservation protocol: the [TDM Reservation Protocol](#), published in December 2022 and developed by the [W3C Text and Data Mining Reservation Protocol Community Group](#).¹⁰ However, ML training was not an imminent use case when the protocol – which consists of a flag for a rights reservation and the ability to provide links to information about available TDM licenses – was developed.¹¹

⁸ See for example Bernt Hugenholtz “[The New Copyright Directive: Text and Data Mining \(Articles 3 and 4\)](#)” from July 2019.

⁹ `robots.txt` is the filename of a file used for implementing the Robots Exclusion Protocol, a standard used by websites to indicate to web crawlers which portions of the website they are allowed to crawl. `robots.txt` was proposed in 1994 and relies on voluntary compliance. Most search engines comply with the standard, as do web scrapers such as [Common Crawl](#) that play an important role in the curation of training data for ML models. The underlying Robots Exclusion Protocol was proposed as an IETF standard ([RFC 9309](#)) in 2022.

¹⁰ W3C Community Groups should not be confused with W3C Working Groups. [According to the W3C website](#) “Community Groups enable anyone to socialize their ideas for the Web at the W3C for possible future standardization.” The development of actual [W3C web standards](#) takes place in W3C Working Groups. The W3C Text and Data Mining Reservation Protocol Community Group was set up and predominantly consists of representatives of organizations from the European publishing sector.

¹¹ The possible application of the protocol to ML training was first discussed during [the April 2022 meeting of the Community Group](#).

Since ML training is a specific application of the more general concept of text and data mining, it logically follows that the rights reservations expressed through the protocol would include the use of the works in question for ML training. The protocol is designed as an open solution, applicable to different content types (text, video, images) and granularity levels in a digital environment. It applies primarily to web pages/websites,¹² making it more useful as a tool for publishers and other intermediaries who have full control over their publishing infrastructure, and less suitable for individual creators in case they intend to opt out for specific works (or entire repertoires) that may be spread across a number of web sites and other publishing channels. At the time of writing, there is some initial uptake by European press publishers¹³ and in the educational publishing sector, and several manifestations of interest by book publishers associations. In July, Spawning (see below) partially integrated the TDM reservation protocol into their API.

Spawning and ai.txt

[Spawning](#) is a Berlin-based initiative launched in September 2022 as a collaboration between artists and machine learning engineers (Mat Dryhurst, Holly Herndon, Patrick Hoepner, Jordan Meyer).¹⁴ Spawning initially offered a single product, the website haveibeentrained.com, which allowed creators to search the LAION 5B training dataset for their own works and subsequently register an opt-out for works they wished to exclude from ML training. The company subsequently announced that the generative AI tool developer [Stability AI](#), which uses the LAION 5B dataset to train the Stable Diffusion image generator, would honor these requests. In March 2023, Spawning announced that it had collected 80 million opt-outs (including 40,000 opt-outs from individual creators).¹⁵

Since then, Spawning has launched two additional products. The first is an API that allows AI developers to integrate the opt-out information collected by spawning.ai into their own (ML training) processes. [The Spawning API](#) provides access to the opt-out information collected by spawning.ai itself but also aggregates information collected by other machine-readable opt-out systems.¹⁶

In May 2023, Spawning also released a specification for [ai.txt](#), which allows anyone running a website to set machine-readable permissions for commercial text and data mining. Similar to

¹² All three implementation techniques specified for the protocol – http headers, well-known file hosted on origin server, and html metadata – require access to the configuration of a website. However, in a meeting in July, the Community Group decided to also explore the possibility to include the TDM protocol in EPUB files.

¹³ As of September 2023, some European press publishers implemented the protocol; recently, GESTE, the federation of French online publishers, recommended the adoption of the TDM protocol in a [position paper on AI](#).

¹⁴ Spawning has since been incorporated as a US-based company that has taken on external funding (see more details in [this blog post](#)).

