bitkom

Update on AI & Copyright?

Bitkom Commentary in light of recent developments

At a glance

Headline

Initial position

The Copyright & AI debate is, first and foremost, a competitiveness question for Europe. The core issue is whether organisations with lawful access to information may also lawfully analyse it at scale. A workable Text and Data Mining (TDM) framework is essential digital infrastructure - for AI, but also for cybersecurity, healthcare, research, and the productivity gains the Draghi agenda calls for. Discussions are ongoing; this paper sets out balanced positions that protect creators while preserving the legal certainty Europe needs to remain competitive.

The most important takeaway

New providers are represented in Bitkom, as are members who are very close to traditional services. Our paper therefore outlines possible compromise lines:

- **Preserve the TDM mechanism**

  The existing EU TDM framework should remain the baseline for AI training on lawfully accessible content. It already balances innovation and rightsholder control and should not be reopened through new (national) layers.

- **Rely on proportionate transparency, not new presumptions**

  The AI Act already introduces meaningful transparency on training data and opt-out compliance. Before adding new obligations, policymakers should let this framework operate and avoid impractical registries, presumptions of use, or blanket remuneration schemes.

- **Focus on workable licensing**

Voluntary licensing for restricted or high-value datasets is growing alongside and because of TDM, there is no need to turn lawful data analysis into a permission-based system.

Content

Policy Paper on AI & Copyright 4

General Remarks 4

Statement on Current Ideas Within the Existing Legal Framework 6

Managing Access and Opt-Outs 6

Transparency 7

Trusted Intermediary 9

(Ir)rebuttable Presumption 10

Help the licensing market? 10

Statements on Current Ideas Outside the Existing Legal Framework 11

Remuneration Obligation 11

Output 13

Copyright Protection for Individual Likeness 14

Call to Action 14

Policy Paper on AI & Copyright

General Remarks

Europe wants to lead in AI, but regulatory uncertainty—particularly around copyright—is actively driving investment elsewhere. The US produced 40 frontier AI models in 2024; the EU produced 3. Only 6 percent of global AI funding flows to European companies vs. 61 percent to the US. As of mid-2025, approximately 75 percent of global GPU cluster capacity sat in the US, just 5 percent in Europe. AI infrastructure investments require 5-10 year planning horizons. Every quarter of regulatory delay widens the gap. Any copyright policy position must be assessed against this competitive reality.

The German Coalition Agreement

The German coalition agreement states: *Authors must be fairly remunerated for the use of their works that are necessarily used in the development of generative AI* (ll. 2826-2827).
This passage should be interpreted in light of current European-level developments and discussions. The coalition agreement does not call for new remuneration obligations or levies. It recognises the existing legal framework. Voluntary AI partnerships are already enabling creators to benefit commercially through voluntary licensing under the current TDM framework—without new legislation.

We emphasise that the coalition agreement refers explicitly to works that are **»necessarily used«** in the development of generative AI: a deliberate and narrow wording that implies a specific, identifiable selection — not a statistical aggregate of everything publicly accessible. Given the scale of data used to train foundation models, it cannot be assumed that each individual work collected during web crawling was »necessary« in this sense. No individual work in a Common Crawl snapshot was necessary; the aggregate was. Any future policy measures must respect this threshold and should not be used to retrospectively alter the balanced framework established under the text and data mining (TDM) exception. A remuneration obligation keyed to that phrase cannot attach to web-scale training without stretching it past its plain meaning. The coalition agreement's mandate calls for legal clarity: consolidating the CDSM exceptions into practical guidance, not a new national layer above the EU floor.

TDM

There is broad agreement that the TDM exception (Article 4 DSM) applies to AI training, as acknowledged by the Hamburg Court Robert Kneschke v. LAION e. V. The issue is expected to be clarified by the CJEU in *Like Company v. Google (C-250/25)*. This paper takes the applicability of Article 4 DSM as the basis for the discussion that follows in line with how former Commissioner Breton, on behalf of the European Commission ↗clarified this point during the AI Act negotiations in 2023, namely that »[the TDM] exceptions provide balance between the protection of rightholders including artists and the facilitation of TDM, including by AI developers.«

This existing framework strikes an appropriate balance: it permits innovation-enabling uses of publicly accessible content while giving rightsholders a mechanism to reserve their rights. Reopening this settled framework risks creating years of legal uncertainty that would harm European AI developers — including SMEs and start-ups — without delivering meaningful additional protection for creators.

