

Warszawa, 8 kwietnia 2026 r.

Szanowna Pani

Marta Cienkowska

Minister Kultury i Dziedzictwa Narodowego

Szanowna Pani Minister,

w imieniu Związku Cyfrowa Polska, zrzeszającego branżę cyfrową i nowoczesnych technologii, przesyłam poniżej nasze odpowiedzi, na kwestionariusz MKiDN, w ramach **przeglądu dyrektywy Parlamentu Europejskiego i Rady (UE) 2019/790 z dnia 17 kwietnia 2019 r. w sprawie prawa autorskiego i praw pokrewnych na jednolitym rynku cyfrowym oraz zmiany dyrektyw 96/9/WE i 2001/29/WE.**

I.1. Text and Data Mining for Research Purposes (Art. 3 of the DSM Directive / Art. 26² of the Copyright Act)

- 1. In your assessment, does the implementation of Art. 3 of the DSM Directive concerning text and data mining for research purposes facilitate cross-border research cooperation and the creation of partnerships between research organisations and industry?**

The Text and Data Mining (TDM) exceptions (Art. 3 and Art. 4 DSM Directive) are the fundamental bedrock for AI development and data-driven research in Europe. The Directive has been crucial in creating the legal certainty required for European AI champions and research organizations to innovate and compete globally. Training AI models on the open web is a transformational, non-expressive use of information. This is



**Porozmawiajmy
o technologii!**



Związek Cyfrowa Polska
ul. Twarda 2, 00-105 Warszawa
e-Doręczenia:
AE:PL-30344-97239-IDARJ-15



+48 (22) 666 22 46
biuro@cyfrowapolska.org
cyfrowapolska.org



KRS: 0000250359
REGON: 140463214
NIP: 5222802518

analysis—reading patterns to understand logic and language. AI models are not information retrieval systems nor databases. They are reading this information to create wholly new content and tools—tools that, in many cases, are of immense value to creators themselves.

The TDM research exception under Article 3 is a critical achievement for Europe's research community and enables European universities and research institutions to stay at the cutting edge of research. Restricting AI model training on publicly available data would severely restrict innovation and run counter to the goals of the European research landscape. Limiting the data that AI systems can learn from risks Europe falling behind in every scientific discipline that AI is transforming, from drug discovery to climate science to materials engineering.

While the corresponding recitals to Article 3 specifically refer to private - public partnerships and endorse them, several grey areas remain in its application. These uncertainties may still prevent industry players to rely on Article 3 in projects with eligible institutions.

- 2. In your view, to what extent has the TDM exception introduced in Article 3 been effective in your Member State in improving the ability of research organisations and cultural heritage institutions to carry out text and data mining for scientific research, including in the context of AI development? In your response, please consider whether the implementation has increased legal certainty for eligible institutions, improved practical access to content for TDM, reduced technical or contractual obstacles to conducting TDM, and led to any increase or change in TDM activities.**

Poland's Article 3 implementation (TDM for research) has created legal certainty for research organizations and cultural heritage institutions. The exception cannot be contractually overridden, enabling computational analysis of copyrighted works for scientific research and culture heritage purposes—including AI development. The work that culture heritage institutions carry out has a natural cross-border impact. Projects like the BBC Archives, the Natural History Museum (UK), and SUCHO (Ukrainian cultural institutions) enable cross-border access to cultural collections, including for Polish organizations, startups and citizens. These projects leverage cloud services to store and manage large collections of cultural assets, making large



scale consultations and uses possible, include for AI models training purposes. Private commercial enterprises play a central role in research, either independently or in close collaboration with other research organisations.

Poland has launched five notable AI initiatives in cultural heritage. The **Picture from Auschwitz** project (€1.5 million) creates brick-level 3D digital reconstructions of the concentration camp. **DRC-AI** at Cyfronet AGH/Jagiellonian University (~9.87 million PLN) uses AI for mass digitization, making 150,000+ digital objects publicly available. The **POLIN Museum Culture Chatbot** offers bilingual AI assistance with 82% student approval. The **Artistic Chatbot** at Warsaw Academy of Fine Arts provides voice-to-voice RAG-powered exhibition support with 60% response relevance. **AIMused** (Interreg South Baltic 2021-2027) develops AI-powered accessibility tools across Poland, Denmark, Sweden, Lithuania, and Germany.