¹⁵ By the time of this writing, Spawning claims to have collected more than 1.4 billion opt-outs.

¹⁶ See <https://github.com/Spawning-Inc/datadiligence> for an overview of the opt-out methods supported by the Spawning API.

the way `robots.txt` works, `ai.txt` files reside in the root directory of a website and provide instructions on whether the images, media, and code hosted on the website can be used to train AI models. Here, `ai.txt` encounters the same limitation as the TDM reservation protocol: It requires control over a website/web server and cannot be used on individual works independently.

Content Authenticity Initiative

In April, the Adobe-led [Content Authenticity Initiative](#) (CAI) and its associated nonprofit standards organization, the [Coalition for Content Provenance and Authenticity](#) (C2PA), [released version 1.3](#) of its technical specification, which includes support for expressing creators' "do not train" intent.

The specification includes a number of entries related to [text and data mining](#). These allow a general opt-out from TDM (`c2pa.data_mining`) and more specific opt-outs from AI training (`c2pa.ai_training`) and generative AI training (`c2pa.ai_generative_training`). Each of these entries can specify whether the activity is allowed, prohibited, or restricted (in the latter case, it is possible to provide information pointing to additional license terms).¹⁷

At the time of writing, it is unclear how widely the "do not train" statements supported by C2PA are used and to what extent the technical standard is implemented in tools that are widely available to creators.

Company specific initiatives

In addition to these general initiatives, there are a growing number of platforms that provide machine-readable "no-ai" information, such as [Artstation](#) and [DeviantArt](#).

In July 2023, Google announced the launch of a process to explore "AI web publisher controls." In [the announcement](#) – which is otherwise vague and invites interested parties to sign up for a mailing list – Google states its belief that "it is time for the web and AI communities to explore additional machine-readable means for web publisher choice and control for emerging AI and research use cases."

In August 2023, Open AI [published documentation](#) on how GPTBot, the web crawler used to collect training data for its GPT series of large language models, can be blocked from accessing content via `robots.txt`. According to the document, GPTBot will honor "allow" and "disallow" rules specifically directed at the crawler. While this allows website owners to provide machine-readable opt-out information, the vendor specific nature of this implementation would mean that website publishers would need to specify opt-out information for each individual crawler

¹⁷ The C2PA also contains a `c2pa.ai_inference` entry that can be used to allow or disallow use of the asset "as input to a trained AI/ML model for the purposes of inferring a result." This permission cannot be derived from the TDM exception in the CDSM Directive, and the question of using works as input for the purposes of inferring a result is a copyright relevant act that is not covered by other exceptions and requires a separate legal analysis that is outside the scope of this document.

collecting ML training data. This approach seems flawed as it would require website owners wishing to opt out to have perfect information about all crawlers collecting ML training data and to frequently use their robots.txt files in response to the emergence of new crawlers. In September 2023, Open AI also started providing creators and other rightholders with the ability to manually opt out images via a web-form.

State of the Art

At the time of this writing in the summer of 2023, about a year after the explosion of generative AI and four years after the adoption of the Copyright in the Digital Single Market Directive, there is no clear or widely-adopted standard for expressing a machine-readable rights reservation that complies with Article 4(3). Instead, the field of machine-readable opt-out standards and tools is rapidly evolving.

Among the approaches highlighted above, [Spawning's](#) seems to have the most traction, although the impressive numbers (1.4+ billion works opted-out) seem to be primarily the result of partnerships with large platforms and rightholders. While these opt-outs are related to works of visual art, the proposed ai.txt standard also covers other types of media, like text, audio, video, and code.

What currently sets Spawning apart is the API that allows its services to be integrated into ML training tools and services¹⁸ and the fact that its API aggregates opt-out information from sources outside of Spawning's own data.