Principle of territoriality

This paper also highlights that the **principle of territoriality is ingrained in international copyright law**. All measures under discussion must therefore be assessed in light of territoriality and the potential consequences for AI development in Europe. Proposals that seek to apply EU copyright rules to training conducted in third countries risk conflicting with established international norms and could provoke retaliatory measures or regulatory fragmentation globally. Extraterritoriality would also be highly unworkable from a practical perspective for AI model developers who objectively can't feasibly comply with several inconsistent laws for the same act of training.

Misleading wording: »Memorisation«

Claims about »memorisation« have become a recurring feature of debates on generative AI and copyright. It is therefore important to clarify what the term does — and does not — mean in this context. In particular, the suggestion that an LLM contains internal copyright-relevant copies of protected works is often used to characterise how these systems function and to support claims to additional rights.

»Memorisation« is a misleading notion as a catch-all description of technical mechanisms. The term suggests there is a direct »copy« stored inside the model. This is inaccurate: generative AI learns statistical patterns represented within the training material and produces output based on probabilistic prediction. The term originates from computer science, where it has a narrow technical meaning — referring to rare cases in which models reproduce verbatim or near-verbatim fragments of training data — but when carried into copyright discourse, the metaphor has been misinterpreted and conflated with the concept of reproduction.

Rare instances of regurgitation are statistical, unwanted anomalies with today's models, not functional features, and require adversarial prompting. Verbatim or near-verbatim reproduction of a specific work in output is materially different from the statistical pattern-learning that characterises ordinary training. The former may give rise to infringement; the latter does not. Conflating the two under the label of »memorisation« within the model layer risks conceding more than copyright doctrine is likely to require. A more defensible threshold is therefore verbatim or near-verbatim reproduction of protected expression in output, rather than the abstract suggestion that a model »contains« copies of works. Recent case law, including GEMA v. OpenAI and Getty Images v Stability AI, shows that courts have approached this issue differently.

Output is therefore better understood primarily as a question of liability rather than as a question of the copyrightability of the training process itself. The relevant question is therefore not whether the training process itself is infringing, but whether particular outputs give rise to liability in specific contexts. Any such assessment should focus on the circumstances in which outputs are generated and used, taking into account user behaviour.

Statement on Current Ideas Within the Existing Legal Framework

As stated above, this paper is based on the legal understanding that the TDM exception applies to generative AI training.

Under this approach, TDM for the purpose of generative AI training is generally permitted unless the rightsholder has opted out using a machine-readable standard.

Managing Access and Opt-Outs

The TDM exception relies centrally on the ability of rightsholders to reserve their rights by doing so in a machine-readable format. Clarity is needed on what constitutes a »machine-readable« reservation of rights, a question that is discussed in international standardisation bodies. It should be noted that sophisticated technical solutions already exist today that move beyond binary opt-out declarations to provide granular, verifiable control over data access. These solutions enable rightsholders to manage crawling activity in real time – either by preventing unauthorised scraping altogether or by facilitating access under specific commercial conditions. To promote broad effectiveness, these technical solutions should be translated into industry standards for the expression of granular TDM preferences, bot transparency and authentication.

Major AI developers already deploy separate crawlers for AI training, allowing website operators to block AI-specific access while maintaining visibility for other services.

Format of a Machine-Readable Opt-Out

A key question is how opt-outs can be expressed. **Robots.txt** remains the most scalable and widely adopted de facto standard: Independent studies such as the 2024 Web Almanac found that robots.txt is used by the vast majority (almost 84 percent) of the sites it measured and that some sites increasingly include rules for AI crawlers such as GPTBot. However, it was not originally designed in 1994 for the complexities of generative AI and offers only a binary choice between allowing or disallowing access. Recently, additional standards have been proposed. The AI Preferences Working Group of the IETF (Internet Engineering Task Force, the international standards body for internet technologies) is in the process of developing a standardised vocabulary for expressing AI-related preferences and a description of attachment mechanisms for content, including, but not limited to, the use of robots.txt and HTTP response header fields. However, commercial TDM users caution that multiplying opt-out channels risks fragmenting compliance and reducing legal certainty. Robots.txt is widely used, easy to implement, highly interoperable and internationally recognised; it is also increasingly referenced in emerging AI governance expectations. Questions about whether the opt-out has been respected are often raised; see the section on transparency below.

