
The Battle for Data

An Economic Analysis of Law Across Different Regulatory Approaches to Text and Data Mining for AI Training

Bartłomiej Biga, PhD

Associate Professor at Krakow University of Economics

Text and data mining (TDM) has shed its reputation as a niche technical concern and emerged as a strategic inflexion point in the global technology race.¹ As the *Artificial Intelligence Index Report 2025* makes clear, AI is exerting ever-greater influence on economies and global governance² — and underpinning that influence is access to high-quality information resources.

Legal restrictions on TDM directly undermine this foundation. They render the curation and digitisation of local knowledge archives either legally impossible or prohibitively risky. The consequences are stark: countries that erect restrictive barriers forfeit their opportunity to develop AI models capable of understanding their own cultural and linguistic contexts. At the same time, protecting creators' rights must be an integral element of any balanced legal framework. Consequently, compromise solutions are required that enable the development of AI while ensuring that creators preserve their rights.

The opt-out mechanism adopted in the European Union provides a balanced approach, although it does not capture the maximum potential gross domestic product (GDP) gains a fully unconstrained TDM exception could deliver. On one hand, it allows the AI sector to operate without major barriers to data access such as the need to obtain licenses for every single piece of content (opt-in) or a rebuttable presumption of use of copyrighted content. On the other hand, it allows rights holders to retain control over their work, something that would be impossible under the most liberal frameworks, where TDM would be subject to no restrictions at all. However, to be effective, an opt-out mechanism must minimise uncertainty, with clear criteria establishing when a creator's reservation of rights is considered valid and enforceable. In practice, this requires standardised, machine-readable reservations.

1 This policy brief is the result of a six-month research project carried out by Prof. Biga for the Lisbon Council, under the coordination of Caroline de Cock, and financially supported by Meta Platforms, Inc. Any errors of fact or judgement are the lead author's sole responsibility.

2 Nestor Maslej, Loredana Fattorini, Raymond Perrault, Yolanda Gil, Vanessa Parli, Njenga Kariuki et al., "Artificial Intelligence Index Report 2025," *arXiv*, 20 November 2025.

Framing the regulatory choice as a search for “balance” may, however, obscure the true nature of the problem. In essence, regulating access to data for TDM purposes should be thought of in terms of trade-offs rather than equilibrium: easier access to training data results in smaller direct benefits for data owners, and vice versa. There is no regulatory configuration that maximises both values simultaneously — every arrangement distributes gains and losses between the AI sector and rights holders in a different proportion. The key question, therefore, is not how to reconcile the two sides, but whether it is worth sacrificing a significant share of the additional economic benefits that unrestricted AI development could be expected to deliver in the name of strengthening the position of creators. When assessing this trade-off, it is important to recognise that creators’ rights do not disappear even under

‘Regulating access for TDM purposes should be thought of in terms of trade-offs rather than equilibrium.’

liberal TDM regimes: copyright protection continues to operate at the output stage, applying to what a trained AI system generates. The input stage, i.e., where works are used in the training process, is thus not the only, nor necessarily the most effective, point at which the interests of rights holders can be safeguarded.

Beyond the structural merits of the opt-out mechanism, a critical argument for retaining it lies in the principle of regulatory continuity. European AI developers have already committed significant capital and designed complex operational workflows based on the current legal framework established by the CDSM directive. Altering the legal basis for TDM mid-deployment — for instance, by transitioning towards a mandatory licensing regime — would introduce significant regulatory uncertainty. That uncertainty would itself impose a distinct innovation cost, independent of the specific technical requirements of any new framework. For firms navigating the capital-intensive cycle of model training, such a shift would represent a “regulatory shock” that could devalue existing investments and chill future development. This could potentially drive talent and capital towards more stable jurisdictions just as the European AI ecosystem begins to mature. Even the *prospect* of regulatory change creates prolonged uncertainty: any change to the European Union’s current TDM framework would require a full legislative cycle, including trilogue negotiations and subsequent transposition across all member states. The resulting changes would therefore take considerable time to enter into force. For firms operating on multi-year model-training timelines, this implies a sustained period in which the rules of the game remain unsettled.

This report sets out to conduct an economic analysis of law across the different regulatory approaches to TDM for AI training purposes. More specifically, it develops a preliminary research framework for assessing the impact of an opt-out-based approach on the European Union’s economy, capacity for innovation, creative sectors and broader public access to knowledge. The analysis is comparative in nature, drawing on the following jurisdictions: the United States, Japan, Singapore, Germany, Italy and Poland, with particular focus on the European Union member states. The report looks exclusively at the input stage, i.e., the acts involved in training AI models. The question of whether TDM constitutes a standard field of exploitation, and what implications that carries for AI-generated outputs, falls outside its scope.

Key Findings:

- Among regulatory models that respect creators' rights, the opt-out model delivers the most balanced economic outcome: in the central scenario, this approach preserves €942 billion in potential additional Gross Domestic Product (GDP) in the European Union by 2035 compared to an opt-in regime that would substantially slow down development of artificial intelligence (AI). This represents 79% of the €1.2 trillion potential additional GDP a fully unconstrained text and data mining (TDM) exception would deliver by 2035. The remaining €258 billion constitutes the potential GDP gains lost under opt-out restrictions — the quantifiable price of maintaining creators' reservation rights.³
- Grounding European Union TDM regulation in an opt-out mechanism represents a middle ground between near-unrestricted data access (as seen in Asian models) and a mandatory prior-licensing requirement (opt-in). This balanced approach is enshrined in article 4 of the copyright in the digital single market (CDSM) directive.⁴ It enables AI development without paralysing firms with barriers to entry, while still respecting the agency of rights holders.
- From the perspective of the economic analysis of law, the key advantage of a well-designed opt-out mechanism over an opt-in model is a radical reduction in transaction costs. Under an opt-out model, the reservation procedure must be unambiguous — leaving as little doubt as possible about whether a reservation is valid — and automated, meaning machine-readable. However, the frictions an opt-out model introduces require rigorous scrutiny, particularly in terms of transaction costs and the impact of such a model on AI development in the European Union, including the risk of permanently locking certain actors out of the innovation ecosystem.
- Excessive tightening of the system — for example, through a universal licensing requirement (opt-in model) — risks creating a situation in which every piece of content requires authorisation. This creates too many veto-holders, making any effective exploitation of data resources impossible and resulting in the suboptimal use of a key economic resource. In economics, this outcome is known as the tragedy of the anticommons.
- Given the capital intensity of the AI sector, the economically sound approach is to first create the conditions for mature, profitable products to emerge. Introducing mandatory fees at an early stage of technological development via changes to European law may not only stifle innovation and reduce GDP but, paradoxically, deprive creators of any future revenue by undermining the viability of the entire ecosystem.
- Technology capital is highly mobile. Legislators must therefore consider the risk that stringent TDM restrictions will trigger the relocation of firms to more permissive jurisdictions, thus harming European players and the European economy in particular. While some may frame this as a race to the bottom, the ease of relocating data-intensive activities means that policymakers cannot realistically ignore these competitive pressures.

'Stringent TDM restrictions will trigger the relocation of firms to more permissive jurisdictions.'

³ For more detail on the economic analysis see page 18.

⁴ Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market and Amending Directives 96/9/EC and 2001/29/EC (Text with EEA Relevance).

- Restrictions on TDM freedom carry real economic costs, as quantified throughout this report. The key economically relevant question is whether these costs are proportionate to the values they protect, which include concern for creative sectors, respect for property rights, and considerations of fairness.
- For Germany, opt-out in the most probable scenario means retaining 59% of the maximum potential AI benefits — i.e., approximately €167.5 billion out of a fully unconstrained ceiling of €283.9 billion. For Italy, this figure stands at 79%, amounting to €113.1 billion out of a maximum of €143.2 billion, while for Poland it reaches 91%, translating into €53.1 billion out of a fully unrestricted maximum of €58.4 billion in additional GDP by 2035.
- The ultimate decision on the appropriate level of copyright protection — and its cost in terms of GDP dynamism — is therefore political and philosophical in nature. From an economic standpoint, however, the maximum potential GDP gain would be delivered via a fully unconstrained TDM exception while opt-out is the model that allows Europe to currently avoid the scenario of a complete stagnation of AI sector development.

Regulatory Mapping

When it comes to the regulation of TDM, the countries under analysis fall broadly into two groups: those with a liberal approach (the United States, Japan and Singapore) and those that have adopted the more restrictive opt-out mechanism (Germany, Italy and Poland). Within each group, however, there are meaningful differences between individual jurisdictions — differences that can have significant economic consequences. It is worth emphasising that the regulatory landscape is evolving rapidly and as such, the classifications used here should be understood as simplified snapshots of the current situation rather than fixed positions. The “liberal” classification of the U.S. regime, in particular, is contingent on the outcome of ongoing litigation: depending on how courts ultimately interpret the fair use doctrine, significant parts of the AI training process may fall outside its protective scope, materially altering the comparative picture presented below.

Table 1. Comparison of regulations for all-purpose TDM

Country	Current TDM legal frameworks	Opt-out requirement	Permitted commercial use	Planned directions for further regulation	Impact on innovation and creators
United States	Fair use within U.S. copyright law	No	Yes, provided it meets (flexible) fair use criteria	Policy assuming minimal regulations (restrictions) regarding AI	Favours rapid technological development, with creator control determined case-by-case through fair use criteria rather than by default
Japan	Art. 30-4 of Japanese copyright law (broad exception for TDM)	No	Yes	Monitoring impact on the creative content market; debate on creator protection	Beneficial for AI innovation; perceived as risky by rights holders
Singapore	Automatic analysis exception in Singapore's copyright act of 2021	No	Yes, provided access to the materials is lawful	Continuous development of guidelines within Singapore's smart nation strategy	Attracts tech companies thanks to legal clarity; limits creator claims
Germany	§44b & §60d of Germany's act on copyright and related rights (UrhG), through which the CDSM directive was implemented	Machine-readable opt-out for purposes other than scientific, including commercial uses; scientific research TDM under section 60d of the UrhG falls outside the opt-out mechanism and cannot be blocked by rights holders	Yes, provided access to the materials is lawful, unless an opt-out has been applied	The practical threshold for a valid machine-readable opt-out reservation is being established through litigation rather than statute. The Hamburg Court of Appeal (OLG Hamburg) ruled in December 2025 (<i>Kneschke v. LAION</i>) that a natural-language terms-of-service reservation does not satisfy §44b's machine-readability requirement. The case is on appeal to the federal court of justice (BGH)	Strives for balance, with a strong emphasis on interpretability and copyright
Italy	Art. 70-ter and 70-quater of Italy's copyright act (LdA), through which the CDSM directive was implemented in 2021; Art. 70-septies (introduced by law No. 132/2025, in force 10 October 2025), which expressly extends the TDM exception to AI training, including generative AI systems, subject to opt-out requirements; Art. 171-a-ter, which introduces criminal liability for TDM violations; amended Art. 1 limits copyright protection to works of human creativity	Opt-out does not have to be machine-readable	Yes, provided access to the materials is lawful, unless an opt-out has been applied	Ongoing supervisory alignment with EU AI act requirements; judicial and regulatory clarification expected on whether "extraction" under Art. 70-septies covers comprehensive LLM training or only narrower analytical uses; practical threshold for opt-out compliance uncertain given the absence of a statutory machine-readability definition in Italian law	Supports the protection of cultural heritage and creators; imposes unclear obligations on AI developers
Poland	Late implementation of the CDSM directive (Art. 3 & 4) into Poland's copyright act	Machine-readable opt-out for purposes other than scientific	Yes, provided access to the materials is lawful, unless an opt-out has been applied	In Poland the emerging direction is not a new statutory right but a market practice layered onto the individual opt-out: collective management organisations have begun issuing blanket reservations on behalf of represented rights holders. However, their validity under the machine-readable requirement remains unclear	Gives creators commercial control tools while supporting scientific research

Source: Author's own study

The U.S. model, while broadly permissive, leaves considerable room for judicial interpretation. The prevailing case law suggests that, provided access to the training data was lawful, TDM can proceed without the rights holder's consent — even when undertaken for commercial purposes. Some courts have carved out an exception where TDM might produce an output that directly competes in the market with an existing work. Overall, the U.S. framework creates an environment that enables relatively free AI development.

Japan, by contrast, takes freedom of exploitation as its starting point, allowing works to be used to the extent necessary for information analysis, provided the purpose is not “enjoyment.” In this context, “enjoyment” refers broadly to consuming content in a way that delivers the kind of satisfaction typically associated with engaging with copyright-protected works. The goal cannot be merely “personally enjoying, or enabling others to enjoy, the thoughts or feelings expressed in a work.” This approach made the Japanese model one of the most liberal in the world. In 2024, however, the framework was refined to clarify that this freedom does not extend to situations where: (1) the TDM is ultimately intended to deliver similar “enjoyment” to end users; (2) it aims to produce works closely imitating a specific author’s style; or (3) it knowingly relies on pirated content. The first of these conditions, in particular, may pose a substantial obstacle for AI companies, as the notion of “enjoyment” is broadly defined and difficult to engineer around at the model-design stage. The second condition also raises concerns, particularly for providers of consumer-facing generative AI tools susceptible of producing outputs similar to copyright-protected works. However, these risks can generally be addressed through guardrails at the output stage, such as refusing prompts that explicitly invoke a specific author’s name or filtering outputs whose similarity to training data exceeds a defined threshold. The third condition can be addressed through standard dataset hygiene to ensure that training data was lawfully obtained and is similar to a specification of the lawful access requirement.

