

**OPEN
_FUTURE**

A VOCABULARY FOR OPTING OUT OF AI TRAINING AND OTHER FORMS OF TDM

**OPEN FUTURE #8
POLICY BRIEF**

Author: Paul Keller

7 MARCH 2025



INTRODUCTION

This document consists of two parts: (1) a proposal for a vocabulary for opting out from AI training and other forms of text and data mining (in Annex 1) and (2) a paper that elaborates on the process of coming up with the vocabulary that places the vocabulary in the context of our previous work related to machine-readable opt-outs¹.

This paper and the proposed vocabulary reflect the position of Open Future. Both parts are based on extensive interactions with various stakeholders involved in discussions about the implementation of the legislative framework related to the use of copyrighted works for the purpose of training AI models.

We are publishing this paper and vocabulary proposal to facilitate the development of robust, interoperable standards for machine-readable rights reservations on copyrighted works used for AI training.

PREVIOUS WORK

We initially identified the need for a vocabulary as part of our policy brief on “[Considerations for Implementing Rightholder Opt-Outs by AI Model Developers](#)”. This policy brief listed four areas that required further convergence: (1) identifiers, (2) vocabulary, (3) infrastructure, and (4) the effects of opt-outs. With regard to the vocabulary issue, the paper concluded that:

All of this points to the need for a common vocabulary for machine-readable opt-outs. Defining such a vocabulary should be a relatively straightforward step that equally benefits rightholders and AI training companies. Based on the considerations above, it seems desirable that any compliance policies should be based on a vocabulary that distinguishes between a full TDM opt-out (no-tdm) and an opt-out from generative AI training that applies to the use of works for no-generative-ai is a more specific version of the no-tdm opt-out, for the purpose of training generative AI models, either of these two would signal an opt-out to the model training the subset of AI models described in recital 105 of the AI act (no-generative-ai).

In our follow-up work to the considerations paper we focussed on the vocabulary question because there is a shared understanding among most stakeholders that addressing the vocabulary issue represents a logical first step and has the potential to unlock various follow up steps. This understanding was in part based on the fact that the question of defining a vocabulary also came up in other contexts such as [a workshop convened by the IETF's Internet Architecture Board](#) that was held in Washington DC in September 2024.

¹ An overview of our work on this issue can be found [on our website](#).

WHAT PROBLEM ARE WE ADDRESSING WITH THE VOCABULARY?

On its most basic level, a vocabulary would provide a standardized set of terms that describe types of use of protected works and other subject matter in the context of text and data mining and the training of AI models. The purpose is to provide a vocabulary of use cases that can be used to describe the scope of statements of intent such as rights reservations, opt-outs, or opt-ins (including licenses).

There is currently a multitude of systems that allow creators and other rightholders to make machine-readable rights reservations/opt-outs. The following table provides an overview of systems that were discussed during the IETF AI-control workshop in September 2024:

	Draft EU vocab	IBM	TDM-REP	IPTC/PLUS	TDM-AI	Spawning	CZPA	Open AI	Google	MSFT	Meta	Apple	Perplexity
TDM	Full opt out in line with Art 4(3)ODSMD		TDM-no Full opt out in line with Art	Data Mining Allowed / Prohibited			c2pa_data_mining Can any text or data content be extracted from the source for purposes of determining "patterns, trends and correlations".						
Gen AI training	Opt out from training generative AI models	AI training - all training-related activity, including pretraining, fine/instruction /etc. tuning,		Prohibited for generative AI/ML training	TDMAI: false	DNT / ai.txt (allows differentiation per media type)	c2pa_ai_generative_training Can the asset be used as training data to an AI/ML model that could produce derivative assets.	GPTBot used to crawl content that may be used in training our generative AI foundation	Googlebot-extended (sites help support search engines and various AI generative APIs, including future generations of models that power those products) - Google-CloudVertexBot	noarchive/ nochache no cache litem opt out / no archive full opt out	Meta-ExternalAgent crawls the web for use cases such as training AI models or improving products by indexing content directly.	AppleBot extended opt out website content being used to train	
AI training	Opt out form training non generative			Prohibited for AI/ML training			c2pa_ai_training Can the asset be used as data to train non-generative AI/ML models, such as those used for classification, object detection,						
AI inference / RAG / agents		ai-knowledge-base (the use of content as a knowledge base for an AI system, e.g., RAG)					c2pa_ai_inference Can the asset be used as input to a trained AI/ML model for the purposes of inferring a result.	ChatGPT-user may visit a webpage to help answer and include a link to the source in its response		noarchive /nochache no cache litem opt out / no archive full opt out	Meta-ExternalFetcher includes a list of individual links to support specific product functions. Because the fetch was initiated by a user, this crawler may bypass robots.txt files.		
AI search		AI-other?						GAI-SearchBot used to link to and surface websites in search results in the SearchGPT prototype, and OpenAI search features					Perplexitybot index search results that allow their AI Assistant to answer user questions.
Search	Via specific permissions / robots.txt		*TDM-Rep assumes that indexing for search is not covered by TDM	Prohibited except for search engine indexing					GoogleBot	BingBot		AppleBot	
Synthetic Data creation		ai-synthetic: the creation of a derivative data product such as synthetic data based on the content of a Web site											