While it is unclear to what extent tools using the C2PA specification are currently in use, the C2PA approach is unique in that it addresses different types of ML systems. The ability to opt out of TDM altogether, to opt out only for training generative ML systems, or to opt out of all ML training (while allowing other types of TDM) is interesting in that it addresses a need to distinguish between different types of ML training that has been articulated by some types of rightholders (see below).

While creators and other rightholders have an increasing number of systems for expressing rights reservations from which to choose, the most important shortcoming of the current situation is the fact that there are no clear commitments from most¹⁹ of the major entities that train generative ML systems. This results in significant uncertainty for creators and rightholders who wish to exclude their works from ML training. There also is a risk that creators and other rightholders are confronted with an increasing number of AI company-specific opt-out mechanisms that require constant vigilance from them.

¹⁸ See for example Hugging Face's Data Sourcing Reports for data sets containing image URLs [that are powered by the Spawning API](#).

¹⁹ The notable exceptions here are [Stability AI](#), which has committed to respect opt-outs registered via [Spawning](#) when training the next generation of the Stable Diffusion image generator and Hugging Face, which has committed to observing opt-outs and shows a Spawning data report on every data set uploaded to its platform.

STAKEHOLDER PERSPECTIVES

There are at least two groups of direct stakeholders²⁰ when it comes to the use of copyrighted works as training data for generative ML systems: Creators and other rightholders, and the entities involved in building and deploying such systems. The following section discusses concerns that have been articulated by representatives of each of these groups, as well as a number of concerns that are relevant to both contexts.

Creators and Other Rightholders

Many creators and other rightholders are increasingly concerned about the impact of generative ML on their livelihoods and business models. In the EU policy debate, many rightholder organizations recognize that the TDM exceptions introduced by the CDSM Directive provide the relevant legal framework,²¹ but they have also identified two shortcomings:

1. The lack of clear and recognized standards for effective reservations of rights, and
2. The lack of transparency regarding the training data used by ML developers, which makes it difficult to assess the extent to which ML developers comply with the requirements of the exceptions (respect for opt-outs, use of works on the basis of lawful access).

In response to the latter issue, organizations representing creators and other rightholders are lobbying for the inclusion of training data transparency requirements in the upcoming AI Act.²²

In the context of the expressed need for clear standards for effective rights reservations, organizations and rightholders have identified a number of other issues and requirements that such standards would need to address.

- Machine-readable rights reservations must be able to differentiate (or express different policies) for different uses. There needs to be a clear distinction between opting out works from use in ML training and from indexing/crawling works and websites for purposes such as providing search engines. Some stakeholders have also expressed the need to distinguish between the use of works for training generative ML models and the use of works for training other types of ML systems.

²⁰ There are a number of other groups that have at least an indirect stake in this discussion. This includes (but is not limited to) the users of generative ML systems (a group that will have a substantial overlap with creators) and those who research and study ML systems.

²¹ See for example these recent statements by [organizations representing individual creators](#) and [organizations representing intermediary rightholders](#). For an organization that does not recognize the TDM exceptions as the relevant regulatory framework (but nevertheless points to a lack of “technical [and] contractual arrangements”) see [this statement by the European Writers Council](#).

²² As a result, the European Parliament’s report on the AI Act contains a provision requiring providers of generative ML models to “make publicly available a sufficiently detailed summary of the use of training data protected under copyright law.” There are no corresponding requirements in the European Commission’s original proposal or the position adopted by the Member States.

- Machine-readable rights reservations should be standards-based and must be freely available and user-friendly for creators and other rightholders.
- There is considerable uncertainty about how machine-readable rights reservations can account for the fact that multiple copies of works (“duplicates”) exist in the online environment. To what extent can machine-readable rights reservations deal with duplicates (or altered versions) of a work? Does the rights reservation apply to the work as such, or does it need to be made for each instance of a work that is available?
- Some stakeholders representing individual creators have also pointed to the need to be able to combine machine-readable opt-outs with collective remuneration schemes for creators that allow the use of opted-out works subject to the payment of remuneration.²³ For these stakeholders, the opt-out is not an end in itself, but a means to reach remuneration arrangements.