Interoperability matters here as much as the legal rule. A widely adopted, internationally recognised, machine-readable signal, such as robots.txt, produces a predictable compliance environment for developers and rightsholders. Proliferating bespoke or jurisdiction-specific opt-out channels, without broad industry adoption through consensus-based standardisation, would fragment the European data environment, raise costs for SMEs and weaken Europe's position relative to jurisdictions with clearer rules. The TDM exception should therefore be viewed in light of these international developments.

Opt-Out Registry

Some stakeholders propose allowing rightsholders to register opt-outs in a centralised registry, for example via the EUIPO. In our view, such a registry would likely be overly bureaucratic. A registry adds cost and complexity without solving the enforcement problem it purports to address. For technology companies, a voluntary registry would only be partially helpful: registered opt-outs would be clear, but the absence of a registration would carry no legal certainty—developers could not safely conclude that unregistered content is available for use.

Proposals for a central opt-out registry also raise concerns: (1) high implementation and maintenance costs for all parties; (2) practical difficulties in keeping an EU-wide repository accurate and fully up to date, given the scale, frequent rights changes and granular reuse possibilities; and (3) the likelihood that parallel systems (a registry plus robots.txt) would complicate rather than simplify compliance—especially for SMEs—by requiring the reconciliation of multiple, potentially conflicting signals. Registries also create single points of failure and, without robust identity verification, are vulnerable to potentially false opt-out declarations.

In addition, a registry may create risks of legal misinterpretation: it could encourage an overly broad understanding of how opt-outs may be expressed, potentially diluting the »machine-readable« requirement in Article 4(3) DSM and encouraging impractical or non-standard declarations (e.g. unilateral website statements or email requests). Finally, even robust signalling cannot fully prevent deliberate non-compliance by bad actors, and a registry could introduce new vulnerabilities (e.g. if depositing precise file sets could be misused to assemble high-quality datasets). A registry would, in effect, layer a new compliance regime on top of a machine-readable standard, adding cost and complexity without delivering proportionate benefits to rightsholders and disproportionately burdening start-ups and SMEs.

Transparency

On transparency, it should be noted that discussions are taking place alongside the newly agreed General-Purpose AI Code of Practice and the template under the AI Act. Calls for »full and detailed transparency concerning all copyright-protected content used to train the system, irrespective of jurisdiction« run into the same challenges discussed during the Code of Practice negotiations and the broader AI Act debates. For example, for some third-party datasets, it is not feasible to provide itemised transparency; in other cases, disclosure could compromise trade secrets and potentially compromise security. Any transparency obligation must respect trade secrets and avoid competitive disadvantages. It should be noted that public datasets are especially important for SMEs, as they are still a key resource for LLMs. Projects like OpenEuroLLM, for example, also rely on third-party datasets. Rightsholders can request deletion from datasets, for example from Common Crawl, one of the most widely used datasets.

Overly burdensome compliance requirements may pose significant challenges, particularly for SMEs and innovation-driven start-ups, which are crucial to Europe's economic landscape. Statements claiming that transparency should be granted »regardless of jurisdiction« overlook the principle of territoriality and risk putting European companies at a disadvantage.

The template, the Code of Practice and the AI Act must be considered together. Their effectiveness cannot yet be fully assessed, as these measures only took effect in August 2025. Demanding changes before these rules have even had time to prove their effectiveness undermines the regulatory process.

Transparency obligations should be designed in a risk-based and proportionate manner, be interoperable with the General-Purpose AI Code of Practice and the AI Act template, and should not be expanded before the existing framework has had time to operate and be evaluated. Calls for additional transparency obligations—particularly extraterritorial ones—would layer yet more compliance costs on European AI companies and start-ups, widening the competitive gap with the US and China. A premature recalibration would create legal uncertainty, encourage firms to relocate development activities outside the EU and run directly counter to the Draghi competitiveness agenda.