Figure 1: Use cases covered by systems that were discussed during the IETF AI controls WS. For illustrative purposes only (as of Q3 of 2024, specific information may be outdated).

Some of these systems allow users to specify the scope of a reservation/opt-out, while others operate in a binary fashion or only apply to the actions of a specific user (such as the operator of a specific crawler). This has resulted in a situation where it is often unclear what types of use are covered by an opt-out or rights reservation.

This situation raises a number of issues:

- There is a lack of conceptual interoperability between these systems. They use the same terms without the same meaning. As a result, it is often not possible to translate expressions made through one system into another.

- From the perspective of creators and other rightholders, this means that they will need to invest in understanding the nuances of each system and cannot rely on a shared understanding of key concepts.
- From the perspective of model developers, who will likely need to comply with expressions made via multiple systems, this means that they will need to invest in understanding the nuances of each system and will need to deal with data that comes with different sets of constraints and cannot easily be aggregated.

All of this points to the need for some level of standardization at the vocabulary level. Given the fact that formal standardization processes tend to be relatively slow and that there is an urgent need to create more convergence right now, we have worked with opt-out systems providers, rightholders, AI model providers, and other public interest technology providers on a proposal for such a vocabulary (see Annex 1). We foresee two possible uses for the proposed vocabulary:

1. It can be taken up by opt-out model providers as an external source to define the scope of opt-outs/rights reservations expressed via their systems.
2. It can be used as input for ongoing standardization processes such as the work of the [IETF AI preferences Working Group](#) that was chartered in February of this year².

In addition, we also consider the publication of the vocabulary to be a contribution to the ongoing discussions about how machine-readable opt-outs from Generative AI training and other forms of TDM can be implemented in practice and how they can give creators and other information producers more control over how their works are used in the context of AI training.

LIMITATIONS OF OUR PROPOSAL

The discussions about the proposal have focussed on the functioning of the vocabulary in the context of the debate about the use of copyrighted content. They have been heavily influenced by the EU regulatory framework for the use of copyrighted works for the purpose of AI training that consists of Article 51(1)c of the AI Act and Article 4 of the Copyright in the Digital Single Market directive. Nevertheless, the vocabulary is intended to function both in contexts where opt-outs expressed give rise to legal obligations under copyright law and in contexts where this is not the case.

Being based on copyright considerations, the proposed vocabulary has not been designed to be used with regard to other rights and interests such as those related to the protection of personal data or personality rights.

In the current state, the use cases contained in the proposed vocabulary are also limited to uses of copyrighted works in the context of training AI models. There are currently no defined use cases that deal with preferences that content owners may have in relation to the use of their works by AI systems at inference time. This limitation is briefly discussed in some more detail at

² For this purpose we will also publish the proposed vocabulary as an [Internet Draft](#).

the end of this paper, and the possibility of adding additional (inference-related) use cases at a later stage is also acknowledged in the proposed vocabulary itself.

KEY CONSIDERATIONS UNDERPINNING THE PROPOSAL

When we described the need for a vocabulary in our previous policy brief, we noted that “defining such a vocabulary should be a relatively straightforward step”. In practice, this initial assessment has proven to be overly optimistic. While the general contours that we had outlined in the previous policy brief (“a vocabulary that distinguishes between a full TDM opt-out and an opt-out from generative AI training”) have made it into the proposed vocabulary, there have been a number of thorny issues that we have had to address to arrive at the current proposal.