Although copyright issues are not expressly mentioned among the key barriers slowing AI adoption, Poland is 2.5 years behind on transposing EU copyright law, creating legal uncertainty that reduces technology investment by 48%. Polish museums lack systematic approaches and cohesive digital strategies. 56% of museums cite funding shortages as a major barrier; Poland's basic digital skills rate (44.3%) lags the EU average (55.6%); enterprise AI adoption stands at just 3.7% versus the EU's 8%.

Official sources:

- Picture from Auschwitz: <https://www.theguardian.com/film/2025/may/16/new-historically-accurate-digital-replica-will-allow-films-to-be-set-within-auschwitz>
- DRC-AI: <https://www.cyfronet.pl/en/projects/national-projects/drc-ai>
- POLIN Chatbot: <https://polin.pl/en/culture-chatbot>
- Artistic Chatbot: <https://arxiv.org/html/2509.00572v1>
- AIMused: <https://aimused.eu/>

Poland's implementation of Article 3 has established a framework enabling research organizations and cultural heritage institutions to perform computational analysis of copyrighted works for scientific research purposes, including AI development. Poland has launched notable AI initiatives in cultural heritage, including the Picture

from Auschwitz project (€1.5 million), DRC-AI at Cyfronet AGH/Jagiellonian University (~9.87 million PLN), POLIN Museum Culture Chatbot, and AIMused (Interreg South Baltic 2021-2027).

Current EU copyright rules protect the interests of creators while also protecting the rights of secondary users to learn from and be inspired by works so as to create new works. The TDM exception allows organizations to undertake training on substantial volumes of copyrighted works without requiring individual permissions, whilst acknowledging rightsholders' expressed opt-out preferences. Safeguarding the exception for text and data mining remains a key element for AI-related innovation to flourish within the EU market. The right to opt out, when applicable, should be made compatible with this objective, e.g., expressed in a machine-readable format. International standardization can help, which is why industry-led machine-readable standards and protocols to express a TDM rights reservation with rightsholders, AI providers and other relevant stakeholders merit further discussion.

I.2. Text and data mining for non-research purposes

3. To what extent has your national implementation of the Article 4 TDM exception been effective in improving the ability to carry out text and data mining, including in the context of AI? Are you aware of specific situations / use cases in which the exception has been used? Please specify?

Poland has transposed the exception almost verbatim, and the Article 4 TDM exception has been effective. The Directive deliberately struck a balance by allowing AI model developers to utilise the TDM while giving rightsholders the option to decide against this for non-scientific uses via machine-readable opt-outs (Art. 4(3)). Companies developing AI have facilitated the ease with which publishers can opt-out and proactively respects this balance through machine-readable controls utilizing the established and scalable robots.txt protocol. This tool allows web publishers to easily manage whether their site content is used to improve our generative AI foundational models. This demonstrates that the existing legal framework of TDM exceptions combined with opt-out capabilities functions effectively in practice.

However Poland only introduced the Directive, and hence the commercial exception for AI development in September 2024. Despite this delay—which resulted in a longer period of legal ambiguity for local

innovation—the recent enactment of Article 26³ finally integrates Poland into the harmonized European framework. This crucial step ensures that the balanced ecosystem envisioned by the Directive can now operate effectively within the Polish market. Consequently right holders and AI developers can confidently navigate the legal framework, enabling developers to train models responsibly and legally.

The EU AI Act confirms in its recitals that TDM covers data use for AI model training. Consistent TDM exceptions across both research and commercial contexts remain essential to enable innovation while respecting rightsholder opt-outs expressed in machine-readable formats. In the existing legal framework, AI developers and rightsholders have the option to voluntarily enter into licensing agreements, including for the use of copyrighted works in AI training. Which specific data has value for AI training and should be licensed must be mutually agreed upon by the contracting parties. Information about licenses is naturally sensitive business information and is typically treated confidentially in the interest of both licensors and licensees, which is why public information about such agreements is limited.