In the wake of the adjustments to Japan’s approach, Singapore can now lay the strongest claim to the world’s most liberal approach to TDM. The country provides an exceptionally broad exception for Computational Data Analysis (CDA) — broad enough, in fact, that data owners cannot prohibit TDM through contractual means such as terms of service. The analytical phase, i.e., extracting statistical features, learning patterns and recognising objects, is entirely unrestricted. Some caution is nonetheless warranted at the output stage. A model must not produce outputs that are too faithful a reproduction of works used in training, in a way that would deliver an “enjoyment” experience comparable to the original. Clear limits therefore apply to outputs here as well, though they are not especially onerous: it is sufficient that the output not too closely resemble the work of any particular author.

Turning to the European Union member states shifts the focus squarely to the opt-out mechanism adopted across the European Union. Under this model, TDM is permitted without rights holders’ consent unless they have explicitly reserved their rights in machine-readable format. Crucially, rights holders cannot opt out of non-commercial TDM conducted for scientific research purposes. Within the implementation of the CDSM directive and in judicial practice, however, the three European Union jurisdictions under analysis diverge in some important respects. A detailed analysis of the differences is provided later in this report.

Economic Analysis of Law

A Context of Global Rivalry

At the outset of this economic analysis of law, one critical caveat must be stated plainly: AI is a global phenomenon. Any jurisdiction that introduces stringent TDM regulation must therefore consider the real possibility that activity will simply migrate to more permissive jurisdictions. This is a serious problem, as it gives rise to what is known as a “race to the bottom” — a dynamic in which individual countries compete to offer AI developers ever more favourable conditions, disregarding the associated costs, in a bid to remain more attractive than their rivals. Without global coordination — which is currently out of reach — there is no obvious escape from this dynamic.

Moreover, the data suggests that stricter AI regulation correlates with weaker economic outcomes.⁵ Demonstrating a causal relationship is admittedly far from straightforward. However, in the arena of superficial public debate, a country’s position in high-profile rankings can nevertheless exert sufficient pressure on policymakers to push regulation toward leniency. Extracts from the most significant rankings relating to the economic context of AI are included in the annex to this report.

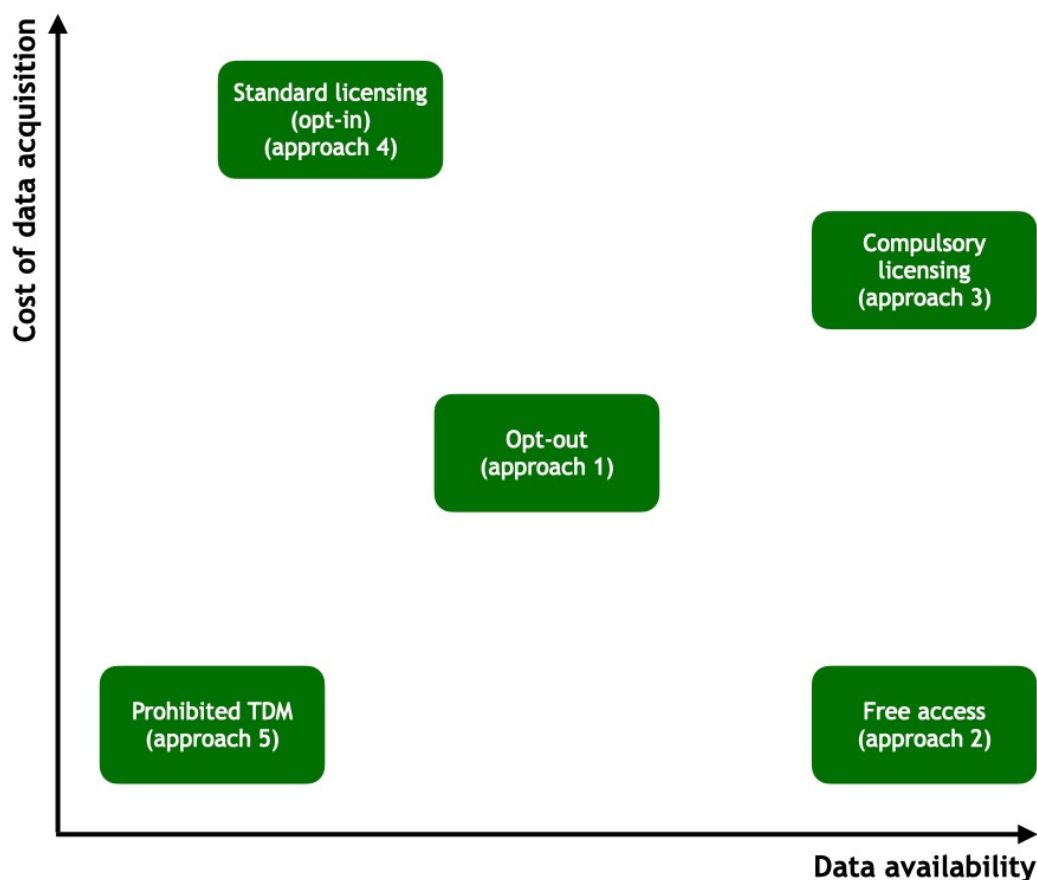
Two Core Dimensions of Analysis

Five regulatory approaches to TDM can be identified: the opt-out mechanism (1), free access (2), compulsory licensing (3), standard licensing requirements (opt-in) (4), and an outright prohibition on TDM (5). For the purposes of economic analysis, it is useful to assess these approaches along two dimensions: the availability of data for AI training and the cost of acquiring that data.

Figure 1 illustrates the different regulatory approaches and shows the opt-out mechanism as a political compromise that achieves an optimum in the economic sense by attempting to minimise the combined cost of two competing interests: data access for AI developers and reservation rights for creators. The free access model (approach 2) maximises data availability at zero acquisition cost, but it does so at the price of a complete external transfer of risk onto creators. Licensing models (approaches 3 and 4), by contrast, push us towards high transaction costs that paralyse innovation. The strategic positioning of opt-out at the centre of the graph illustrates an attempt to minimise the combined sum of two types of loss: access costs for AI developers and the costs of loss of control borne by creators. It is precisely this “centrality” that makes opt-out the most cost-efficient regulatory position available — one that addresses both extremes: the risks stemming from insufficient creator agency on the one hand and technological stagnation on the other.

⁵ Nestor Maslej, Loredana Fattorini, Raymond Perrault, Yolanda Gil, Vanessa Parli, Njenga Kariuki et al., “Artificial Intelligence Index Report 2025,” arXiv, 20 November 2025.

Figure 1: Availability of data for TDM and cost of acquisition in different regulatory approaches



Source: Author's own study

Opt-out (approach 1) represents a middle path. On one hand, it establishes a default presumption that data may be used for TDM. On the other, it introduces a straightforward mechanism for rights-holders to express their reservation of rights in machine-readable form. In terms of data availability, this approach's position along the X-axis may shift considerably depending on how easy the mechanism is to invoke and whether it is available in all circumstances. In terms of data acquisition costs, use within the exception is of course free of charge. However, some form of negotiated payments in exchange for not exercising an opt-out reservation are not uncommon. The position on the X-axis also reflects costs arising from uncertainty about whether any given dataset has been reserved.

'Too many veto-holders make any effective exploitation of data resources impossible.'

Free access (approach 2) is the most advantageous solution from the perspective of AI training companies. It permits the use of all data for TDM purposes without incurring any acquisition costs beyond purely technical ones. However, the analysis of TDM should not be conflated with the analysis of output generation: these are distinct legal phases, and rights holders retain a separate set of remedies at the model-release and output stages — increasingly operationalised through technical guardrails — regardless of how the TDM stage is regulated. The relevant question regarding approach 2 is therefore narrower than it might initially appear: whether the input/TDM stage should be left entirely unconstrained. Even with this narrower framing, eliminating the opt-out mechanism completely raises substantive

concerns. While output-stage controls remain available, removing creators' ability to reserve their works ex ante from training corpora deprives them of any means of expressing dissent before their content is ingested at scale. From an economic standpoint, this may give rise to "productive base erosion:" where creators have no means of influencing how their work is used in AI training, the standard incentive logic of intellectual property — that the prospect of returns motivates the production of high-quality content — is weakened. If this weakening is material, the quality and volume of content on which AI systems depend could diminish, creating a paradox whereby the productivity gains from unconstrained data access are partly offset by a decline in content quality over time.

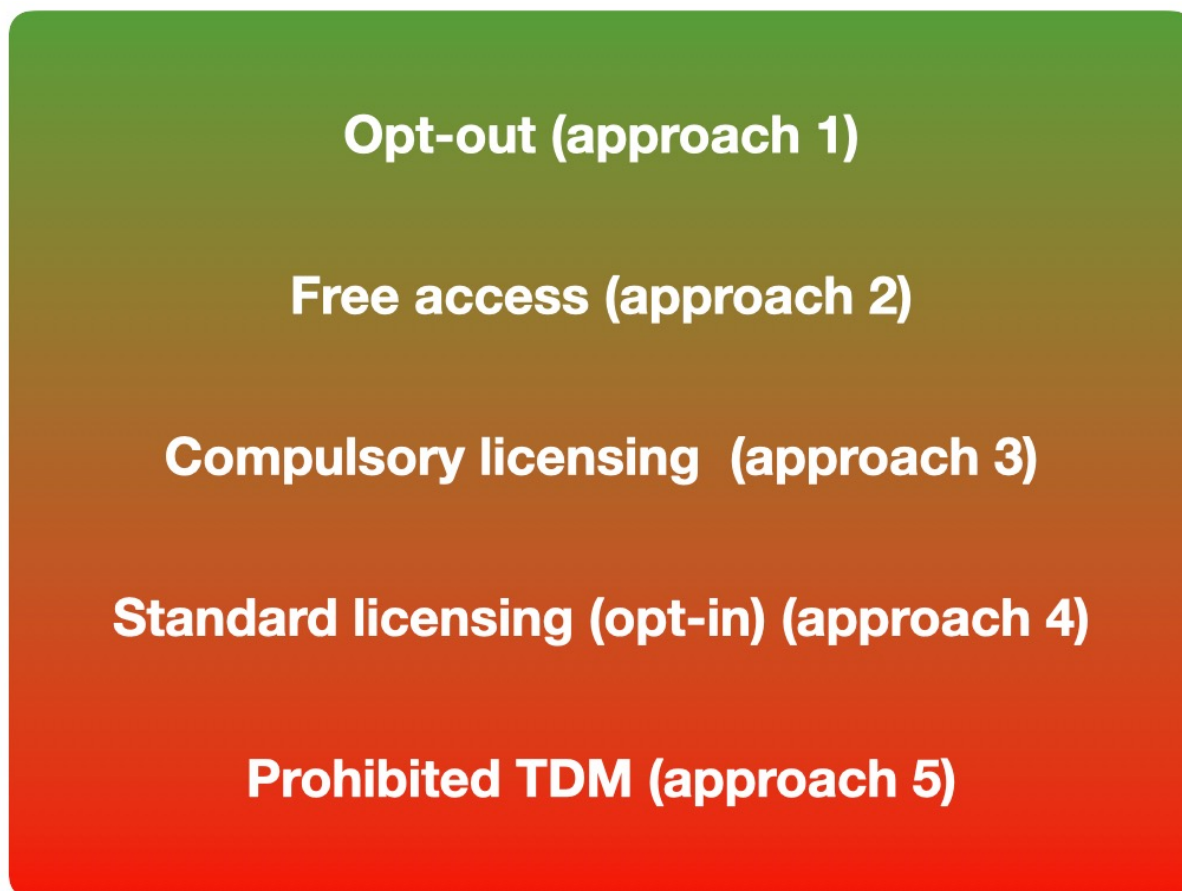
Compulsory licensing (approach 3) ensures that all data remains available, with rights holders neither needing to grant consent nor being able to prevent TDM. This comes, however, at significant cost, since every use carries a fee. These fees would not reach the levels typical of standard licensing (approach 4) given the lower transaction costs made possible by uniform terms and public oversight of the process, which should also prevent the abuse of negotiating power. Given the already substantial technical costs (e.g., hardware and energy), such additional costs could nonetheless impose a burden significant enough to meaningfully delay AI investment and development. In practice, such a system would presuppose some form of collective rights management, with one or more collecting societies (or a dedicated body) setting tariffs, gathering fees, and distributing them among rights holders. The position of this approach on the Y-axis would, of course, vary considerably depending on the fee-setting policy adopted.

Standard licensing (opt-in) (approach 4) requires prior consent from rights holders for every use of protected content, meaning that no work may be mined or otherwise exploited unless authorisation has first been obtained. This would generate enormous transaction costs, severely restrict data availability, and amplify the scale of problems already recognised in intellectual property law — the tragedy of the anticommons, patent trolling and IP thickets — each of which is covered in more detail later in this analysis.