Article 53(1)(d) of the AI Act requires providers of general-purpose AI models to publish a **»sufficiently detailed summary«** of the content used for training, following a mandatory template issued by the European Commission's AI Office.

The summary must include:

- **Categories of data sources** — public datasets, licensed datasets, crawled/scraped web content, user-submitted data, synthetic data, and other sources
- **Approximate scale by modality** — token ranges for text, approximate image/video/audio counts
- **Top contributing domains** — for scraped content, providers must list the main websites or domain groups that contributed to the dataset (to the extent feasible)
- **Licensing and rights information** — whether licensing agreements exist for private datasets
- **Text-and-data-mining compliance measures** — how the provider identified and respected opt-outs under Article 4(3) of the DSM Directive

The intent is to provide rightsholders with sufficient information to assess whether their content could plausibly have been included, without requiring disclosure of full datasets (which would raise trade secret and security concerns). This is important and should be introduced into daily business with proportionate effort. Maintaining this requirement remains essential because only meaningful transparency enables rightsholders to realistically assess whether their protected content has been used and to exercise their rights effectively. The obligation is proportionate, as it requires aggregated, high-level information rather than disclosure of individual works or trade secrets, thereby balancing transparency with business interests. Given the scale and impact of general-purpose AI models, a reasonable compliance effort is justified to ensure accountability and uphold existing copyright and data mining safeguards.

We see a lack of understanding when it comes to these mechanisms, rightsholders must:

1. **Review the published summary**

   Check whether your domain, platform, or data source appears in the provider's listed categories or top-domain disclosures. If a provider states they scraped large portions of the open web and your content was publicly accessible without a machine-readable opt-out, there's a reasonable likelihood it was ingested.

2. **Cross-reference known datasets**

   Many GPAI models are trained on well-documented datasets (Common Crawl, The Pile, LAION, etc.). If the summary names specific datasets, rightholders can consult third-party tools and indices that catalogue the contents of those datasets.

3. **Request further information from authorities**

   The AI Office and national competent authorities can request the full technical documentation (which contains more granular detail than the public summary) when investigating complaints. Rightholders can lodge a complaint or qualified alert, prompting regulatory scrutiny.

4. **Use membership inference or output probing (emerging practice).**

Trusted Intermediary

Some suggest that an institution like the EUIPO should act as a trusted intermediary to check whether a work has been used in the training of an AI model.

To verify whether a specific work was used, the EUIPO would need access to information that is often **more sensitive than anything companies could publish**, such as:

- detailed source lists (URLs, vendor datasets),
- hashed identifiers, deduplication logs and filtering rules,
- training runs, checkpoints and lineage records,
- vendor contracts and licensing constraints.

This proposal does not eliminate the disclosure problem; it **shifts it** from »public disclosure« to »regulator/intermediary access.« From a business perspective, the risk to trade secrets remains (e.g., exposure through oversight processes, FOI-type access issues, cybersecurity risk, or downstream disputes). The AI Act reflects this tension by framing documentation and information-sharing duties alongside protections for IP, confidential business information, and trade secrets. These challenges do not mean that such a mechanism would be impossible, but rather that they would have to be addressed.

Any trusted-intermediary mechanism would have to be designed according to security-by-design principles and include robust safeguards for confidential business information and trade secrets. Concentrating sensitive training-data information within a single regulator or third-party body creates new attack surfaces and supply-chain risks; the policy benefit must be weighed against those security and competitiveness costs.

(Ir)rebuttable Presumption

The Voss Report calls for a rebuttable presumption that content has been crawled, including for inference and retrieval-augmented generation purposes, where AI systems handle user queries. Other suggestions, such as a recent French proposal, call for an irrebuttable presumption in cases of non-compliance with EU transparency requirements regarding training data: it would be automatically presumed that copyrighted material was used, without the possibility of rebuttal.

Such a rule would shift the burden of proof in a way that, under established legal principles, is permissible only in very narrow exceptional cases. An irrebuttable presumption would leave no legal basis for considering contrary evidence. Decisions should not be based solely on such a presumption without proper examination of the facts. Furthermore, legal costs and expenses should continue to be allocated according to applicable **national civil procedure rules**, rather than being imposed automatically on AI providers.