4. Has the national implementation of Article 4 enabled text and data mining in areas beyond large language model development? Please indicate concrete cases or use cases you are aware of and any obstacles encountered.

We stand at a pivotal moment in the development of artificial intelligence. AI has the potential to fundamentally change the ways we live and work, through its capacity to assist, complement, empower, and inspire people in almost every field of human endeavor. As seen with the Nobel Prize, AI is poised to usher in a new era of scientific discovery, revolutionizing how we conduct research and accelerating the pace of scientific breakthroughs. While LLM development is a key area, the resulting models are versatile, facilitating text-and data mining in various applications. But this all rests on the ability of AI models to train on vast amounts of data. Balanced copyright frameworks seek to protect a private rightholder interest while also enabling and incentivising a public good which can include:

- **Scientific Research & Discovery:** Enabling researchers to analyze massive, complex datasets to accelerate advancements in a wide range of fields, including medicine, climate science, and materials engineering.



- **Economic Innovation & Competitiveness:** Empowering developers—from local startups to global enterprises—to train various AI models, as well as deploy applications built upon them that drive new technologies, optimize industries, and fuel national economic growth.
- **Preservation of Cultural Heritage:** Facilitating the digitization, analysis, and deep contextual understanding of large cultural datasets, ensuring local languages and histories are accurately represented in the digital age.
- **Access to Information:** Powering various search tools to ensure the transparent flow of information necessary to foster an informed and engaged public discourse.
- **Fighting Online Piracy:** Leveraging fingerprinting technologies and the automated filtering of takedown notices for the large-scale detection of pirated content to ensure the rapid, accurate removal of infringing content and the protection of right holders' interests.
- **Everyday Digital Services & Accessibility:** Powering the routine, non-generative AI tools that society relies upon daily, such as vital cybersecurity threat detection, advanced machine translation, and accessibility tools for users with disabilities.

To fully realize these public goods, the legal mechanisms governing data access—specifically TDM exceptions—must be practically workable and legally certain. When a copyright framework successfully balances clear, statutory permission to mine data with the provision of standardized, machine-readable tools for right holders to opt out, it fosters a predictable and thriving digital ecosystem.

Ultimately, ensuring that TDM exceptions function smoothly in practice is not merely a matter of regulatory compliance; it is a cornerstone of technological innovation and local competitiveness. Frameworks that embrace this balance ensure that the next wave of AI-driven research, innovation, and economic benefit is cultivated locally.

1.3. Use of works and other subject matter in digital and cross-border teaching activities

5. To what extent has your national implementation of Article 5 been effective in improving the use of works and other subject matter in digital and cross-border online teaching activities?



**Porozmawiajmy
o technologii!**



Związek Cyfrowa Polska
ul. Twarda 2, 00-105 Warszawa
e-Doręczenia:
AE:PL-30344-97239-IDARJ-15



+48 (22) 666 22 46
biuro@cyfrowapolska.org
cyfrowapolska.org



KRS: 0000250359
REGON: 140463214
NIP: 5222802518

Poland implemented Article 5 of the DSM Directive as Article 27 of the Copyright Act, establishing a mandatory exception for digital and cross-border teaching activities. This provision allows educational institutions to use copyrighted works for illustration purposes in teaching without individual licenses, provided proper attribution is included.

We have limited documented use cases specific to Article 5. However, Polish educational institutions using cloud services have focused primarily on digital infrastructure development for remote learning platforms rather than leveraging copyright exceptions for cross-border content sharing.

I.4. Permitted Use for the Purposes of Preservation of Collections (Art. 6 of the DSM Directive / Art. 28 of the Copyright Act)

6. To what extent has the implementation of Art. 6 of the DSM Directive into Polish law increased the capacity of cultural heritage institutions to preserve works or other protected subject matter? In your answer, please take into account the following issues:

- practical preservation activities made possible by the exception provided for in Art. 6 (e.g. digitisation, archiving, format conversion, circumvention of technological protection measures);
- any observable increase in protected resources;
- examples of successful collection preservation projects undertaken by cultural heritage institutions since the implementation of Art. 6 of the DSM Directive
- whether the implementation of Art. 6 of the DSM Directive has reduced legal, technical or organisational barriers to the preservation of cultural heritage?