The first two issues are at least partially addressed by the current state of the art. None of the approaches described in the previous section charge authors or other rightholders for using machine-readable rights reservations. Moreover, both the `ai.txt` approach offered by Spawning and the TDM reservation protocol operate alongside/in addition to `robots.txt`. And the TDM-related entries in the C2PA specification allow creators/rightholders to target uses for training specific types of ML models.

The other two concerns (duplication/multiple instances of a work and remuneration) are not currently addressed by the systems discussed above. The TDM reservation protocol and the C2PA specification provide the ability to link to permission information, but there is little evidence of existing collective licensing arrangements that leverage prior opt-outs.

ML Developers

On the side of the entities involved in building and deploying (generative) ML models and systems,²⁴ there is a less clearly defined set of concerns.

The most obvious issue is that there is currently no clear commitment to machine-readable rights reservations on the part of the majority of entities providing (generative) ML tools and applications.²⁵ As a result, there is also no information on which sources of rights reservation information are consulted/supported.

²³ Such voluntary (collective) licensing arrangements seem to be compatible with the system introduced by the CDSM Directive. There are also more far reaching calls for mandatory remuneration schemes/levies for the use of copyrighted works for ML training that are clearly incompatible with the existing regulatory framework and are therefore out of scope for this paper.

²⁴ In the context of this paper, these will most often be companies but can also be open source projects or nonprofit entities that do not fall within the scope of the exception in Article 3 of the CDSM Directive.

²⁵ The exceptions from this are Stability AI, which has committed to respect opt-outs collected by Spawning from future models, and Open AI, which indicates that its GPTBot crawler will honor custom instructions in `robots.txt`.

This lack of industry commitment and alignment in turn creates a situation where creators and other rightholders face considerable uncertainty in determining how to express rights reservations. The aforementioned announcement by Google to develop web publisher controls seems to indicate that addressing this issue is also seen as important by some of the larger players in the field.

Other Issues

In addition, there are more systematic issues that are likely to affect the development of best practices and/or standards for machine-readable opt-outs from the use of copyrighted works for ML training.

While Article 4 of the CDSM Directive entrusts creators and other rightholders with the ability to reserve their rights, the validity of such machine-readable reservations will be difficult to assess unless there are measures in place to prevent false claims by persons or entities other than the actual rightholders.

All of the approaches described above involve attaching some kind of metadata either to digital representations of copyrighted works or to entire websites, and offer relatively low barriers to abuse: Spawning claims to maintain a know your customer (KYC) verification process for opt-outs above a certain threshold of opted-out works. The TDM reservation protocol requires control of the server in question. The C2PA specification has a strong focus on provenance, making it likely that opt-out information will be accompanied by provenance information.

None of these systems appear to be designed to prevent accidental or systematic abuse, such as someone deliberately opting out works of another creator or rightholder, or works that are not copyrighted or freely licensed. With respect to the latter, a registration system for works that are out of copyright or freely licensed (and thus can be used for ML training²⁶) could provide some protection against misuse.

THE WAY FORWARD

In the current situation, it is not clear how the opt-outs from ML training based on the machine-readable rights reservation foreseen in Article 4(3) of the CDSM Directive will work in practice. While there are a number of potential solutions that allow creators and other rightholders to communicate their rights reservations in a machine-readable way, it is unclear if and how opt-outs expressed through these tools will be respected by ML model developers.²⁷ This in turn means that there is considerable uncertainty for creators and rightholders as to the practical benefits of investing in working with any of these tools. As a result, the forward-looking legal

²⁶ Such a system would be structurally very similar to the systems [that we have previously proposed in the context of the implementation of Article 17 of the CDSM Directive here](#).