Both an irrebuttable and a rebuttable presumption are technically unworkable: copyright in the EU is unregistered, AI models train on billions of data points, and no complete index of online content exists. Such a rule would effectively block AI research and development in Europe — the exact opposite of Europe's sovereignty goals. France's Darcos bill, which creates a similar presumption, has already been cited by investors as a reason to redirect capital away from European AI.

Help the licensing market?

In our view, there is no clear need to intervene in the licensing market at this stage, as licensing and partnership activity is already growing. Markets are emerging for access to specific restricted datasets, such as data that exists behind paywalls. These markets play an important role in facilitating access to specialised datasets and are likely to expand further as the AI ecosystem develops. Demand for specific data exists, but it is often highly specialised. At the same time, licensing cannot be the sole mechanism for accessing data, as it does not scale to the volume and diversity of inputs required for modern AI systems. Instead, licensing frameworks should operate alongside a functioning TDM exception, enabling both broad analytical use of lawfully accessible data and targeted access to restricted datasets where licensing is appropriate. Some stakeholders advocating for licensing agreements do not necessarily offer the kinds of data typically required for model training.

Calls for adequate remuneration for works that were **necessary** for AI training (e.g. under the German coalition agreement) should be understood in the context of the existing TDM framework, which permits the use of lawfully accessible works for analytical purposes. Market-driven initiatives and companies can already make use of voluntary collective licensing arrangements as a flexible, scalable tool to accurately value contributions and secure legal certainty. Specifically, voluntary collective frameworks can help determine fair value by establishing objective quality tiers, applying contribution-based distribution formulas and resolving the high transaction costs of tracking individual micro-contributions—provided these mechanisms are technically viable and commercially feasible for both developers and rightsholders. Public policy solutions that continue to allow those arrangements to thrive can benefit the entire market. Given the scale of data used to train foundation models, it cannot be assumed that each individual work collected during crawling was »necessary« in that sense. Companies can already make use of voluntary collective licensing where feasible, i.e. by offering their copyrighted content on a voluntary basis as a collective to AI companies. Where licensing is appropriate, such as for access to restricted datasets, markets are already emerging to facilitate this. However, requiring remuneration at the level of all training data would risk undermining the functioning of the TDM exception and introducing a de facto »permission-to-compute« model, creating friction, cost and legal uncertainty at the data-analysis layer, with knock-on effects across strategically important sectors, including cybersecurity, healthcare and research.

Voluntary AI partnerships are already scaling under the current EUCD framework and extend beyond remuneration for training to include output aspects not connected to copyright law. Multi-year partnerships between major AI companies and European publishers have been negotiated under the existing framework, with no mandatory licensing and no levy. Premature intervention would replace these voluntary, market-rate arrangements with administered, below-market terms and throw the commercial market into years of uncertainty.

AI policy should reject mandatory remuneration schemes or levy systems for lawful text and data mining and instead uphold full contractual freedom in licensing. This means no imposed mediation, arbitration or standard contractual clauses—allowing parties to negotiate terms independently and on a voluntary basis.

Statements on Current Ideas Outside the Existing Legal Framework

Remuneration Obligation

Outside the existing framework, different remuneration models are being discussed.

As discussed above, this paper reflects the understanding that the use of publicly accessible content for training generative AI models takes place within existing copyright exceptions, in particular the TDM exception under Article 4 of the CDSM Directive. This provision allows rightsholders to opt out via a machine-readable mechanism. This system offers a balanced approach between protecting creative content and enabling innovation. Introducing additional authorisation or remuneration requirements would create significant burdens for smaller providers and disproportionately restrict access to publicly available information.

Introducing additional authorisation or remuneration requirements on top of the TDM framework would: (i) create significant burdens for SMEs and start-ups, which cannot sustain complex licensing operations; (ii) disproportionately restrict access to publicly available information that underpins data analytics across the economy; (iii) increase fragmentation across Member States; and (iv) encourage firms to develop and train systems outside the EU. Each of these outcomes runs directly counter to the Draghi competitiveness agenda.