These include the level of application of the vocabulary, the definition of text and data mining, the relationship with search, and the distinction between generative AI and other forms of AI, which will be discussed in the following sections.

Level of application

Discussions about opting out work from AI training and other forms of TDM often mix two different levels of use of copyrighted works in the process: (1) The acquisition of training data and (2) the subsequent use of the training data in the context of training AI models. Most public discussion about opt-outs is currently focussed on the acquisition of training data via web-crawling. As a result, most of the discussion focuses on the question of how the behavior of crawlers can be controlled. This is partially due to the fact that for many rightholders and web publishers, the act of crawling is the most “visible” touch point between the works that they make available and entities collecting data for purposes such as AI training.

Despite this, there are clear advantages to having a vocabulary of use cases that relate to how the collected works are subsequently used instead of how they have been collected. There are at least three reasons for this:

1. The act of crawling has largely been uncontroversial for the past 20 years and has enabled a wide range of useful services and activities. It is the use of crawled web data for the purpose of training generative AI systems that has given rise to issues that rightholder opt-outs seek to address.
2. Many crawlers are multi-purpose and it is therefore not always possible to clearly identify downstream uses of the crawled data at the point of crawling³.

³ Some of the most widely known and encountered crawlers are in fact multi-use. This is true for Common Crawl which makes its crawls available for multiple purposes (including AI training but also to create search indexes or for scientific research) but also for the crawlers used by Google. The Google-Extended control that Google [has introduced as a control for AI training uses](#) of the crawled data is a flag that applies to content crawled by GoogleBot. Data that is collected by GoogleBot that carries the ‘Google-Extended’ flag will not be used by Google to “improve Gemini Apps and Vertex AI generative APIs”.

3. There can be substantial delays between the time of crawling and the actual use of the crawled data in the context of training AI systems, which means that opt-outs expressed (or not expressed) at the time of crawling will often be outdated.

In addition to these concerns that arise in the context of location based opt-outs, the proposed vocabulary is also intended to be used with asset-based opt-outs⁴. In the case of asset-based opt outs crawling is only one of the possible data acquisition strategies and reading the opt-out will by definition require access to the content.

This means that the proposed vocabulary must target the actual use of the data independently from how it has been acquired. This would allow it to function in situations where data acquisition has a clearly defined intended use, but also in situations where the intended use cannot be determined at the time of crawling. For the latter to function reliably, crawler operators that make crawled data publicly available would need to commit to publish any opt-out information together with the crawled data.

The definition of Text and Data Mining

Any vocabulary for opting out from AI training and other forms of text and data mining that is intended to function in the EU regulatory context will need to be hierarchical in its structure. This is needed to accommodate the fact that the various types of use of copyrighted works in the context of AI training and deployment are instances of the much broader legal concept of Text and Data Mining (TDM) as defined in Article 2(2) of the 2019 CDSM directive.

As a result in its most basic form, a vocabulary will need to consist of an overarching Text and Data Mining category and a number of more specific categories addressing various types of use in the context of AI training:

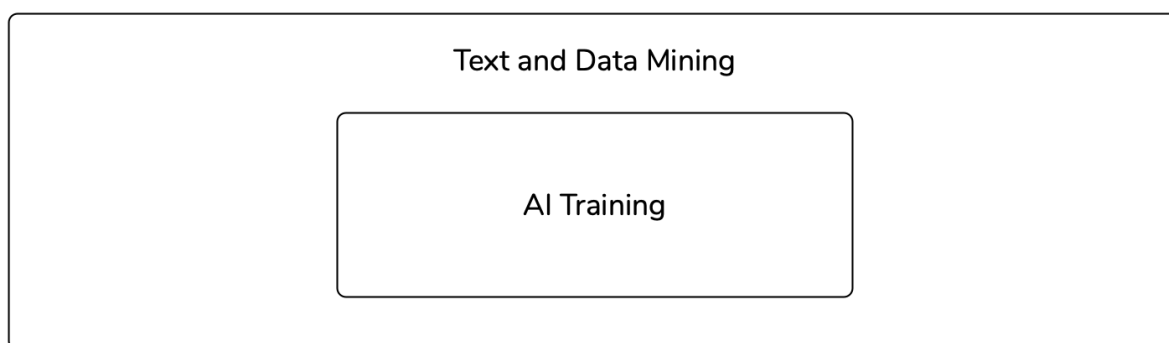


Figure 2: AI training as a subset of Text and Data Mining

While the definition of what constitutes Text and Data Mining seems straightforward at first glance, there is currently no agreement with regard to the exact scope of the activities that