Prohibited TDM (Approach 5) refers to a scenario in which the act of text and data mining is banned outright as such, i.e. it may not be carried out under any conditions, regardless of consent or remuneration. If enacted, a ban on TDM would simply cause AI development activity to relocate to other jurisdictions. Without legal access to training data, such activity would be too legally precarious to sustain.

The above comparison clearly demonstrates that the opt-out mechanism is the only solution that facilitates AI development while simultaneously ensuring creators retain control over their work. Approach 2, while the most beneficial for AI growth due to full data availability at no cost, would mark a significant departure from the regulatory tradition that has shaped European Union copyright policy to date. Conversely, all available alternatives that would further strengthen the position of creators (approaches 3 and 4) increase data acquisition costs to the point that, in practice, the development of the AI sector would be impossible within such jurisdictions. Moreover, prohibited TDM would be the worst available outcome: it would eliminate the AI sector's development potential within the jurisdiction while delivering no compensating benefit to creators.

Figure 2. Comparison of regulatory approaches to TDM



Source: Author's own study

In the context of European Union regulations, if the political consensus remains aligned with the current TDM framework, policy efforts should focus specifically on refining the opt-out mechanism. This is further supported by the fact that the regulatory approaches from other continents, as presented earlier, are not without their own flaws and limitations. Moreover, the issue of legal certainty is paramount; since the European Union has already committed to this regulatory path, it is prudent to build on this established approach.

From the perspective of economic analysis of law, the enactment or modification of legislation must primarily focus on:

- 1) ensuring low transaction costs,
- 2) preventing the suboptimal allocation of resources (e.g. the Tragedy of the Anticommons), and
- 3) providing legal incentives for socially desirable behaviours, such as artistic and scientific endeavours.

The following three subsections examine these issues within the context of the opt-out mechanism for TDM. They provide an analysis of the potential threats and identify methods to neutralise or minimise those threats to ensure that the European Union's opt-out mechanism continues to deliver the most favourable cost-benefit ratio.

While this report focuses on the growth potential of the AI sector, it should be emphasised that the opt-out mechanism serves to optimise overall economic welfare (i.e., net growth). It does so by preventing the negative externalities that could be generated by the unrestricted use of data under a free-access model. By protecting the rights of creators, opt-out safeguards the revenue base of the creative sectors, enabling them to continue investing in content that is, paradoxically, indispensable for training future, even more advanced AI models.

That this trade-off is workable in practice is already evidenced by the rapidly expanding licensing market between AI developers and rights holders, both within the European Union and elsewhere. For example, in 2025 alone, OpenAI concluded agreements with The Guardian, Schibsted, Axios, The Washington Post and Condé Nast; Google signed its first such deal with the Associated Press; Mistral partnered with Agence France-Presse; and Amazon entered multi-year arrangements with The New York Times, Condé Nast and Hearst, allowing the use of their content in AI-related products and services. The March 2026 report by the UK Publishers Association concluded that the market is already maturing at a scale far beyond previous estimates, with multi-million-dollar deals becoming commonplace.⁶ This suggests that the opt-out approach is delivering tangible benefits for both sides without requiring a shift to more interventionist regulation.

Minimising Transaction Costs

Sound regulation, in the economic analysis of law tradition, is characterised by the minimisation of transaction costs.⁷ In the TDM context, the main driver of transaction costs is legal uncertainty over the validity of opt-out reservations. The opt-out mechanism reduces transaction costs most effectively when reservations are expressed in a strictly standardised, machine-readable format, with robots.txt remaining the universal de facto standard, complemented by emerging structured-metadata schemes such as TDMRep, C2PA, and IPTC.⁸

If, by contrast, natural-language reservations or terms-of-service clauses were deemed acceptable, every dataset would require multi-stage legal verification, which would raise costs of due diligence and discourage downstream investment.

The Tragedy of the Anticommons

One of the most significant threats to AI development is the growing risk of a tragedy of the anticommons. This phenomenon occurs when excessive fragmentation of property rights, combined with the ability to block access to resources, prevents those resources from being used. The transaction costs involved in attempting to “consolidate” fragmented resources for large-scale AI development become so prohibitive that projects can become economically unviable.

⁶ Publishers Association, Content Superpower: UK publishing and the AI licensing market (Publishers Association, 2026).

⁷ Robert Cooter and Thomas Ulen, *Law and Economics* (Scott, Foresman & Co., 1988).

⁸ A robots.txt file is a text file located in the root directory of a website that instructs automated web crawlers and search engine bots which pages or folders they can or cannot visit. TDMRep (Text and Data Mining Reservation Protocol) is a machine-readable protocol to signal copyright opt-outs for automated web scraping and AI training, complying with European Union copyright rules. C2PA (Coalition for Content Provenance and Authenticity) is an open technical standard that embeds tamper-evident, cryptographically secure metadata into media files (images, video, audio) to track asset history, creation tools, and authenticity. IPTC (International Press Telecommunications Council) is a standards body providing traditional photo and news metadata frameworks updated with specific fields for AI generation context and data mining restriction flags.

Two pivotal analytical breakthroughs provide the basis for evaluating the consequences of European Union regulation of TDM. The first is Garrett Hardin's *Tragedy of the Commons*, which describes the overexploitation of a shared resource by "free riders."⁹ For decades, this was also the dominant paradigm in intellectual property theory: the right to exclude was seen as a safeguard against exploitation and a necessary incentive for creation.

The second breakthrough is Michael A. Heller's *Tragedy of the Anticommons*.¹⁰ Heller described a phenomenon operating in precisely the opposite direction: when too many owners hold the right to exclude others from using a resource, the result is its chronic underuse — even when exploitation would be socially valuable. Heller illustrated this with a vivid historical example: the iron chains stretched across the Rhine by feudal barons in the Middle Ages, each extracting a separate toll, made navigation impossible, despite the individual charges being "justified" in isolation.

Table 2. Comparison between The Tragedy of the Commons and The Tragedy of the Anticommons

Feature	Tragedy of the commons	Tragedy of the anticommons
Core definition	Individuals acting independently deplete a shared resource through overuse.	Multiple owners hold the right to exclude others, leading to underuse.
Property rights	Non-excludable; no one has the right to prevent others from using the resource.	Fragmented; too many people have the right to say "no" (veto power).
The problem	Too many people use the resource because it is "free."	The resource remains unused because it is too difficult to get permission from everyone.
Incentive	Use it now before someone else does (race to the bottom).	Hold out for a better deal or simply block others (coordination failure).
Outcome	Depletion or ruin: the resource is exhausted (e.g., collapsed fish stocks).	Waste or stagnation: the resource is wasted by not being used (e.g., empty buildings).
Classic example	Overgrazing a public pasture or pollution of the atmosphere.	"Patent thickets" in biotech or fragmented land plots that prevent development.

Source: Author's own study based on Hardin (1968) and Heller (1998)

Also in 1998, Heller and Rebecca S. Eisenberg brought a second foundational contribution, demonstrating through the case of fragmented rights in biomedical research, the precise causal mechanism of the tragedy of the anticommons: the greater the number of veto players standing along the path to innovation, the higher the aggregate transaction costs — and the greater the probability that no single actor will see the research process through to completion.¹¹ Each additional barrier, even if rational in itself, contributes to systemic innovation paralysis. James M. Buchanan and Yong J. Yoon went on to formalise this insight mathematically, showing that the social optimum lies between the polar extremes of full commonality and full privatisation, and that deviation in either direction generates inefficiency.¹²

In digital data sectors, where the marginal cost of reproduction approaches zero, a move towards excessive propertisation is particularly destructive as the resulting social loss is entirely detached from any logic of scarcity.¹³ Although the works cited address other domains of economic activity, their insights can be meaningfully extrapolated to TDM,

9 Garrett Hardin, "The Tragedy of the Commons: The Population Problem Has No Technical Solution; It Requires a Fundamental Extension in Morality," *Science*, 162(3859), 1968, pp. 1243–1248.

10 Michael A. Heller, "The Tragedy of the Anticommons: Property in the Transition from Marx to Markets," *Harvard Law Review*, vol. 111, no. 3, 1998, pp. 621–688.

11 Michael A. Heller and Rebecca S. Eisenberg, "Can Patents Deter Innovation? The Anticommons in Biomedical Research," *Science*, 280(5364), 1998, pp. 698–701.

12 James M. Buchanan and Yong J. Yoon, "Symmetric Tragedies: Commons and Anticommons," *The Journal of Law and Economics*, 43(1), 2000, pp. 1–14.

13 Michael Boldrin and David K. Levine, *Against Intellectual Monopoly* (Cambridge University Press, 2008).

where the negative effect may be even more pronounced. The economics of intangible resources emphasises that when the marginal cost of copying data is zero, artificially erecting proprietary barriers no longer serves to manage a scarce resource — data, unlike a pasture, is not depleted by use — but functions solely as a mechanism for extracting rent, generating pure deadweight loss.¹⁴

‘Data, unlike a pasture, is not depleted by use.’

Unlike the goods described by Hardin, data is a non-rival good: its use by an AI algorithm does not diminish its availability to other actors.¹⁵ From an economic perspective, it is precisely the opt-in model, with its requirement of prior licensing, that generates the most destructive form of artificial scarcity. It transforms the free flow of information into a system of pervasive rent-seeking, in which the mere fact of holding data becomes a barrier that blocks technological development rather than an impulse for the creation of new value.

The patent literature demonstrates how non-practising entities (NPEs) — so-called “patent trolls” — accumulate rights not in order to create, but to set legal traps for innovators.¹⁶ The opt-in model would constitute an ideal breeding ground for “data trolls.” In a system where the absence of explicit consent is tantamount to a legal infringement, trolls can acquire low-value databases *en masse* and subsequently seek enormous damages from AI market leaders. For large players with substantial financial resources, the mere threat of project delays can have catastrophic consequences for market valuation, making them easy targets for extortion.

In this context, economic analysis points unequivocally to the conclusion that the opt-out mechanism is not the problem but the necessary solution, provided it is properly designed. To avoid the risks of data thickets and troll activity, an efficient system must be built on two foundations:

- **Strict technical standardisation:** a requirement that any opt-out reservation be expressed exclusively in a machine-readable and unified format (e.g. metadata, technical protocols, robots.txt instructions).
- **Elimination of uncertainty:** a clear determination that the absence of a technical reservation is equivalent to implied consent for TDM purposes.

Incentive-Based Paradigm

Incentive-based theory justifies intellectual property protection as an instrument for eliminating the public goods problem: because creative works are intangible and non-excludable, without legal protection they would be underproduced, since creators would be unable to capture the benefits of their labour. Copyright, therefore, grants a temporary monopoly, allowing creators to recover the costs of creation and generating a surplus that motivates future creation.

¹⁴ Ian Hargreaves, *Digital Opportunity: A Review of Intellectual Property and Growth* (Intellectual Property Office, 2011).

¹⁵ Bartłomiej Biga, *The Economics of Intellectual Property and Openness: The Tragedy of Intangible Abundance* (Routledge, 2021).

¹⁶ Lauren Cohen, Umit G. Gurun and Scott Duke Kominers, “Patent Trolls: Evidence from Targeted Firms,” *Management Science*, 65(12), 2019, pp. 5461–5486.

The incentive-based paradigm has, however, faced sustained criticism in legal and economic scholarship on several fundamental grounds. First, empirical research shows that creators are driven by a mixture of motivations: reputation, artistic passion, the need for expression, academic prestige and commercial commissions.¹⁷ Property rights are one element among many in this motivational structure and are often not the dominant one. Many musicians, bloggers, academics and other creators produce valuable content even in the absence of, or with only limited, legal protection.

Second, copyright protection is, by definition, prospective in its operation: it can serve as an incentive for future creation, but it cannot rationally justify compensation for works already produced. Those works came into existence without any knowledge of their subsequent use by AI; retroactive remuneration did not, and could not, have influenced the decision to create them. The incentive-based argument therefore loses its force as a justification for retrospective claims against AI companies concerning historical content.

That said, AI models trained on cultural and informational content generate enormous commercial potential. It must be stressed, however, that the value supplied by content to a general-purpose AI model is collective and emergent, not per-work. Contemporary

‘The value supplied by content to a general-purpose AI model is collective and emergent.’

foundation models are trained on corpora measured in trillions of tokens, and the marginal contribution of any single work to a model’s capabilities is, in the overwhelming majority of cases, negligible. Value emerges instead from the cumulative contribution of an entire ecosystem of creators, publishers and information producers.

The case for routing some share of AI-generated value back to the creative sector remains strong. Introducing TDM-related fees too early, however — through opt-in or other more restrictive approaches — would be highly risky, given that the current phase of AI development involves enormous costs (e.g., energy, infrastructure) that are only marginally offset by revenues from AI product sales. This phenomenon is, not without reason, known as “AI’s \$600 Billion Question.” As analysts at Sequoia Capital have observed, for investments in chips — primarily from Nvidia — to break even, the AI industry would need to generate approximately \$600 billion in annual revenue.¹⁸ Current revenues from AI software (in the form of Software as a Service (SaaS)) represent a fraction of that figure: OpenAI alone recently surpassed \$3.4 billion in annual revenue — a sum that, measured against the scale of the entire industry’s expenditure, amounts to a drop in the ocean.