²⁷ While there is a clear legal obligation to respect opt-outs, it is currently unclear under which legal regime (and in which jurisdiction) ML training efforts take place. This is likely one of the aspects creating the uncertainties when it comes to compliance.

framework for ML training established with the adoption of the CDSM Directive currently lacks meaningful practical application and must be considered largely ineffective.

At the same time, there seems to be a growing recognition among stakeholders that there is indeed a need to identify best practices for the communication of opt-outs in accordance with Article 4(3) of the CDSM. Such best practices need to address both the supply side (providing certainty to creators and rightholders on how to express opt-outs) and the demand side (incentivizing entities developing ML models to respect opt-outs).

This requires an actor with sufficient credibility to provide some form of guidance on how to express machine-readable rights reservations. In the current constellation, the entity best placed to take on this role is the European Commission, which is responsible for ensuring the implementation of the provisions of the CDSM Directive. The European Commission could do this either directly or through an executive agency such as the EUIPO.

The CDSM Directive does not explicitly provide for the Commission to issue guidance on this topic,²⁸ but given the urgency of the situation, this should not withhold the Commission from providing guidance. At this stage, this could be limited to a relatively small intervention.

At a minimum, this would consist of an online information resource listing data sources, protocols and standards that allow authors and rightholders to express a machine-readable rights reservation in accordance with Article 4(3) CDSM, that are freely available, and whose functionality is publicly documented. Entries should be added at the request of the developers/maintainers of such standards and should be validated by the Commission (or EUIPO) before being added.

This online information resource would serve two purposes: It would provide guidance to creators and rightholders seeking means to opt out of ML training, and it would provide more certainty to ML developers seeking to understand what constitutes best efforts to comply with their obligations under Article 4(3) of the CDSM.

Maintaining such a list would make it possible to provide guidance in a situation where there are no clear standards supported by all stakeholders. The list approach, which would treat all providers of data sources, protocols, and standards equally, would also provide an opportunity for services that aggregate opt-out information (such as the Spawning API discussed above). Aggregators that bundle information from all services included in the list can significantly reduce the compliance burden for ML developers and provide certainty in an evolving landscape.

In the longer term, this approach should be superseded by the emergence of robust standards supported by all relevant stakeholders. Here, the experience with the robots.txt standard

²⁸ The only case where the directive explicitly requires the Commission to provide guidance is with regards to the implementation of Article 17 of the Directive, which was published in 2021 in the form of a formal Communication from the Commission to the European Parliament and the Council: [Guidance on Article 17 of Directive 2019/790 on Copyright in the Digital Single Market](#).

could serve as a template. A standard (or set of standards) that is maintained independently of any direct stakeholders seems to have the best potential to be trusted by all relevant stakeholders.

In the short term, however, the priority for the Commission and all other stakeholders must be to provide a minimum level of clarity. As highlighted above, the time lag built into the existing legal framework makes swift action even more necessary, and the EU legislator has a responsibility to provide creators who feel threatened by the current technological environment with effective means to exercise the rights they have been granted in 2019. Otherwise, the EU regulator risks undermining the legitimacy of the current legal framework, which would likely lead to a review, resulting in an even longer period of uncertainty. Europe is fortunate to have a balanced legal framework for the use of copyrighted works for ML training and should not squander the opportunity to put it into practice.



ABOUT OPEN FUTURE

[Open Future](#) is a European think tank that develops new approaches to an open internet that maximize societal benefits of shared data, knowledge and culture.

[Paul Keller](#) is a co-founder and director of policy at Open Future. His work focuses on the intersection of copyright policy and emerging technologies. He works on policies and systems that improve access to knowledge and culture and protect the digital public sphere.

[dr Zuzanna Warso](#) is the Research Director at Open Future. She has over ten years of experience with human rights research and advocacy. In her work, she focuses on the intersection of science, technology, human rights, and ethics. She holds a Ph.D. in International Law from the University of Warsaw.



This report is published under the terms of the [Creative Commons Attribution License](#).