⁴ Asset-based opt-outs are opt-outs where the opt-out signal is directly or indirectly linked to a media file. For the distinction between location-based and asset-based opt outs see: [Consideration for implementing Rightholder Opt-Outs by AI Model Providers](#), Open Future (2024).

should be considered to constitute TDM and that could therefore be affected from opting out from Text and Data Mining.

At the core of this issue lies the fact that the relationship between uses for the purpose of TDM and uses that enable search and discovery of online content is contested between different types of stakeholders.

On the one side, it is clear that the relatively broad language used to define the concept of TDM in Article 2(2) of the CDSM directive makes it hard to exclude with certainty any type of computational analysis from its scope:

‘text and data mining’ means 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 means that from the perspective of the user of copyrighted works, it seems prudent to assume that TDM encompasses almost any form of computational analysis of copyrighted works, including operations (such as ranking and classifying) that are commonly employed as part of online search and discovery services.

On the other hand, there are concerns that if the definition of TDM is understood to include such operations, rights reservations based on Article 4(3) of the CDSM Directive (i.e full TDM opt-outs) would also result in an exclusion of the opted-out works from search and discovery services. For rightholders in many creative economy sectors, the discoverability that results from being included in the search and other discovery services is a desirable outcome and an essential element of existing business models.

In line with this, some stakeholders involved in the discussions about a vocabulary hold that the concept of TDM was not intended to apply to “search and discovery” and that such uses should therefore be considered out of scope of the definition of TDM.

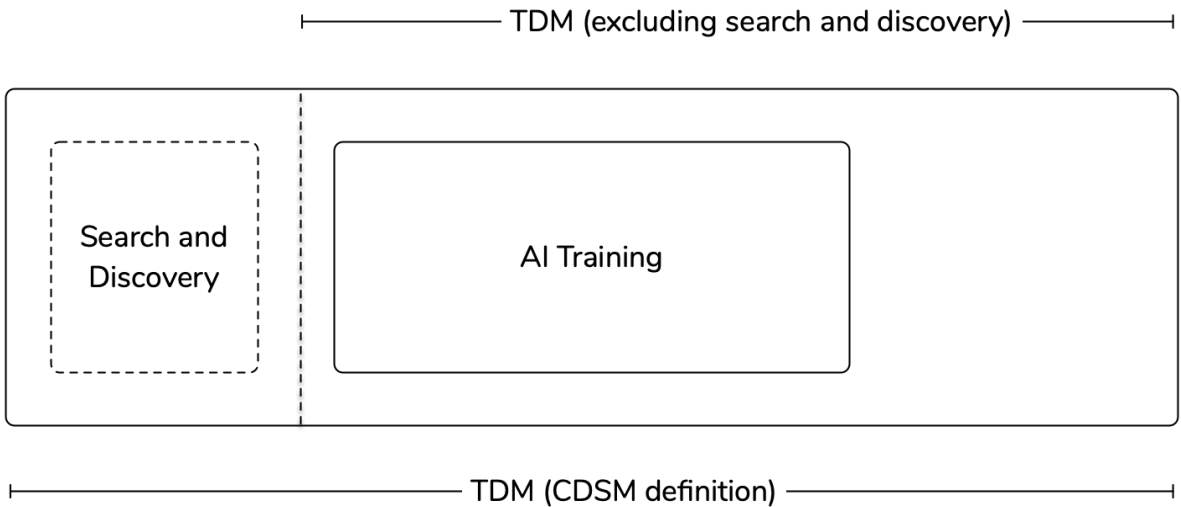


Figure 3: Competing conceptualizations of the scope of TDM

To be useful in the EU context, where creators and other rightholders have the ability to reserve the right to use their works for the purpose of TDM, the proposed vocabulary needs to include a definition of TDM, which must be the same definition in the CDSMD.