Imposing TDM fees now, when the industry is running a deficit of hundreds of billions of dollars, could drive smaller players into insolvency and accelerate the monopolisation of the market by the wealthiest actors.

¹⁷ William M. Landes and Richard A. Posner, *The Economic Structure of Intellectual Property Law* (Harvard University Press, 2009).

¹⁸ David Cahn, “AI’s \$600B Question,” *Sequoia Capital*, 20 June 2024.

Empirical Evidence: Licensing Markets

The theoretical case for opt-out, set out above, is increasingly supported by empirical observation. Two trends are particularly relevant. First, a market for AI training rights has emerged within the European Union's opt-out regime, supporting the proposition that legal certainty and creator agency can coexist with viable AI development. Second, deployment cases across the news, audiovisual and language-services sectors suggest that AI tools, when integrated into existing creative and editorial workflows, tend to drive efficiency, audience expansion and reinvestment into core creative work, rather than functioning as substitutes for human labour. Together, these observations indicate that the opt-out framework is capable of supporting its dual objectives. However, as discussed at the end of this section, whether it does so optimally and without recourse to mandatory licensing remains a normative question that the empirical evidence alone cannot resolve.

The licensing agreement between Axel Springer — publisher of Bild, Die Welt, Politico, and Business Insider — and OpenAI, concluded in late 2023, provides an early example of the kind of voluntary commercial arrangement that the European Union framework is designed to accommodate. This agreement was one of the first of its kind and was followed by many others. Under the agreement, OpenAI obtains rights to use Axel Springer content for training and to provide up-to-date journalism in ChatGPT responses. In return, the publisher receives compensation that flows directly into newsroom investment. The full financial terms have not been officially disclosed; however, press estimates, including reporting by Bloomberg, place the value in the tens of millions of euros over the multi-year term.¹⁹

In the music sector, the agreement between Universal Music Group and AI music platform Udio provides another example. Announced in October 2025 as a settlement of earlier copyright litigation, the agreement establishes a framework under which Udio operates as a licensed AI music creation platform drawing on Universal's catalogue, with revenue sharing structured to flow to rights holders and, through them, to artists.²⁰ The settlement is significant because it demonstrates that even in the segments of the creative industries where concerns about stylistic substitution are most acute, durable commercial settlements can be reached.

These two cases are illustrative rather than exhaustive. Publicly disclosed deals in the news, magazine and information sectors suggest that the market for voluntary AI licensing has expanded significantly since 2023, with public disclosures increasing sharply during 2024 and 2025.²¹ Any precise count depends on the criteria used — what qualifies as a distinct agreement, how grants and revenue-share arrangements are treated, and how multi-publisher deals are aggregated — and the figures cited in industry trackers should accordingly be read as indicative rather than definitive.

¹⁹ Bloomberg, "OpenAI to Pay Axel Springer Tens of Millions to Use News Content," 13 December 2023.

²⁰ Universal Music Group, "Universal Music Group and Udio Announce Udio's First Strategic Agreements for New Licensed AI Music Creation Platform," press release, PR Newswire, 29 October 2025.

²¹ Tow Center for Digital Journalism, *AI deals and lawsuits*, n.d.

Where financial terms have been disclosed, several high-profile cases involve reported amounts in the tens of millions: Meta’s 2026 arrangement with News Corp is reported at up to \$50 million (approximately €46 million) annually over at least three years;²² Perplexity’s Comet Plus revenue-sharing pool is reported at \$42.5 million (approximately €39 million);²³ and the Axel Springer–OpenAI deal cited above sits in a comparable range.²⁴ Beyond bilateral contracts, intermediated structures have also emerged: Factiva/Dow Jones reports licensing rights from more than 8,000 source publications for generative AI use;²⁵ Microsoft’s Publisher Content Marketplace operates as a multilateral exchange in which publishers define usage terms and are remunerated on a delivered-value basis;²⁶ and the Bria–News/Media Alliance (NMA) arrangement of March 2026 extends a revenue-share template to NMA member publishers under a single contract.²⁷ These developments are consistent with a market in which transaction costs — central to the theoretical objection to *ex ante* licensing — have been at least partially absorbed through standardised contractual templates, supply consolidation and revenue-share architectures that reduce the need for per-work valuation.

Taken together, the case studies and the broader publicly disclosed record indicate that the opt-out mechanism can function as an effective infrastructure for negotiation: it establishes a default rule that favours AI development while preserving rights holders’ bargaining

‘Voluntary licensing markets have emerged within the opt-out framework rather than despite it.’

position where they choose to monetise their works. This emerging deal flow constitutes substantial evidence that the European Union framework is capable of supporting both objectives. Whether it is the optimal or sufficient model for the entire European Union, however, is a separate, normative question — one that the empirical record can

inform but cannot definitively resolve. What the empirical record does resolve, however, is the direction: voluntary licensing markets have emerged, are growing, and are doing so within the opt-out framework rather than despite it.

The Opt-Out Rule: Estimating the Results for the European Union’s Economy

Assumptions

Martin Thelle et al. estimate that, if the European Union widely adopts generative AI, a potential increase in GDP of €1.2 to €1.4 trillion over ten years is achievable.²⁸ This represents the increase above baseline GDP following a ten-year adoption period. As the estimate reflects an optimistic scenario in which the technology develops without significant impediments, the current policy brief therefore uses the conservative end of this range, i.e., €1.2 trillion, as the maximum potential additional GDP by 2035. Importantly, this estimate was calculated at a time when the opt-out mechanism already constituted the regulatory baseline for TDM in the European Union. What Thelle et al. could not have anticipated,

22 Karissa Bell, “Meta Signs a Multimillion Dollar AI Licensing Deal with News Corp,” *Engadget*, 03 March 2026.

23 Ian Carlos Campbell, “Perplexity Has Cooked up a New Way to Pay Publishers for Their Content,” *Engadget*, 25 August 2025.

24 Bloomberg, “OpenAI to Pay Axel Springer Tens of Millions to Use News Content,” 13 December 2023.

25 Dow Jones, “Dow Jones Factiva Surpasses 8,000 Licensed Sources for GenAI Use,” press release, *Business Wire*, 20 January 2026.

26 Tim Frank, “Building Toward a Sustainable Content Economy for the Agentic Web,” *Microsoft Advertising*, 03 February 2026.

27 Sara Guaglione, “News/Media Alliance Signs AI Licensing Deal to Unlock Recurring RAG Revenue for Small and Mid-Sized Publishers,” *Digiday*, 23 March 2026.

28 Martin H. Thelle, Anders Thor Lundberg, Bodile Emilie Hovmand, Hans Henrik Woltmann, Laura Virtanen, Nikolaj Tranholm-Mikkelsen et al., *The Economic Opportunity of AI in the EU. Beyond the Hype – Boosting EU Competitiveness with Generative AI* (Implement Consulting Group, 2024).

however, was the pace and scale at which rights holders would exercise this option. More recent observations confirm that the volume of data excluded from TDM is growing rapidly; if that trend continues, it could significantly undermine the assumptions underlying the projection, meaning that the economic gains from artificial intelligence could ultimately prove substantially smaller than Thelle et al. assume. The estimated gains could be further constrained by a range of regulatory obstacles, such as a potential shift towards an opt-in mechanism — which would render the European Union's TDM framework exceptionally restrictive by global standards.

The opt-out mechanism will likewise fall short of delivering the maximum possible AI-driven gains for the European Union's GDP. Unlike more restrictive regulations, however, it will not entirely block AI development in Europe and will offer the European Union a realistic chance of capturing a substantial share of the expected benefits. To calculate the extent to which opt-out could reduce those gains relative to the optimistic scenario, it is necessary to estimate what proportion of content valuable for AI training will be excluded because of this mechanism. Research by the Data Provenance Initiative shows that within just one year (2023–2024), 25–28% of high-quality sources in major training corpora were already subject to exclusions.²⁹ Looking at the scale of AI crawler blocking, it is reasonable to assume that this phenomenon will grow rapidly, particularly in the European Union, where there is a robust legal basis for such restrictions.³⁰ For the purposes of the analysis that follows, three scenarios are therefore adopted, in which 30%, 50% and 70% of valuable content is excluded respectively.

The relationship between data volume and model quality is not merely theoretical. A concrete illustration is provided by the GPT-NL project in the Netherlands. This publicly funded effort to build a sovereign Dutch-language model adopted a deliberate opt-in data sourcing policy, training exclusively on licensed content and abstaining from the open web, including Wikipedia. This reduced its Dutch-language training corpus to under 55 billion tokens, approximately 10% of what would have been lawfully available under the TDM exception. Benchmark results published in late 2025 and analysed by Edwin Rijgersberg show the model scoring near random-guessing level on factual knowledge in Dutch, and underperforming on all four Dutch-language indicators against models of comparable or smaller size that trained on unrestricted corpora.³¹ These models included the Swiss Apertus model, which complied with opt-outs but otherwise used data available under the applicable laws.³² GPT-NL's experience illustrates that the delta (δ) parameter in this report — used to express the percentage of data lost — is not a hypothetical: voluntary data restriction of the kind a mandatory licensing or opt-in regime would produce at scale generates measurable, substantial degradation in model capability.

29 Shayne Longpre, Robert Mahari, Ariel Lee, Campbell Lund, Hamidh Oderinwale, William Brannon et al., "Consent in Crisis: The Rapid Decline of the AI Data Commons," *arXiv*, 24 July 2024.

30 Richard Fletcher, "How Many News Websites Block AI Crawlers?" *Reuters Institute for the Study of Journalism*, 22 February 2024.

31 Edwin Rijgersberg, "GPT-NL as a Dutch Sovereign Alternative? The Preliminary Benchmarks Don't Look Good," *GoingDutch.ai*, 18 June 2026.

32 Paul Keller, "GPT-NL Respects Copyright — Cui Bono?" *Open Future*, 26 June 2026.

The current analysis adopts $\delta = 0.5$ (i.e., exclusion of 50% of valuable data) as the central scenario, which is grounded in the trend observed in 2023–2024 and described by the above-mentioned report by The Data Provenance Initiative. Given the growing availability of automated opt-out tools and the progressive professionalisation of copyright management across the digital ecosystem, extrapolating this trend to a 50% level by 2035 is a conservative assumption. This scenario thus represents the centre of gravity between the current situation ($\delta = 0.3$) and a high-market-restrictiveness model ($\delta = 0.7$).

The next step requires estimating the degree to which restricted access to valuable data degrades the AI models trained on it. This analysis draws on neural scaling laws, originally formulated by Kaplan et al.³³ and subsequently revised in the context of computational optimisation by Hoffmann et al.³⁴ Building on work by Sorscher et al.³⁵ and Covert et al.³⁶ on the marginal contribution of individual data points, the current analysis adopts an elasticity coefficient (ϵ) range of 0.2–0.5. This range accounts for the fact that data is not homogeneous: the removal of high-information-density datasets — so-called “gold data” — produces a far steeper decline in the predictive and cognitive capabilities of a model than the removal of redundant data.

Finally, the GDP impact estimates presented in the current analysis should be interpreted as upper-bound estimates of the potential economic cost of TDM restrictions. These estimates rest on the assumption that access to high-quality training data is a binding constraint on AI development during the period under analysis. In practice, the degree to which data access translates into real economic growth depends on a range of structural factors, including the availability of computing power (GPUs), skills gaps in the labour market, energy infrastructure and the private sector’s capacity to commercialise innovation. In circumstances where these constraints are significant, the actual negative impact of TDM restrictions may be lower in absolute terms; nevertheless, their role as a systemic barrier remains critical for long-term digital competitiveness.

Calculations

This model is intended as a transparent scenario tool, not a precise forecast. Its aim is to clarify the direction and rough scale of the trade-offs at stake — and to make its assumptions open to challenge — rather than to pin down an exact figure.

The elasticity coefficient ϵ is *inspired* by neural scaling laws rather than derived from them: those laws describe how a model improves as data is *added*, whereas in the current paper a similar mathematical form is applied to the *removal* of high-value data and the result used as a proxy for model quality. The range adopted ($\epsilon = 0.2$ – 0.5) is therefore intended to represent a set of plausible scenarios, not to reproduce a specific published value. The research on the value of individual data points is cited only to support one qualitative point — that not all data is equal, so losing high-quality data has a greater effect than losing redundant data — and not to fix the value of ϵ .

33 Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B. Brown, Benjamin Chess, Rewon Child et al., “Scaling Laws for Neural Language Models,” *arXiv*, 20 January 2020.

34 Hoffmann, Jordan, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford et al., “Training Compute-Optimal Large Language Models,” *arXiv*, 29 March 2022.

35 Ben Sorscher, Robert Geirhos, Shashank Shekhar, Surya Ganguli and Ari S. Morcos, “Beyond Neural Scaling Laws: Beating Power Law Scaling via Data Pruning,” *arXiv*, 29 June 2022.