Mandated licensing would affect model quality: any obligation imposed on developers to license one class of content—in instances where collecting societies govern the necessary economic rights in question—could create an incentive for developers not to train on those classes of works and to restrict training to works that do not require a licence (e.g. works in the public domain, open-source works or non-copyrightable material). This would have a negative impact on the development and quality of multimodal models.

It must be ensured that innovative technologies retain access to the data necessary to advance AI development. A premature remuneration obligation would significantly hinder technological progress, especially for SMEs. Existing regulations should be reviewed and implemented through a balanced process that adequately considers both the creative industries and the AI sector.

Some propose a **levy system**, similar to mechanisms used in other areas. However, a levy modelled on traditional private-copy schemes is unlikely to be future-proof for AI training. Levies would not only have an extraterritorial effect in the context of AI training but would also ignore the fact that we have a working and balanced system in place. Furthermore, they would add years of legal uncertainty and drive AI investment away from Europe. Such schemes can be blunt instruments: they often generate high administrative and litigation costs, offer limited transparency regarding allocation, and can lead to »double payment« where content is already licensed through existing channels. Applied to AI, levies are also poorly targeted because training practices, data sources and uses vary widely, while AI development is inherently cross-border—raising risks of fragmentation, compliance complexity and competitive disadvantages. Policy should therefore prioritise scalable licensing pathways, clear rules on data provenance and rights information, and enforceable transparency and redress mechanisms, as provided for under the AI Act and existing procedural law, rather than extending legacy levy models to a fundamentally different context. Compensation should be closely aligned with the actual value and significance of the licensed content. Rightsholders should be remunerated fairly and proportionately within the TDM framework, reflecting the qualitative contribution of their works to the training process, rather than through rigid, blanket levy systems. Market-driven initiatives and companies can already make use of voluntary collective licensing arrangements as a flexible, scalable tool to accurately value contributions and secure legal certainty.

More broadly, the Commission's initiative on the optional scientific research exception under the InfoSoc Directive should prompt a wider assessment of whether the Directive's exceptions remain effective in today's digital environment. That assessment should also revisit the private copying exception in Article 5(2)(b), where divergent national levy regimes create legal uncertainty, administrative burdens and risks of internal-market fragmentation, and where the concepts of »private copying«, actual harm and fair compensation require renewed scrutiny. Targeted infringement remedies may be more precise than general remuneration schemes, particularly in cases where memorisation results in verbatim or near-verbatim reproduction.

This reflects a fundamental principle of copyright law: the scope of copyright is limited to the expression of an idea, not the underlying ideas, facts or concepts themselves, nor the ability to derive insights by identifying patterns across large datasets. TDM operates precisely at this level, extracting statistical and analytical insights rather than substituting for expressive works. Introducing a broad remuneration obligation risks extending copyright beyond its intended scope while delivering limited economic benefit when dispersed across large numbers of authors. By contrast, a carefully scoped TDM exception supports the development of a thriving AI ecosystem, enabling new business models, including those for the content industries, for example by powering innovative forms of content creation and distribution.

More fundamentally, layering new remuneration obligations on top of lawful analysis would treat a broadly employed economic activity—reading data with machines—as a licensing event. That is the wrong design choice for a competitiveness agenda built on diffusing digital tools across the economy.

What we are not arguing: Preserving a workable TDM framework is not a call to weaken copyright protection, nor is it a demand for free access to paywalled or otherwise restricted content. It is the preservation of a clear, long-standing principle: the analysis of lawfully accessed information is not, in itself, an infringement. Creators retain their rights, their remedies and, through the machine-readable opt-out, meaningful control over whether their content is used for AI training. This control is further reinforced by the AI Act's transparency obligations, which require clear disclosure of AI training practices and thereby strengthen the practical enforceability of rightsholders' opt-out choices.

Output

Some stakeholders argue that outputs should be subject to clearer guardrails. Proposals include: clarifying that purely AI-generated material that does not meet the EU »work« threshold should not receive copyright protection; reducing infringement risks from outputs that reproduce protected content; addressing harmful deepfake or digital replica outputs that imitate a person's likeness or voice without consent; requiring services to act against illegal dissemination; and introducing labelling requirements for purely AI-generated content.