Without further clarification regarding search and discovery, using the CDSM definition of TDM would mean that any prudent users (including AI model developers) would need to interpret a full TDM opt-out as an opt-out from search and discovery services as well. As a result, creators and other rightholders who want to prevent the use of their works for text and data mining but want to ensure that they are available to be included in search and discovery services could not make use of the TDM category and would need to opt-out via a more specific category.

There are two ways to resolve this issue: The first one would consist of introducing a “TDM minus search and discovery” category that would exist as a subcategory of an overarching TDM category defined in line with the broader conceptualization of the definition of TDM under the CDSM directive. However, this approach poses two problems:

1. To be meaningful, it would require a definition of the scope of “search and discovery,” which is likely to be a very contested exercise in itself, and
2. It would introduce additional complexity for users of the vocabulary by adding another hierarchical layer to the vocabulary.

The second option—which is implemented in the proposed vocabulary—is to include a clarification in the non-operative parts of the vocabulary, stating that the vocabulary is not intended as a control mechanism for crawling websites for the purpose of building a search index as there are existing and more specific standards and protocols for this purpose, including but not limited to the Robots Exclusion Protocol ([RFC 9309](#)). This clarification can be further backed by including in the text of the proposed vocabulary the principle that when using the vocabulary, more specific instructions have preference over less specific ones.

In practice, this would mean that (the more specific) instructions related to search and discovery uses would continue to be made via the existing means, such as RFC 9309, and with the aim of ensuring that full TDM reservations do not impact a site's inclusion in search engines.

Differentiating between Generative AI and non-generative AI

The policy discourse on regulating the use of copyrighted works for AI training emerged in response to the first publicly available AI systems capable of generating synthetic content, including outputs resembling or matching various types of copyrighted works. These systems were trained – among other data sources – on large amounts of copyrighted works obtained without permission of their creators or other rightholders. These services were built on top of a new generation of so-called Generative AI models (transformer-based Large Language Models and diffusion models).

However, these are not the only types of AI models that exist. Many other models have been in use for years without raising similar concerns among creators and other rights holders. The

same applies to large-scale web crawling, which was also not a major concern before the advent of this new generation of generative AI models.

It is therefore desirable that a vocabulary allows to make a differentiation between the uses of copyrighted works for the purpose of training generative AI models and uses for the purpose of training other – non-generative – types of AI models.

There are broadly two main ways to define generative AI models in the context of a use-based vocabulary:

1. A capability-based definition: (‘models that have the capability to generate synthetic content’).
2. A purpose- or intent-based definition: (‘models that are intended to / have the purpose of generating synthetic content’).

It seems clear that in the current context, which is among other things characterized by a lack of trust between creators and rightholders on the one side and model developers on the other side, a definition solely based on intent-/purpose will not work as rightholders do not trust AI model developers to honor purpose limitations.

In addition to this lack of trust, there is also the issue that in the case of open-weight models (and to a lesser degree, proprietary models), it is not possible to enforce any purpose limitations after the training has taken place and the model has been made available for use by third parties who are not bound by opt-outs.

On the other hand, a capability-based definition of generative AI models runs into the issue that models may have the capability to generate synthetic content even if they are not designed or deployed for this purpose. For example a model that is being used as an image classifier will also have the capability to generate images.

This functional overlap is exacerbated by the current practice of training large foundational / General Purpose AI models. Based on the pre-training on large amounts of copyrighted works and other data, such General Purpose Models have the capacity to generate synthetic content. Before they are deployed, they will generally undergo fine-tuning with additional data (specifically selected for the intended purpose of the system) for a specific use case which may or may not involve the generation of synthetic content. Some of these use cases, such as filtering of copyrighted works, may very well be in the interest of rightholders to allow.

To resolve this situation we have based the definition of ‘Generative AI Training’ in the proposed vocabulary on the concept of ‘General Purpose AI Models’ but have expanded that definition with other models that are trained specifically for the purpose of generating synthetic content.

This approach has the benefit of being based on a category of AI models that is defined in the European Union’s AI Act and to which the copyright compliance obligations for Model developers are tied. At the same time it also acknowledges that not all models capable of generating synthetic content are also General Purpose AI models.