36 Ian Covert, Wenlong Ji, Tatsunori Hashimoto and James Zou, “Scaling Laws for the Value of Individual Data Points in Machine Learning,” *arXiv*, 30 May 2024.

The transmission coefficient ($\tau = 1$) is, likewise, a simplifying baseline, not an empirical fact; it assumes that a given drop in model quality translates one-to-one into a proportional drop in AI-driven GDP. It is used as a neutral baseline precisely because the true relationship is almost certainly non-linear and sector-specific, and estimating a different value would require additional contestable assumptions. The resulting GDP figures should therefore be read alongside the full range of scenarios in the tables as illustrative magnitudes conditional on these assumptions — not as exact valuations. Their value lies less in any single euro figure than in a consistent pattern: under every plausible combination of parameters, opt-out preserves a substantially larger share of the potential AI-driven GDP gain than the licensing-based alternatives.

The impact of data loss on model performance can be expressed using a simplified multiplicative model:

$$Q_{\text{new}} = Q_{\text{old}} \cdot (1 - \delta)^\epsilon$$

Where:

Q_{new} (quality new): The projected performance or quality of the model following partial data loss. In economic analysis, this can be interpreted as “AI-assisted productivity.”

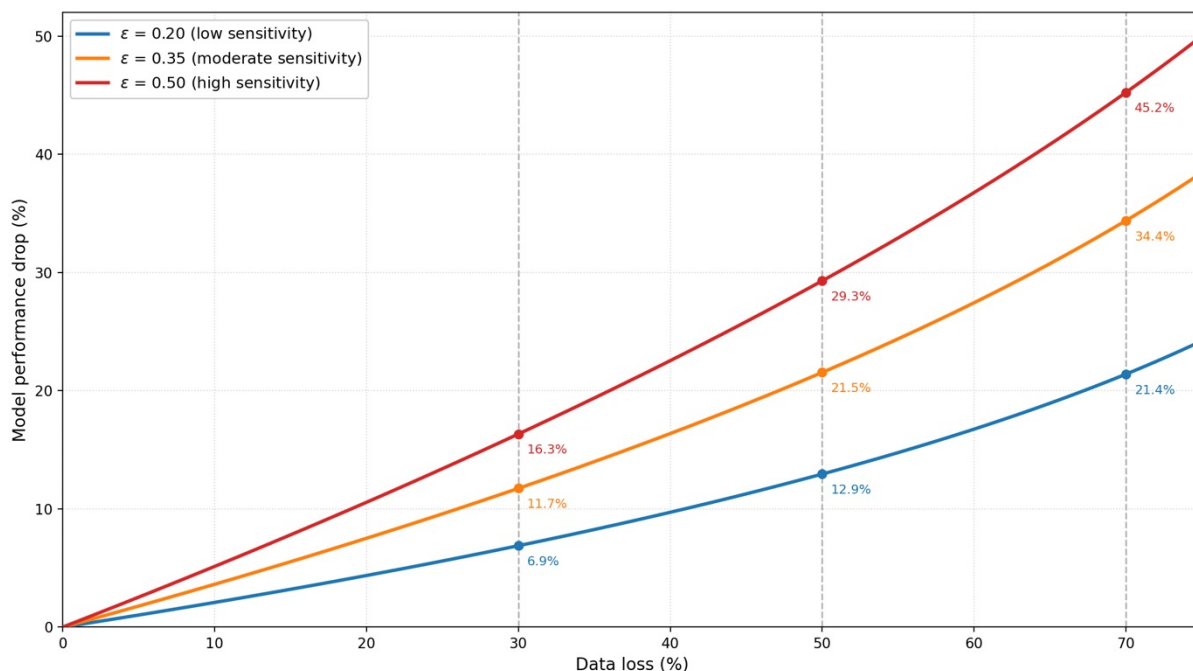
Q_{old} (quality old): The baseline model performance (reference point). This is typically set to 1 (or 100%), so that the result is expressed as a fraction of baseline performance.

δ (delta): The percentage of data lost, expressed as a decimal fraction.

$(1 - \delta)$: The remaining share of the data resource. This is the base that is raised to the scaling exponent.

ϵ (epsilon): The performance elasticity coefficient with respect to data — the key parameter governing the model’s “sensitivity” to information loss. It is derived directly from research on neural scaling laws, where ϵ corresponds to the power-law exponent.

Figure 3. Impact of data loss on AI performance



Source: Author's own study based on Covert et al. (2024) and Sorscher et al. (2022)

The figure above visualises these relationships, showing clearly that as elasticity increases, the quality degradation curve becomes progressively steeper. It is worth noting that at a 50% data loss under the high-sensitivity scenario, the model loses nearly one third of its original performance.

'AI systems are vulnerable to the removal of large portions of their training datasets.'

These figures demonstrate that AI systems are relatively resilient to minor data shortfalls, but highly vulnerable to the removal of large portions of their training datasets.

Results

In analysing the impact of data resource constraints on the European Union economy, the conservative assumption of unit transmission elasticity ($\tau = 1$) has been adopted, meaning that a relative degradation in AI model quality translates into a proportional decline in projected GDP benefits for the European Union. This hypothesis treats artificial intelligence as digital intellectual capital that directly determines the marginal productivity, particularly in professional services and the knowledge economy, which form the backbone of the European Union's internal market. With this framing, every degradation in a model's cognitive capabilities requires workers to spend more time verifying and correcting AI-generated outputs, which at the macroeconomic level effectively erodes the gains from automation in proportion to the decline in model quality.

It should be noted, however, that the real-world relationship between technological precision and economic growth may be non-linear and vary across sectors. In sectors with stringent safety requirements and low error tolerance — such as medical diagnostics, autonomous transport or precision engineering — data quality losses may produce so-called threshold effects, where a drop in precision below an acceptable level results in the complete abandonment of the technology, dramatically amplifying losses ($\tau > 1$). Conversely, in tasks

with low operational complexity, workers' ability to adapt may compensate for weaker AI performance, limiting the economic impact ($\tau < 1$). Despite these sources of variation, adopting a 1:1 transmission elasticity coefficient provides the most stable and objective reference point for modelling the regulatory consequences, avoiding speculative overestimation of the European technology sector's resilience to constraints on access to information resources.

It is worth emphasising, however, that the assumption of unit elasticity should be regarded as a deliberately cautious lower bound rather than a best estimate. While the precise shape of the transmission function undoubtedly requires further in-depth empirical research, several considerations suggest that the actual relationship between model degradation and forgone GDP benefits may well be steeper than proportional — and potentially closer to exponential. Threshold effects of the kind described above are unlikely to remain confined to a handful of safety-critical sectors: as AI systems become embedded in interdependent value chains, the abandonment of the technology in one link may cascade into others, compounding the aggregate loss. Moreover, many of the highest-value applications — precisely those responsible for the bulk of the projected GDP gains — are also those with the lowest tolerance for error, meaning that quality degradation disproportionately erodes the most economically significant use cases first. If this reading is correct, the estimates presented in this analysis, far from overstating the costs of restricting access to training data, may in fact substantially understate them.

By subtracting the opt-out costs calculated as described above from the unconstrained AI development scenario, it is possible to derive the value of the additional potential GDP preserved by opt-out.

Table 3. Additional potential European Union GDP in 2035 under different opt-out scenarios, €

	Delta (δ) = 0.3	Delta (δ) = 0.5	Delta (δ) = 0.7
Epsilon (ϵ) = 0.20	1,117,200,000,000	1,045,200,000,000	943,200,000,000
Epsilon (ϵ) = 0.35	1,059,600 000 000	942,000,000,000	787,200,000,000
Epsilon (ϵ) = 0.50	1,004,400,000,000	848,400,000,000	657,600,000,000

Source: Author's own study

In the central scenario, the opt-out mechanism preserves €942 billion in additional potential GDP compared to a scenario in which more restrictive copyright and AI regulation would have blocked its development in the European Union. Expressed from the other direction, opt-out forgoes approximately €258 billion in additional potential GDP relative to a scenario of fully unrestricted data access, quantifying the cost of retaining creators' reservation rights within this framework.

It should be noted, however, that this comparison is made against the most restrictive regulatory scenario. In reality, the extent of this loss depends on the details — specifically, the degree of restrictiveness — of regulations based, for example, on a licensing obligation. It is nevertheless impossible to determine where the tipping point lies: that is, at what level of costs (arising from licence fees and associated transaction costs) the development of AI in the European Union would grind to a halt.

Under these assumptions, opt-out across different scenarios preserves the following share of the €1.2 trillion maximum potential additional GDP:

Table 4. Share of additional European Union GDP in 2035 preserved through opt-out

	Delta (δ) = 0.3	Delta (δ) = 0.5	Delta (δ) = 0.7
Epsilon (ϵ) = 0.20	93%	87%	79%
Epsilon (ϵ) = 0.35	88%	79%	66%
Epsilon (ϵ) = 0.50	84%	71%	55%

Source: Author's own study

Comparative Analysis of the Impact of AI Regulation in Germany, Italy and Poland

Assumptions

This section deepens the analysis of differences in how Germany, Italy and Poland have each implemented the CDSM directive's TDM provisions. The aggregate cost estimates for restricting lawful AI training in Europe, developed above, are now disaggregated to the country level.

The country-level economic analysis is based on the following assumptions:

- The reference point is each country's share of European Union GDP in 2025.
- Data loss (δ) is expected to vary across countries. As in the European Union-wide calculations, three scenarios are considered: 30%, 50% and 70%.
- Structural differences between the three economies also justify varying the transmission coefficient (τ). For the European Union average, $\tau = 1$, meaning that the relative degradation in AI model quality translates proportionally into a decline in projected GDP benefits.
- As with the European Union-wide model, three scenarios are presented, differentiated by the key parameter governing the model's sensitivity to information loss (ϵ). The middle value of 0.35 is treated as the baseline.

Table 5. Assumptions for country-level calculations

Country	Delta (δ)	Tau (τ)	% EU GDP
Germany	0.7	1.2	23.77%
Italy	0.5	1.0	12.01%
Poland	0.3	0.8	4.88%

Source: Author's own study

Germany

Regulations

Of the three European Union member states examined in this report, Germany was the first to implement the CDSM directive's TDM provisions on schedule and the first to see them tested before the courts. Germany's implementation of the directive is anchored in §44b (commercial TDM) and §60d (TDM for scientific research) of the country's copyright act, *Urheberrechtsgesetz* (UrhG). The requirement of "machine readability" (*maschinenlesbar*), however, has proved far from self-explanatory in practice. A pivotal moment came with the appellate ruling of the Hamburg Higher Regional Court (OLG Hamburg) in *Kneschke v. LAION* (10 December 2025), which held that a reservation expressed in natural language in a website's terms of service did not meet the machine-readability requirement.

AI and the German Economy

In the latest edition of the *Global AI Index*, Germany ranks seventh in the world, placing it among the "second-tier leaders" — just behind the United States and China and a handful of other high-performing advanced economies.³⁷ Structurally, this reflects a country with solid scientific and institutional foundations, enabling it to hold a strong position in the global technology race. Yet a closer look at the underlying indicators reveals meaningful constraints within Germany's AI ecosystem.

Despite its highly skilled and well-educated workforce, Germany underperforms in areas such as the operational environment, deployment infrastructure and technology commercialisation. In practice, this translates into persistent difficulties converting research potential into scalable market solutions and a comparatively lower dynamism in the private sector relative to the world's leading AI nations.

The result is a "strong science, moderate deployment" model — characteristic of several European economies — which, while delivering stability and research quality, may ultimately constrain Germany's ability to keep pace with more aggressively investing ecosystems such as the United States or China.

'A "strong science, moderate deployment" model, characteristic of several European economies.'

Germany today presents a picture of a two-speed economy in its adoption of AI. World Bank data reveals a deep structural fracture: overall AI adoption among firms stands at just 11–12%, yet in the large-enterprise segment (companies with more than 250 employees) the figure surges past 35%.³⁸ Industrial giants are integrating algorithms efficiently into their production lines, while Germany's economically vital *Mittelstand* — the broad base of small and medium-sized enterprises — is falling behind, at growing risk of further competitive erosion. The German industrial model finds itself caught in a trap: without a change in trajectory, high corporate adoption may prove little more than a cosmetic fix masking a deepening dependence on external technology providers.

³⁷ Joe White and Serena Casareo, *The Global AI Index*, Tortoise Media, 19 September 2024.

³⁸ World Bank, *Digital Progress and Trends Report 2025: Strengthening AI Foundations* (World Bank, 2025).

Cologne-based **DeepL** has built one of the most commercially successful European language-AI businesses by focusing relentlessly on translation quality. Its neural translation engine, trained on extensive multilingual corpora, consistently outperforms larger generalist models on benchmarks for European-language pairs and has been adopted by tens of thousands of enterprise clients across regulated sectors, including law, finance and public administration. DeepL's success demonstrates a point easily overlooked in debates dominated by frontier generative models: proportionate access to training data also underpins the European technology mid-cap segment that delivers concrete productivity gains across the European Union economy — not only headline-grabbing chatbots, but the maybe less glamorous, high-value tools embedded in everyday professional workflows.