BBC Archives Technology and Services manages 16 million assets spanning 100 years of content. The BBC digitized and migrated 25 PB of archives to cloud storage, making the content more accessible through digitization and modern storage solutions. Their goal is to safeguard content so it remains accessible for



**Porozmawiajmy
o technologii!**



Związek Cyfrowa Polska
ul. Twarda 2, 00-105 Warszawa
e-Doręczenia:
AE:PL-30344-97239-IDARJ-15



+48 (22) 666 22 46
biuro@cyfrowapolska.org
cyfrowapolska.org



KRS: 0000250359
REGON: 140463214
NIP: 5222802518

another 100 years, using machine learning tools like speech-to-text and facial recognition to enhance searchability.

The Natural History Museum houses 80 million specimens across five main collections (botany, entomology, mineralogy, palaeontology, and zoology). The museum is digitizing collections to enable global access and transform research capabilities. It is collaborating with cloud service providers to analyze environmental DNA at scale and better understand biodiversity, making their research accessible to international scientists.

Saving Ukrainian Cultural Heritage Online (SUCHO) is a volunteer initiative of over 1,500 international cultural heritage professionals who archived web content from 3,000+ Ukrainian cultural institutions during the 2022 invasion. They preserved 50 terabytes of data from museums, libraries, and archives, making digitized Ukrainian cultural heritage accessible globally through open access collections. The archived material will be repatriated to Ukrainian institutions when they can rebuild.

I.5. Cross-cutting questions regarding Articles 3-7 of the DSM Directive

7. Do you have any other information (including court rulings) that could be considered relevant to the implementation of Articles 3-7 of the DSM Directive into Polish law?

Regarding specific information and court rulings relevant to the implementation of Articles 3-7 of the DSM Directive, it is vital to note recent European jurisprudence that will guide Polish application. Notably, the April 2024 *Kneschke v. LAION* decision by the Hamburg Regional Court confirmed that scraping internet content to compile datasets for AI training explicitly falls within the scope of the CDSM Directive's TDM exceptions. The case *Like Company v Google* is still pending in the CJEU.

Beyond this, we are aware that some press publishers advocate for payment or licensing for the use of publicly accessible content in AI model training. While the EU's copyright law includes a Text and Data Mining (TDM) exception – allowing generative AI training unless a rightsholder explicitly opts out, as recently [confirmed by Commissioner Virkkunen](#) and reinforced by the aforementioned LAION court ruling – this framework does not preclude new commercial partnerships between AI developers and right holders for access to content: negotiating agreements and partnership deals for a variety of situations, including programmatic access to custom APIs, access to data, digitisation, etc. Given the early stages of this technology, of commercial

decisions, and the complexity of license deals themselves, it would be premature to introduce new measures to mandate licensing. Early signs show that a market is developing and as such a voluntary licensing market should be given time to develop and grow.

II.1. Use of Works Considered Out of Commerce (Arts. 8-9 of the DSM Directive / Division 6² of the Copyright Act)

8. To what extent has the implementation of Art. 8 of the DSM Directive into Polish law facilitated the licensing of out-of-commerce works and improved the access of cultural heritage institutions to such works? Please assess whether the implementation has in practice enabled the conclusion of licensing agreements and led to greater availability or use of out-of-commerce works.

Not applicable, no data

9. Do the licences concluded in practice cover cross-border use of out-of-commerce works for the purposes of cross-border access in accordance with Art. 9 of the DSM Directive?

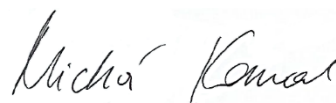
Not applicable, no data

10. Do cultural heritage institutions in practice make use of the exception provided for in Art. 8(2) (Art. 35²² of the Copyright Act et seq.) when making works considered out of commerce available? If so, in what situations?

See reply in Q. 6

Z wyrazami szacunku

Michał Kanownik



Prezes Zarządu

Związek Cyfrowa Polska