In addition the definition of Generative AI Training in the proposed vocabulary also contains a component that covers models which lack such generality in purpose but are nevertheless capable of generating synthetic content. Since any models that have the capacity for multiple purposes are covered by the concept of general purpose AI, this part of the definition can be purpose-based.

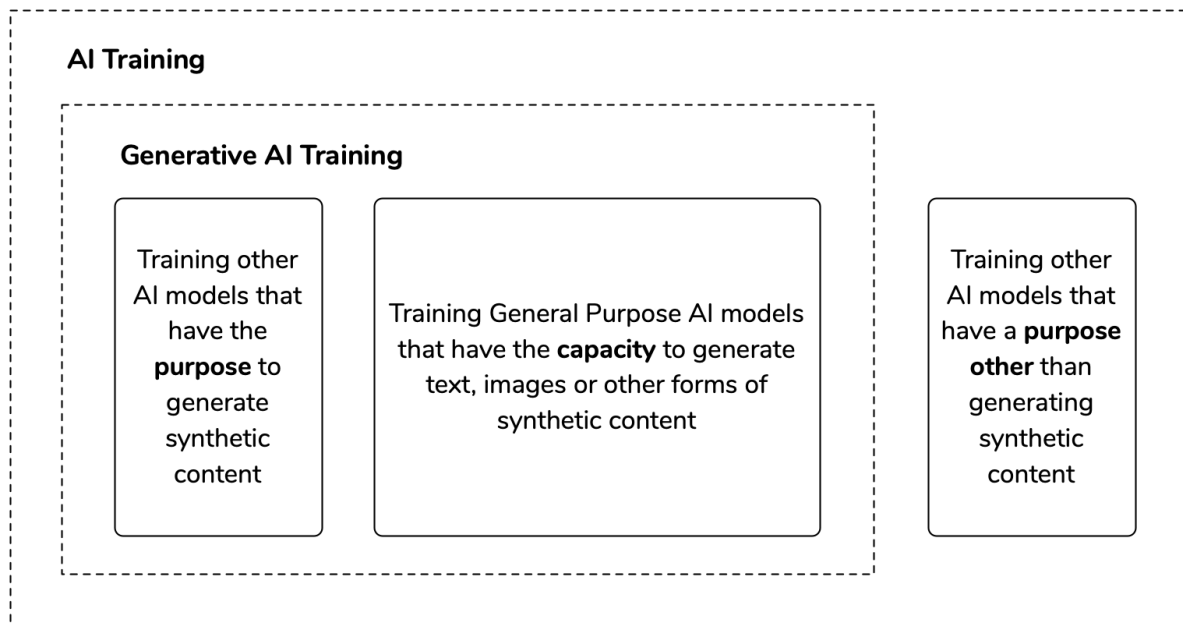


Figure 4: Generative AI training as a subset of AI training

To allow the expression of preferences related to the use of copyrighted works for the purpose of training non generative AI models, a vocabulary must also contain a category for other types of AI training. There are two ways to implement such a definition:

1. It can be constructed as a category that exists alongside (on the same hierarchical level) as the definition of Generative AI Training: ('The act of training AI models to perform specific tasks other than generating text, images, or other forms of synthetic content').
2. Or it can be constructed as a definition that exists as a broader category that encompasses the Generative AI Training category: ('The act of training AI models').

In the first case this would require rightholders who intend to opt out from all forms of AI training to explicitly opt-out of both Generative AI Training and AI Training whereas in the latter approach the same intent would be covered by a simple AI Training opt out.

The drawback is that with the hierarchical approach rightholders that want to opt-out of non-generative AI but allow generative AI training would be required to opt-out for AI Training and an opt-in for Generative AI Training. Since this scenario is rather unlikely (rightholders have primarily expressed concerns about uses of their works to build generative AI models) the proposed vocabulary implements the hierarchical approach.

Training vs. inference

The vocabulary that applied to the context of AI the proposed vocabulary only deals with the use of copyrighted works for the purpose of training AI systems. At this time it does not directly deal with the use of copyrighted works by trained AI models models and systems at inference time. Such uses include the use of works as input for answering search queries, the creation of summaries or translations of works. While much of the discussions about AI and copyright have been focussed on the use of copyrighted works for the purpose of AI training, creators and rightholders have also expressed the wish to be able to control uses of their works by trained models and systems.