Germany ranks eleventh in the *Global Innovation Index* and remains a European powerhouse in research and development (R&D) expenditure.³⁹ Yet the German economy is currently navigating a painful phase of digital recalibration. Traditional industrial conglomerates such as Volkswagen and Mercedes-Benz are being forced to trim budgets in the face of declining revenues — a stark contrast to the aggressive expansion of global information and communication technologies (ICT) hardware giants. Shifting R&D budgets away from traditional manufacturing and toward AI-intensive sectors is therefore not merely a cost-optimisation exercise. For the German economy, it is an urgent act of self-preservation.

Originally a spin-off from the Technical University of Munich (TUM), **Celonis** is now Germany's most highly valued technology start-up. The company created and came to dominate the category of process mining, establishing itself as a strategic pillar of Germany's digital transformation. Its process intelligence platform acts as an "X-ray" for major corporations — including Siemens, BMW, and Deutsche Telekom — enabling the visualisation and real-time optimisation of billions of business operations.

In an era of intense pressure on efficiency, Celonis equips German industry with tools to achieve dramatic reductions in both costs and CO₂ emissions, demonstrating that Germany is capable of building global deep-tech champions that do not merely compete with Silicon Valley, but set their own data engineering standards for it to follow. Celonis is pushing the boundaries of innovation further still, deploying so-called "agentic AI" — systems that do not simply analyse data, but autonomously propose and execute corrective actions across supply chains and financial operations. By creating "digital twins" of entire organisations, the company enables the safe deployment of artificial intelligence grounded in hard process data rather than in probabilistic language models alone.

³⁹ World Intellectual Property Organization (WIPO), *Global Innovation Index 2025: Innovation at a Crossroads* (World Intellectual Property Organization, 2025).

The Economic Consequences of TDM Restrictions

Of the countries analysed in this report, the highest expected data loss (δ) is in Germany, at a level of 70%. This follows primarily from the fact that opt-out mechanisms were introduced at an early stage in Germany, and both the publishing sector and the collective rights management organisations occupy an exceptionally strong position. The trajectory of case law, including the ongoing appeal in *Kneschke v. LAION*, will be a key variable: stricter requirements for effective opt-out reservations could further concentrate data exclusions among well-resourced rights holders, reinforcing the high- δ scenario for Germany.

For the German economy, the safest assumption is a tau (τ) above unity, at 1.2. Germany is an economy built on industry 4.0, precision engineering, and the automotive sector. In these industries, the margin for error is minimal. Should AI model quality fall below a certain threshold, autonomous systems and algorithms optimising production lines become operationally unusable on safety grounds — a phenomenon known as threshold effects.

Table 6. Economic consequences for Germany, €

	Delta (δ) = 0.3	Delta (δ) = 0.5	Delta (δ) = 0.7
Epsilon (ϵ) = 0.20	261 622 128 000	241 084 848 000	211 990 368 000
Epsilon (ϵ) = 0.35	245 192 304 000	211 648 080 000	167 492 928 000
Epsilon (ϵ) = 0.50	229 447 056 000	184 949 616 000	130 525 824 000

Source: Author's own study

For Germany, then, the most probable scenario suggests that opt-out will enable AI development to generate an additional €167.5 billion in GDP by 2035. This represents 59% of the maximum potential additional GDP that would be generated under a scenario of unconstrained AI development in the European Union. The lower share of retained AI development benefits in Germany compared with the European Union as a whole stems from the specifics of the German economy and the strong position of the intellectual property rights holder community; in the baseline scenario, this translates into more severe restrictions on data availability for TDM purposes. It is worth keeping in mind that regulatory approaches that are more burdensome than opt-out — such as standard licensing requirements (opt-in) — could deprive the German economy of any benefits from the development of this technology altogether.

For Germany, the opt-out mechanism enables the preservation of the market value of the “*Mittelstand* of knowledge” — the thousands of medium-sized academic and specialist publishers that underpin the strength of the German innovation system. Without this mechanism, the short-term gains of the AI sector could be offset by a long-term decline in the quality and availability of specialist German-language literature.

Italy

Regulations

Italy has taken a pioneering path, going from a standard implementation of the CDSM directive in 2021 to the adoption of a comprehensive artificial intelligence act (the Italian AI law), which entered into force on 10 October 2025. The Italian model is distinctive in its ambition to directly synchronise national copyright law with the transparency requirements of the European Union's AI act. The centrepiece of this framework is article 70-*septies* of the *Legge sul diritto d'autore* (LdA), which unambiguously confirms the applicability of TDM exceptions to generative AI systems — but does so subject to rigorous compliance with opt-outs.

The most radical element of the Italian regime is the introduction of criminal liability (article 171 *a-ter* of the LdA) for violations of TDM mechanisms. This criminalisation sends a powerful political signal, designed to strengthen the negotiating position of rights holders and creators. The Italian system is also characterised by a distinctive form of legal anthropocentrism: the amendment to article 1 of the LdA limits copyright protection — and therefore, the right to opt out — exclusively to works of human creativity. At the same time, Italy has no statutory definition of “machine readability,” meaning that entities conducting TDM operate under considerably greater legal uncertainty than in jurisdictions that favour technical formalism.

AI and the Italian Economy

In the latest edition of the *Global AI Index*, Italy ranks a clear distance from the leading pack — 24th overall — placing it firmly among countries with moderate systemic AI capability.⁴⁰ Its position reflects the comparatively limited scale of the Italian AI ecosystem relative to the front-runners, both in terms of research output and real-world deployment. Yet a closer reading of the index's components reveals an intriguing asymmetry; Italy performs exceptionally well in the so-called operational environment, ranking first globally in a dimension that encompasses regulation, levels of public trust, and openness to new technologies. The principal barrier to growth, however, remains the country's weak capacity to commercialise and scale innovation. Italy scores lower across areas such as technology development, start-up activity and private investment — factors that collectively drag down its overall ranking. This suggests that despite growing awareness of AI's strategic importance and the existence of public AI strategies, Italy has yet to fully mobilise its scientific and institutional potential. The result is a model best described as “enabling conditions, constrained deployment dynamics.” This is characteristic of economies that have built sound

regulatory frameworks but have not yet cultivated the strong private innovation ecosystem needed to compete at a global level.

‘Proportionate TDM rules are not merely a technical matter, but a defence of national innovation identity.’

Italy's exceptional position — ranking first in the world for operating environment in the *Global AI Index* — is

no accident, but the result of a deliberate policy strategy grounded in the concept of a “right to innovation.” The Italian legislature seeks to establish a framework in which technological progress is treated as a cornerstone of economic sovereignty. Yet this prestigious standing is directly threatened by any potential tightening of the interpretation of TDM mechanisms.

⁴⁰ Joe White and Serena Casareo, *The Global AI Index*.

A shift from the flexible opt-out model towards restrictive licensing systems would undermine the very foundations on which Italy's regulatory advantage rests. For Rome, the fight to preserve proportionate TDM rules is not merely a technical matter, but a defence of national innovation identity; if the barriers to entry for AI are raised, Italy risks losing its status as the most innovation-friendly jurisdiction for technology developers.

Project Minerva, developed by a research group at the Sapienza University of Rome in collaboration with the CINECA consortium, represents the first family of large language models trained entirely from scratch with the Italian language and culture in mind. Minerva forms the bedrock of Italy's AI strategy, offering a model that is not merely a localised adaptation of U.S. solutions, but a system deeply embedded in the Italian legal, literary and social context. By harnessing the computational power of the Leonardo supercomputer — one of the most powerful systems in the world — Italian researchers have demonstrated that Europe can build its own transparent, open-source models: systems that guarantee the data security and domain precision that generic, Anglocentric models from global technology giants simply cannot match. Today, Minerva serves as an intelligent support layer for Italian public administration and the broader “Made in Italy” economic ecosystem.

Italy grapples with an even sharper contrast than Germany between its technology leaders and the rest of the economy.⁴¹ Overall AI adoption stands at roughly 6–7%, placing Italy below the Organisation for Economic Co-operation and Development (OECD) average, while large firms maintain a comparatively solid adoption rate of around 25%. This stratification is particularly dangerous for Italy's economic structure, which rests on a foundation of small and medium-sized enterprises (SMEs) — the very businesses that define the distinctiveness of Italian craftsmanship and manufacturing. For the SME sector, such a low level of digitalisation represents a direct threat to productivity and the ability to compete on price in global markets. Without the digital transformation of traditional industries, the backbone of the Italian economy may prove unable to withstand the pressure from competitors who are automating their design and logistics processes faster and at lower cost.

Founded by Italian entrepreneur Matteo Lai in collaboration with Professor Rosalind Picard of the Massachusetts Institute of Technology (MIT), **Empatica** is a world leader in affective computing and medical artificial intelligence. The company stands as a benchmark for the integration of AI with wearable technology, offering — among other products — the Embrace smart wristband, the first device of its kind in the world to receive certification from the U.S. Food and Drug Administration (FDA) for the detection of epileptic seizures. Empatica's systems do more than collect physiological data: through advanced machine learning algorithms, they can predict health crises and monitor vital signs in a manner that is nearly imperceptible to the patient. The company's success demonstrates that European technological ingenuity is capable of dominating even the most demanding and rigorously regulated markets in medical technology — medtech included.

⁴¹ World Bank, *Digital Progress and Trends Report 2025: Strengthening AI Foundations*.

Ranked 28th in the *Global Innovation Index*,⁴² Italy is a textbook example of an innovation-efficient economy: it sits above the trend line, signalling a high capacity to convert inputs into measurable outputs. The country achieves outstanding results on the index's newer indicators relating to university-industry collaboration and international research engagement. The historical foundation of Italy's modern economy is its advanced physical infrastructure. Italy was a pioneer in high-speed rail construction — a development that drove mobility and logistics for decades. This status as an “efficiency leader” in innovation is directly relevant to the country's ability to deploy AI models within its SME sector. Strong university-industry linkages allow Italy to commercialise research faster than economies with higher nominal inputs but less cohesive ecosystems. To sustain this advantage, however, Italy must replicate its historical success in physical infrastructure on the digital frontier — becoming a leader in 5G and AI adoption and thereby preventing its considerable manufacturing prowess from being gradually marginalised.

The Economic Consequences of TDM Restrictions

Italy should be expected to experience a data loss (δ) of 50%. On the one hand, the key sectors that control copyrighted works and data are more fragmented than in Germany. On the other, the severity of sanctions for regulatory violations may induce heightened caution in how data is used — with a chilling effect on TDM activity even where it would otherwise be lawful.

For Italy, a unit transmission elasticity coefficient — a tau (τ) of 1.0 — appears most appropriate. The Italian economy is fundamentally dual in character: on one side sits an advanced machinery and manufacturing sector broadly comparable to Germany's; on the other, SMEs play an outsized role, alongside the creative and luxury industries that collectively define the “Made in Italy” identity. In creative and artisanal sectors, AI functions primarily as a support tool rather than a vehicle for full automation. Human intuition and adaptability are capable of partially absorbing shortcomings of imperfect digital tools — and this resilience offsets the more acute losses likely to be felt in the country's more precision-dependent industrial sectors.

Table 7. Economic consequences for Italy, €

	Delta (δ) = 0.3	Delta (δ) = 0.5	Delta (δ) = 0.7
Epsilon (ϵ) = 0.20	134 175 720 000	125 528 520 000	113 278 320 000
Epsilon (ϵ) = 0.35	127 257 960 000	113 134 200 000	94 542 720 000
Epsilon (ϵ) = 0.50	120 628 440 000	101 892 840 000	78 977 760 000

Source: Author's own study

For Italy, then, the most probable scenario suggests that opt-out will enable AI development to generate an additional €113.1 billion in GDP by 2035. This represents 79% of the maximum potential additional GDP that would be generated under a scenario of unconstrained AI development in the European Union. It is worth keeping in mind that regulatory approaches more burdensome than opt-out — in particular, mandatory licensing requirements (opt-in) — could deprive the Italian economy of any benefits from the development of this technology altogether.

⁴² WIPO, *Global Innovation Index 2025: Innovation at a Crossroads*.

In the Italian context, opt-out serves as a protective barrier for sectors in which cultural identity and heritage are inextricably bound up with economic value. This applies not only to literature, but also to the luxury and artisanal craftsmanship sectors, where “Made in Italy” is underpinned by a unique intellectual contribution. By shielding these assets from unauthorised generative mimicry, opt-out enables Italian creators to monetise their distinctiveness within the new digital ecosystem, protecting the national brand from devaluation.

Poland

Regulations

Poland implemented article 4 of the CDSM directive in September 2024, incorporating TDM mechanisms into articles 26(2) and 26(3) of the country’s copyright act. Despite the delay, the Polish legislature produced a framework widely regarded as one of the most formalistic in Europe. The pivotal provision is article 26(3)(2), which, in defining the concept of machine-readable format, draws directly on Poland’s act on open data and the re-use of public sector information.