While additional GenAI-specific guardrails may appear attractive, they risk duplicating—and potentially confusing—frameworks that already exist in EU law. Copyright infringements in outputs are not a regulatory »gap«: rightsholders already have civil remedies (injunctions, corrective measures, damages, evidence preservation tools and cost recovery) under the IP Enforcement Directive, designed to be effective, proportionate and dissuasive. At the platform layer, the Digital Services Act requires hosting services to offer user-friendly electronic notice-and-action channels and to process notices diligently and in a timely manner, enabling the swift takedown of illegal content (including copyright-infringing material) without introducing a separate »GenAI output« regime. Transparency and provenance are also being addressed horizontally: Article 50 of the EU AI Act imposes transparency duties, including disclosure obligations for deepfakes and requirements related to detectability and machine-readable marking for synthetic content. Finally, rigid bright-line rules (e.g.

blanket public-domain declarations for »purely AI« outputs) risk imposing overinclusive burdens and creating chilling effects on legitimate uses. In practice, where AI tools are used within human creative workflows, copyright's »work« assessment will remain fact-sensitive and will need to be conducted on a case-by-case basis. Guidelines and codes of practice are helpful in encouraging industry best practice to mitigate the risk of problematic outputs. The same principle applies here as elsewhere in this paper: a risk-based, innovation-friendly approach that builds on existing horizontal frameworks is preferable to bespoke, GenAI-specific regimes that risk fragmentation and chilling effects on legitimate uses.

The priority should be the effective enforcement of existing instruments, not the creation of overlapping regulatory layers that increase legal uncertainty without delivering additional protection.

Copyright Protection for Individual Likeness

Individual likeness is not properly a matter of copyright. It is more naturally protected through personality rights, image rights and related rules protecting personal autonomy, dignity and control over the use of identifiable features such as a person's face or voice. In Germany, this is often described as the »general right of personality« (allgemeines Persönlichkeitsrecht) and the right to one's own image, but the underlying concept is not unique to German law.

That is also why the Commission's February 2026 assessment of Denmark's notified proposal (2025/0654/DK) is so important: according to the Commission's view, copyright and related rights are designed to protect creative subject matter, not personal characteristics such as appearance, voice or movements, and using copyright concepts like »making available to the public« for likeness protection creates a clear conceptual mismatch with EU copyright law.

In that sense, the Commission effectively confirmed that copyright has nothing to do with likeness protection as such. The relevant legal concerns arise from the misuse of identity or personal characteristics, not from the exploitation of an intellectual creation.

Call to Action

The stakes are clear. Ongoing regulatory uncertainty directly shapes investment decisions—where data centres are built, where startups scale, and where supply chains take root. Bitkom therefore calls on policymakers to:

1. **Preserve the EUCD's existing TDM opt-out framework;**
2. **Reject presumption-of-use proposals;**
3. **Prevent unilateral national copyright bills that fragment the Single Market;**
4. **Ensure that AI Act transparency requirements do not exceed what was agreed in the Code of Practice.**

Bitkom represents more than 2,200 companies from the digital economy. They generate an annual turnover of 200 billion euros in Germany and employ more than 2 million people. Among the members are 1,000 small and medium-sized businesses, over 700 start-ups and almost all global players. These companies provide services in software, IT, telecommunications or the internet, produce hardware and consumer electronics, work in digital media, create content, operate platforms or are in other ways affiliated with the digital economy. 82 percent of the members' headquarters are in Germany, 8 percent in the rest of the EU and 7 percent in the US. 3 percent are from other regions of the world. Bitkom promotes and drives the digital transformation of the German economy and advocates for citizens to participate in and benefit from digitalisation. At the heart of Bitkom's concerns are ensuring a strong European digital policy and a fully integrated digital single market, as well as making Germany a key driver of digital change in Europe and the world.

Published by
Bitkom e.V.
Albrechtstr. 10 | 10117 Berlin

Contact person
Friederike Michael | Policy Officer Digital Content & Law
P +49 30 27576-099 | f.michael@bitkom.org

Responsible Bitkom Committee
WG Intellectual Property