In the EU some inference uses of copyrighted works will likely be in scope of the overarching TDM category although there may also be legal justifications other than the EU TDM exceptions allowing for inference uses. In addition, depending on the type of use such uses may simply be outside of the control of model developers, for example when they are directly initiated by end users or performed via models that are hosted locally.

On the other hand, it is also clear that inference uses generally do not fall within the scope of either the AI Training or Generative AI Training categories defined in the proposed vocabulary since they do not involve the use of the copyrighted work for the purpose of training a model.

As a result there may be a need for the addition of inference related use cases to be included in the vocabulary at a later stage or in a subsequent version.

USING THE PROPOSED VOCABULARY TO EXPRESS OPT-OUTS

All of the above leads to the following overall structure of the vocabulary:

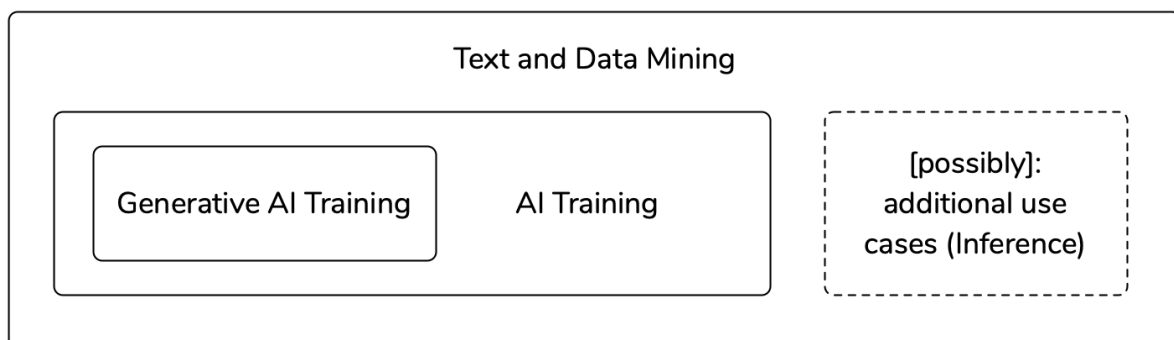


Figure 5: Overall vocabulary structure

For the purpose of opting out from AI training and other forms of text and data mining the proposed vocabulary would enable the following gradations of opt-outs:

- Opt-outs from all forms of Text and Data Mining should be expressed by using the TDM category.
- Opt-outs from any form of AI training should be expressed by using the AI Training category.
- Opt-outs from generative AI training should be expressed by using the Generative AI Training category.

As stated in the introduction and in section 5 of the proposed vocabulary, we do not expect creators or other rightholders to directly use the vocabulary. Instead, it is intended to be referenced by providers of opt-out systems and solutions to ensure that opt-outs from any of the defined use cases that are signaled via their system are functionally equivalent with opt-outs of the same use case made via all other opt-out systems that also use the vocabulary.

We hope that even with this relatively modest scope and relatively abstract level of application, the proposed vocabulary can contribute to the emergence of a set of robust and interoperable machine-readable opt-outs standards that are effective, scalable, and able to meet the needs of both rights holders and AI model developers.

ANNEX 1: PROPOSAL FOR AN OPT-OUT VOCABULARY

1. Purpose

The purpose of this document is to provide a common vocabulary that can be used for machine-readable opt-outs by parties who wish to restrict the use of their assets for the purpose of AI training and other forms of Text and Data Mining (TDM).

The elements of the vocabulary can be used to describe, in a standardized way, the types of uses that a declaring party may wish to restrict (or allow), thereby ensuring that opt-outs can be communicated, processed and stored in a consistent and interoperable manner.

The vocabulary is agnostic to the technical implementations of opt-out systems and is designed to ensure that opt-out information can be effectively exchanged between different systems.

The vocabulary is intended to govern the use of works in the context of training AI models and other forms of TDM but does not concern itself with the collection of training data (crawling). In particular the vocabulary is not intended for expressing instructions or restrictions related to crawling for the purpose of building a search index, as there are already more specific standards and protocols for this purpose including but not limited to [RFC 9309](#).