This technical rigour means that a valid online opt-out reservation must be expressed in one of the following structured formats: XML, JSON, RDF or CSV. Polish law goes further still, imposing an additional requirement to attach metadata. This means that the robots.txt file may be considered insufficient. A Polish publisher must therefore not only make a machine-readable file available, but accompany it with appropriate descriptive information for the reservation to be legally binding. The Polish parliamentary debate, however, exposed deep tensions surrounding generative AI. Attempts to exclude AI model training entirely from the TDM exceptions were ultimately rejected as incompatible with the principle of technological neutrality enshrined in European Union law.

Poland, having transposed the CDSM directive with a delay of several years, finds itself at a unique political and technological juncture. A new opening in Polish digital policy has coincided with the mature phase of the debate on generative artificial intelligence, giving Warsaw the opportunity to leapfrog the stage of “defensive implementation.” Rather than inheriting the anxieties of a decade ago, the Polish legislature has the chance to formulate a legal framework that is not merely reactive, but actively shapes an innovation-friendly environment. For the government in Warsaw — which has staked its reputation on digitalisation and modern public services, as evidenced by the success of mObywatel — maintaining inclusive TDM rules is the key to building a data-driven economy, rather than one reliant solely on the export of low-complexity information technology (IT) services.

The recent parliamentary debate in Poland in which calls for a complete exclusion of AI model training from TDM exceptions were rejected, constitutes fundamental evidence of the political maturity of the Polish ecosystem. The fact that the legislature defended the principle of technological neutrality, despite strong pressure from traditional sectors, sends a clear message to global investors: Poland is choosing the progressive path. This “innovation victory” in parliament should not be treated as an isolated incident, but as a deliberate choice of direction. It makes Poland a leader in the Central and Eastern Europe (CEE) region — one capable of balancing the protection of creators with the real-world need to develop local models such as the Polish Large Language Model (PLLuM), under conditions of competition with global players.

AI and the Polish Economy

In the latest edition of the *Global AI Index*, Poland occupies a mid-table position among AI-developing nations — ranked 36th — clearly trailing the leaders while maintaining a stable presence in the global AI ecosystem.⁴³ This result reflects the country’s growing potential, particularly in human capital: Poland possesses a well-educated pool of IT specialists and a strong foundation in mathematics and computer science. Talent and education indicators rank among its relative strengths, consistent with the broader trend of central and eastern Europe emerging as a competence hub for the technology sector. A closer examination of the index’s component pillars, however, reveals significant structural constraints within Poland’s AI ecosystem. The most pressing challenges remain low levels of private investment, a limited number of globally competitive technology firms, and comparatively weak research commercialisation. The result is a model best characterised as “strong competence base, limited capital firepower” — typical of catching-up economies that possess genuine development potential but require intensified investment and deeper science-business linkages to compete more effectively in the global AI market.

PLLuM, the Polish Large Language Model, is a compelling illustration of why developing local AI solutions matters. It is widely regarded as a pivotal step toward Poland’s digital sovereignty, filling a gap that global technology corporations are structurally ill-equipped to address. While models such as GPT or Llama impress with their versatility, they frequently fall short when it comes to a deep understanding of Polish inflection, cultural nuance, and the specificities of the country’s legal and administrative systems. Developing homegrown AI solutions is therefore essential to ensuring both linguistic precision and data security in strategically sensitive sectors. The real-world impact of this initiative is already visible in its integration with e-government platforms such as mObywatel, where the model functions as an intelligent assistant helping citizens navigate the complexities of administrative regulation.

Poland records one of the lowest overall AI adoption rates among OECD countries — approximately 4% — a figure that stands in sharp contrast to the surprisingly strong

‘The challenge is not a shortage of technology, but the scale at which it transfers into the domestic economy.’

performance of large enterprises, where AI deployment reaches 25%, placing Polish corporate leaders on a par with their Italian counterparts. The country’s highly skilled technical workforce provides the foundation not only for catching up on adoption, but potentially for becoming an originator of solutions for other markets.

The key challenge, therefore, is not a shortage of technology, but the scale at which it transfers from the IT services sector into the broader domestic economy.⁴⁴

Poland ranks 39th in the *Global Innovation Index*,⁴⁵ yet retains its status as one of the fastest-advancing nations in innovation since 2013. The Polish economy has achieved deep diversification, emerging as a regional leader in ICT services exports. The report identifies Poland as one of the top growers, steadily climbing global value chains on the strength

⁴³ Joe White and Serena Casareo, *The Global AI Index*.

⁴⁴ World Bank, *Digital Progress and Trends Report 2025: Strengthening AI Foundations*.

⁴⁵ WIPO, *Global Innovation Index 2025: Innovation at a Crossroads*.

of its knowledge and technology outputs. Poland is undergoing a fundamental shift in its development model — moving from cost-competitiveness to technical competency-competitiveness. Sustaining this momentum will require a substantial increase in domestic R&D expenditure, enabling Polish firms to transition from being mere executors of others' technologies to becoming originating creators of breakthrough AI solutions.

The Economic Consequences of TDM Restrictions

Poland should be expected to experience the lowest data loss (δ) of the three European Union member states studied — 30%. This follows primarily from the late timing of the directive's implementation and the stringent formal requirements imposed on effective opt-out reservations. Every indication suggests that within this legal framework, mass or blanket exclusion of TDM sources will simply not be feasible.

The spectacular rise of **ElevenLabs** stands as compelling evidence that Polish technological ingenuity is setting global standards in generative artificial intelligence. Founded by Mateusz Staniszewski and Piotr Dąbkowski, the start-up has become a world leader in voice synthesis, transforming the media and entertainment industries. Poles also occupy front-rank positions at the world's leading AI organisations: Jakub Pachocki, serving as Chief Scientist at OpenAI, is widely regarded as one of the architects behind the success of the GPT model series, alongside Wojciech Zaremba (co-founder of OpenAI), Szymon Sidor and Aleksander Mądry, who bears responsibility for critical AI safety questions. That this talent was educated in Poland yet operates primarily from London, New York and San Francisco reflects a structural gap rather than a failure of ambition. Poland has the human capital; a stable, open TDM framework is the institutional signal that can convert that advantage into a domestic AI ecosystem rather than a continued export of talent.

For Poland, a transmission elasticity coefficient below unity — a tau (τ) of 0.8 — is the appropriate assumption. Poland is a hub for the business process outsourcing and shared services (BPO/SSC) sector, as well as for IT and software development. In these domains, AI is deployed primarily to enhance efficiency on medium-complexity tasks — code generation, reporting, customer service. As a result, a decline in AI tool quality does not render these tools unusable; it simply means they deliver a somewhat smaller positive impact than they otherwise would.

Table 8. Economic consequences for Poland, €

	Delta (δ) = 0.3	Delta (δ) = 0.5	Delta (δ) = 0.7
Epsilon (ϵ) = 0.20	55 327 488 000	52 516 608 000	48 534 528 000
Epsilon (ϵ) = 0.35	53 078 784 000	48 487 680 000	42 444 288 000
Epsilon (ϵ) = 0.50	50 923 776 000	44 833 536 000	37 384 704 000

Source: Author's own study

For Poland, then, the most probable scenario suggests that opt-out will enable AI development to generate an additional €53.1 billion in GDP by 2035. This represents 91% of the maximum potential additional GDP that would be generated under a scenario of unconstrained AI development in the European Union. It is worth keeping in mind that

regulatory approaches more burdensome than opt-out — in particular, mandatory licensing requirements (opt-in) — could deprive the Polish economy of any benefits from the development of this technology altogether.

‘The opt-out mechanism preserves €942 billion in additional potential GDP’

For Poland, which is in a phase of dynamic growth in modern creative industries — from game development to digital publishing — opt-out represents an opportunity to shape healthy relations between creators

and technology at an early stage. It allows the country to avoid the trap of “technological rent” extracted by global platforms at the expense of local content producers. In doing so, the Polish creative sector can become a partner to, rather than merely a passive supplier for, the AI industry — fostering the development of a diversified, shock-resistant knowledge-based economy.

Annexes

Extracts from Global Rankings for AI Capacity, Adoption and Innovation

Table 9. Selected country rankings from The Global AI Index

	United States	Singapore	Germany	Japan	Italy	Poland
Overall	1	3	7	11	24	36
Talent	1	6	3	23	19	27
Infrastructure	1	3	11	5	27	28
Operating environment	2	48	8	53	1	36
Research	1	3	8	20	21	41
Development	1	5	11	14	45	32
Government strategy	2	10	8	12	13	39
Commercial	1	4	9	14	43	46
Scale	1	11	5	9	16	31
Intensity	3	1	15	31	32	37

Source: Joe White and Serena Casareo, The Global AI Index, Tortoise Media, 19 September 2024.

Overall: The composite score reflecting a country's general standing in the global AI rankings.

Talent: The availability of skilled human capital, from software developers to world-class machine learning researchers.

Infrastructure: The quality of underlying technological foundations, including computational power, data centres and internet connectivity speeds.

Operating environment: The stability of the regulatory framework, levels of public trust, and the ease of conducting technology business.

Research: The intensity and prestige of scientific and research activity, measured by the volume of AI publications and citations.

Development: The effectiveness of translating theory into practice, assessed through patent activity and engagement in software development projects.

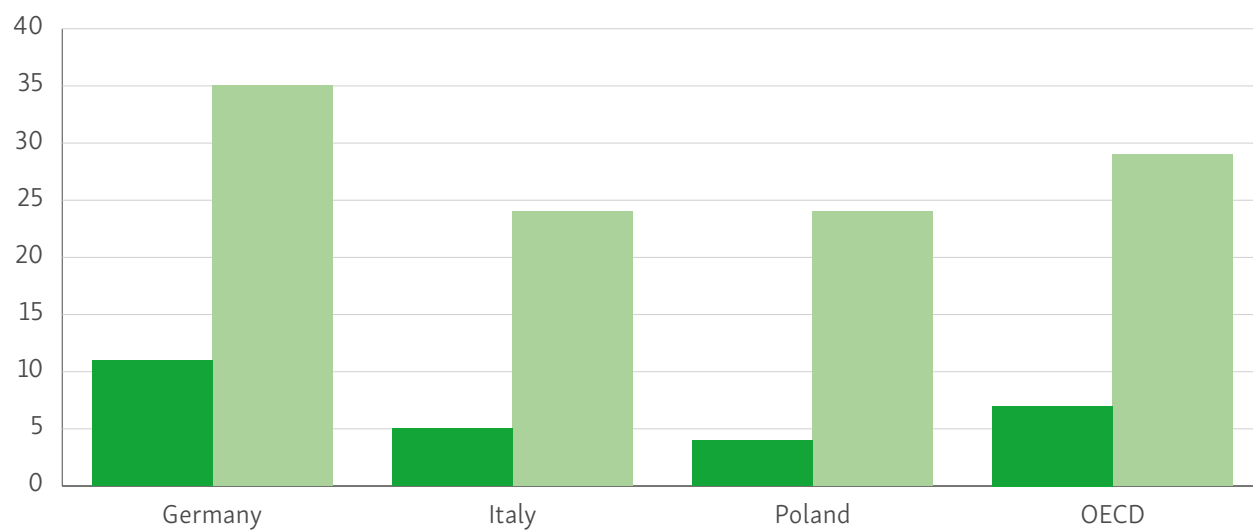
Government strategy: The degree to which the state funds and supports AI technology through national development strategies and plans.

Commercial: The scale of private investment, the dynamism of the start-up market, and the level of venture capital engagement.

Scale: The absolute size of a country's AI market and available resources.

Intensity: The ratio of AI achievements to a country's population size and GDP.

Figure 4. Adoption rates of AI by enterprises with 10 employees or more in the business sector (excluding financial services), 2023



Source: OECD, OECD Digital Economy Outlook 2024 (Volume 1) (OECD Publishing, 2024).

Table 10. Selected Country Rankings from the WIPO Global Innovation Index (GII) 2025

Note: Where a cell contains two values, these are the score out of 100 and the global ranking, respectively.

Indicator / category	Germany	Italy	Poland	United States	Japan	Singapore
GII 2025 rank	11	28	39	3	12	5
Output rank	8	19	36	3	14	9
Input rank	15	37	45	6	12	1
GDP per capita, PPP\$	70.93	60.993	51.629	86.601	53.059	148.186
1. Institutions	70.2 / 23	54.2 / 55	49.7 / 68	74.3 / 16	70.9 / 22	98.7 / 1
1.1 Institutional environment	74.8 / 23	61.7 / 49	60.7 / 50	75.8 / 22	84.9 / 10	98.7 / 1
1.2 Regulatory environment	85.1 / 15	61.7 / 45	64.1 / 40	81.7 / 20	85.1 / 16	97.4 / 1
1.3 Business environment	50.7 / 57	39.4 / 79	24.2 / 113	65.5 / 25	42.6 / 72	100.0 / 1
2. Human capital and research	61.0 / 4	45.1 / 32	38.6 / 43	55.8 / 13	53.5 / 18	63.3 / 2
2.1 Education	61.8 / 32	59.4 / 41	60.9 / 37	60.3 / 40	66.1 / 13	57.4 / 50
2.2 Tertiary education	53.3 / 6	33.4 / 57	30.9 / 65	31.4 / 63	27.4 / 77	70.9 / 2
2.3 Research and development (R&D)	68.0 / 6	42.4 / 25	23.8 / 44	75.7 / 2	66.9 / 7	61.6 / 12
3. Infrastructure	54.9 / 28	54.9 / 27	48.1 / 51	54.3 / 32	57.6 / 17	57.2 / 19
3.1 Information and communication technologies (ICTs)	89.1 / 26	82.6 / 52	88.2 / 33	94.5 / 9	93.2 / 14	96.5 / 4
3.2 General infrastructure	48.9 / 25	42.1 / 39	36.4 / 53	54.1 / 13	53.2 / 15	57.4 / 11
3.3 Ecological sustainability	26.7 / 48	40.0 / 17	19.6 / 73	14.3 / 98	26.4 / 50	17.7 / 84
4. Market sophistication	50.8 / 22	40.4 / 52	37.7 / 64	75.0 / 1	59.4 / 10	61.5 / 6
4.1 Credit	47.0 / 29	25.1 / 77	20.4 / 89	75.2 / 3	66.5 / 8	50.3 / 26
4.2 Investment	16.3 / 38	7.4 / 57	6.9 / 60	55.3 / 4	23.6 / 25	58.6 / 3
4.3 Trade, diversification and market scale	89.2 / 4	88.8 / 6	85.7 / 11	94.6 / 1	88.2 / 7	75.7 / 49
5. Business sophistication	54.5 / 13	40.0 / 31	39.5 / 32	65.9 / 1	57.8 / 6	63.0 / 3
5.1 Knowledge workers	53.1 / 24	39.0 / 59	50.7 / 28	69.9 / 3	55.1 / 20	61.6 / 9
5.2 Innovation linkages	67.5 / 7	47.5 / 26	27.6 / 59	77.0 / 2	68.6 / 6	63.8 / 12
5.3 Knowledge absorption	42.8 / 20	33.5 / 43	40.1 / 27	50.9 / 7	49.6 / 8	63.7 / 1
6. Knowledge and technology outputs	49.6 / 11	40.7 / 17	28.5 / 42	60.0 / 3	48.2 / 12	53.1 / 7
6.1 Knowledge creation	54.3 / 10	38.9 / 24	24.6 / 41	57.0 / 7	56.9 / 8	35.2 / 27
6.2 Knowledge impact	44.7 / 15	34.9 / 38	29.1 / 52	78.7 / 1	35.0 / 35	68.9 / 2
6.3 Knowledge diffusion	49.9 / 14	48.4 / 16	31.7 / 38	44.3 / 20	52.8 / 6	55.2 / 4
7. Creative outputs	55.6 / 8	45.2 / 20	37.0 / 33	56.6 / 5	46.4 / 18	49.2 / 15
7.1 Intangible assets	62.6 / 6	57.7 / 9	40.1 / 32	56.2 / 11	58.0 / 8	39.9 / 33
7.2 Creative goods and services	27.7 / 39	27.5 / 40	23.5 / 49	45.8 / 9	32.9 / 21	48.7 / 6
7.3 Online creativity	69.5 / 10	37.7 / 39	44.2 / 33	68.2 / 11	36.5 / 40	68.2 / 12

Source: World Intellectual Property Organization, Global Innovation Index 2025: Innovation at a Crossroads (World Intellectual Property Organization, 2025).

Bibliography

- Ahmed, Ahmed, A. Feder Cooper, Sanmi Koyejo and Percy Liang. “Extracting Books from Production Language Models,” *arXiv*, 6 January 2026, <https://doi.org/10.48550/arXiv.2601.02671> [accessed 16 August 2026]
- Bell, Karissa. “Meta Signs a Multimillion Dollar AI Licensing Deal with News Corp,” *Engadget*, 03 March 2026, <https://www.engadget.com/ai/meta-signs-a-multimillion-dollar-ai-licensing-deal-with-news-corp-234157902.html> [accessed 16 August 2026]
- Biga, Bartłomiej. *The Economics of Intellectual Property and Openness: The Tragedy of Intangible Abundance* (Routledge, 2021)
- Bloomberg. “OpenAI to Pay Axel Springer Tens of Millions to Use News Content,” 13 December 2023, <https://www.bloomberg.com/news/articles/2023-12-13/openai-axel-springer-ink-deal-to-use-news-content-in-chatgpt> [accessed 16 August 2026]
- Boldrin, Michael and David K. Levine. *Against Intellectual Monopoly* (Cambridge University Press, 2008)
- Buchanan, James M. and Yong J. Yoon. “Symmetric Tragedies: Commons and Anticommons,” *The Journal of Law and Economics*, 43(1), 2000, pp. 1–14, <https://doi.org/10.1086/467445> [accessed 16 August 2026]
- Cahn, David. “AI’s \$600B Question,” *Sequoia Capital*, 20 June 2024, <https://sequoiacap.com/article/ais-600b-question/> [accessed 16 August 2026]
- Campbell, Ian Carlos. “Perplexity Has Cooked up a New Way to Pay Publishers for Their Content,” *Engadget*, 25 August 2025, <https://www.engadget.com/ai/perplexity-has-cooked-up-a-new-way-to-pay-publishers-for-their-content-204255019.html> [accessed 16 August 2026]
- Cohen, Lauren, Umit G. Gurun and Scott Duke Kominers. “Patent Trolls: Evidence from Targeted Firms,” *Management Science*, 65(12), 2019, pp. 5461–5486, <https://doi.org/10.1287/mnsc.2018.3147> [accessed 16 August 2026]
- Cooter, Robert and Thomas Ulen. *Law and Economics* (Scott, Foresman & Co., 1988)
- Covert, Ian, Wenlong Ji, Tatsunori Hashimoto and James Zou. “Scaling Laws for the Value of Individual Data Points in Machine Learning,” *arXiv*, 30 May 2024, <https://doi.org/10.48550/arXiv.2405.20456> [accessed 16 August 2026]
- Dow Jones. “Dow Jones Factiva Surpasses 8,000 Licensed Sources for GenAI Use,” press release, *Business Wire*, 20 January 2026, <https://www.businesswire.com/news/home/20260120489743/en/Dow-Jones-Factiva-Surpasses-8000-Licensed-Sources-for-GenAI-Use> [accessed 15 August 2026]
- Fletcher, Richard. “How Many News Websites Block AI Crawlers?” *Reuters Institute for the Study of Journalism*, 22 February 2024, <https://reutersinstitute.politics.ox.ac.uk/how-many-news-websites-block-ai-crawlers> [accessed 16 August 2026]

- Frank, Tim. “Building Toward a Sustainable Content Economy for the Agentic Web,” *Microsoft Advertising*, 03 February 2026, <https://about.ads.microsoft.com/en/blog/post/february-2026/building-toward-a-sustainable-content-economy-for-the-agentic-web> [accessed 16 August 2026]
- Guaglione, Sara. “News/Media Alliance Signs AI Licensing Deal to Unlock Recurring RAG Revenue for Small and Mid-Sized Publishers,” *Digiday*, 23 March 2026, <https://digiday.com/media/news-media-alliance-signs-ai-licensing-deal-to-unlock-recurring-rag-revenue-for-small-and-mid-sized-publishers/> [accessed 16 August 2026]
- Hardin, Garrett. “The Tragedy of the Commons: The population problem has no technical solution; it requires a fundamental extension in morality,” *Science*, 162(3859), 1968, pp. 1243–1248, <https://doi.org/10.1126/science.162.3859.1243> [accessed 15 August 2026]
- Hargreaves, Ian. *Digital Opportunity: A Review of Intellectual Property and Growth* (Intellectual Property Office, 2011)
- Heller, Michael A. “The Tragedy of the Anticommons: Property in the Transition from Marx to Markets.” *Harvard Law Review*, vol. 111, no. 3, 1998, pp. 621–688
- Heller, Michael A. and Rebecca S. Eisenberg. “Can Patents Deter Innovation? The Anticommons in Biomedical Research,” *Science*, 280(5364), 1998, pp. 698–701, <https://doi.org/10.1126/science.280.5364.698> [accessed 16 August 2026]
- Hoffmann, Jordan, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford et al. “Training Compute-Optimal Large Language Models,” *arXiv*, 29 March 2022, <https://doi.org/10.48550/arXiv.2203.15556> [accessed 16 August 2026]
- Kaplan, Jared, Sam McCandlish, Tom Henighan, Tom B. Brown, Benjamin Chess, Rewon Child et al. “Scaling Laws for Neural Language Models,” *arXiv*, 20 January 2020, <https://doi.org/10.48550/arXiv.2001.08361> [accessed 16 August 2026]
- Keller, Paul. “GPT-NL Respects Copyright — Cui Bono?” *Open Future*, 26 June 2026, <https://openfuture.eu/blog/gpt-nl-respects-copyright-cui-bono/> [accessed 16 August 2026]
- Landes, William M. and Richard A. Posner. *The Economic Structure of Intellectual Property Law* (Harvard University Press, 2009) <https://doi.org/10.2307/j.ctv1q3z2pn>
- Longpre, Shayne, Robert Mahari, Ariel Lee, Campbell Lund, Hamidh Oderinwale, William Brannon et al. “Consent in Crisis: The Rapid Decline of the AI Data Commons,” *arXiv*, 24 July 2024, <https://doi.org/10.48550/arXiv.2407.14933> [accessed 16 August 2026]
- Maslej, Nestor, Loredana Fattorini, Raymond Perrault, Yolanda Gil, Vanessa Parli, Njenga Kariuki et al., “Artificial Intelligence Index Report 2025,” *arXiv*, 20 November 2025, <https://doi.org/10.48550/arXiv.2504.07139> [accessed 15 August 2026]
- Organisation for Economic Co-operation and Development (OECD), *OECD Digital Economy Outlook 2024 (Volume 1)* (OECD Publishing, 2024), <https://doi.org/10.1787/a1689dc5-en> [accessed 17 August 2026]

- Publishers Association. *Content Superpower: UK Publishing and the AI Licensing Market* (Publishers Association, 2026), <https://www.publishers.org.uk/wp-content/uploads/2026/03/Publishers-Association-Content-Superpower-UK-publishing-and-the-AI-licensing-market.pdf> [accessed 16 August 2026]
- Rijgersberg, Edwin. “GPT-NL as a Dutch Sovereign Alternative? The Preliminary Benchmarks Don’t Look Good,” *GoingDutch.ai*, 18 June 2026, <https://goingdutch.ai/en/posts/gpt-nl-als-soeverein-alternatief/> [accessed 16 August 2026]
- Sorscher, Ben, Robert Geirhos, Shashank Shekhar, Surya Ganguli and Ari S. Morcos. “Beyond Neural Scaling Laws: Beating Power Law Scaling via Data Pruning,” *arXiv*, 29 June 2022, <https://doi.org/10.48550/arXiv.2206.14486> [accessed 16 August 2026]
- Thelle, Martin H., Anders Thor Lundberg, Bodile Emilie Hovmand, Hans Henrik Woltmann, Laura Virtanen, Nikolaj Tranholm-Mikkelsen et al., *The Economic Opportunity of AI in the EU. Beyond the Hype – Boosting EU Competitiveness with Generative AI* (Implement Consulting Group, 2024)
- Tow Center for Digital Journalism. *AI deals and lawsuits*, n.d., <https://tow.cjr.org/ai-deals-lawsuits/> [accessed 10 June 2026]
- Universal Music Group. “Universal Music Group and Udio Announce Udio’s First Strategic Agreements for New Licensed AI Music Creation Platform,” press release, *PR Newswire*, 29 October 2025, <https://www.prnewswire.com/news-releases/universal-music-group-and-udio-announce-udios-first-strategic-agreements-for-new-licensed-ai-music-creation-platform-302599129.html> [accessed 16 August 2026]
- White, Joe and Serena Casareo. *The Global AI Index*, Tortoise Media, 19 September 2024, <https://www.tortoisemedia.com/data/global-ai> [accessed 16 August 2026]
- World Bank. *Digital Progress and Trends Report 2025: Strengthening AI Foundations* (World Bank, 2025), <https://doi.org/10.1596/978-1-4648-2264-3> [accessed 16 August 2026]
- World Intellectual Property Organization (WIPO). *Global Innovation Index 2025: Innovation at a Crossroads* (World Intellectual Property Organization, 2025), <https://doi.org/10.34667/tind.58864> [accessed 16 August 2026]

About the Author

Bartłomiej Biga, PhD, associate professor at the Krakow University of Economics, is an attorney, an expert at the Analysis Centre of the Jagiellonian Club, Vice-Chair of the Programme Council of the Open Eyes Economy Summit and an active populariser of economic knowledge (bartlomiejbiga.pl). His work spans the economic analysis of law across three dimensions: theoretical and academic, practical and applied, and public and journalistic. His primary areas of focus are intellectual property, behavioural economics and the digital economy.

Copyright © The Lisbon Council 2026



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International Licence

The Lisbon Council asbl

IPC-Résidence Palace
155 Rue de la Loi
1040 Brussels, Belgium
T +32 2 647 9575
www.lisboncouncil.net
info@lisboncouncil.net