The vocabulary is intended to both work in contexts where such opt-outs expressed to the declaring party give rise to legal obligation (such as rights reservation made by rightholders) and in contexts where this is not the case. It is without prejudice to applicable laws and the applicability of exceptions and limitations.

2. Definitions

- **Asset:** A digital file or stream of data, usually with associated metadata.
- **Declaring party:** The entity that expresses an opt-out with regards to an Asset.

3. Vocabulary Structure

The vocabulary consists of the overarching TDM (Text and Data Mining) category and a number of specific use cases that can be addressed independently. The overarching category “TDM” is based on the definition of Text and Data Mining in [Article 2\(2\) of the European Union Directive on Copyright in the Digital Single Market](#).

4. Proposed Vocabulary

The following categories are defined for use in the opt-out vocabulary:

- **TDM:** Text and Data Mining. The act of using one or more assets in the context of 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.

- **AI Training:** The act of training AI models
- **Generative AI Training:** The act of training General Purpose AI models that have the capacity to generate text, images or other forms of synthetic content, or the act of training other types of AI models that have the purpose of generating text, images or other forms of synthetic content.

This list of specific use cases may be expanded in the future, should a consensus emerge between stakeholders, to include categories that address additional use cases as they emerge. In addition to these categories defined in the vocabulary, it is also expected that some systems implementing this vocabulary may extend this list with additional categories for their particular needs.

4.1 RELATIONSHIP WITH MORE SPECIFIC INSTRUCTIONS

The vocabulary does not preclude the use of other specific categories.

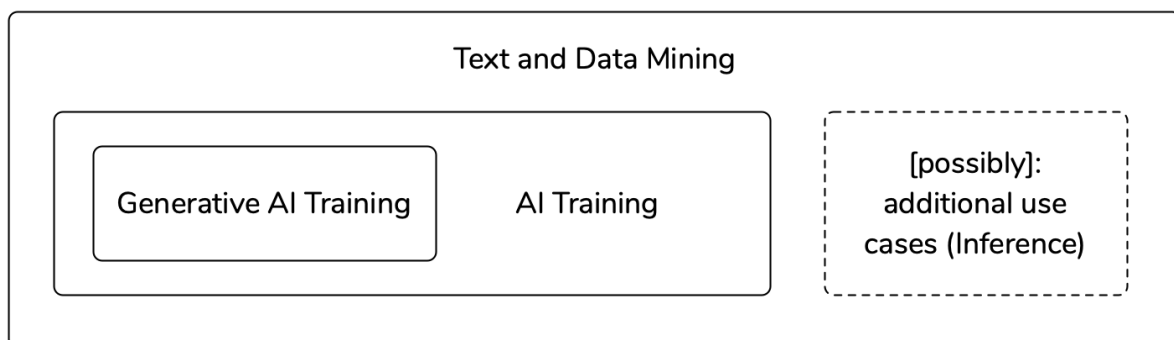
Any opt-outs based on this vocabulary shall not be interpreted as restricting the use of the work(s) strictly for the purpose of search and discovery as long as no restriction is declared through search-specific means such as RFC 9309.

When using this vocabulary more specific instructions – either based on the vocabulary or derived from other protocols – should be given preference over less specific ones.

4.2 RELATIONSHIP BETWEEN CATEGORIES

The TDM category is the overarching category that includes the AI training category. Generative AI training is a subset of the AI training category. Both AI training and generative AI training are considered to be forms of TDM. As such, when a Declaring Party opts out of TDM, they also opt out of these categories. AI model developers processing opt-outs must therefore interpret an opt-out from TDM to also mean an opt-out from Generative AI Training and AI Training.

The figure below shows the relationship between the currently defined categories:



Systems referencing the vocabulary must not introduce additional categories that include existing categories defined in the vocabulary or otherwise include additional hierarchical relationships.

5. Usage

The vocabulary may be used by declaring that an opt-out system or entity expressing or processing opt-outs uses the terms defined in Section 4, directly or via mappings, in accordance with how they are defined in this document.

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.

European
Artificial Intelligence
& Society Fund

This publication was made possible with support from the [European AI & Society Fund](#). Open Future is solely responsible for its content, which may not necessarily reflect the views of NEF or the Partner Foundations.



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