Focus "Digitalisation in Science" by the Alliance of German Science Organisations

Arrangements on artificial intelligence in licence agreements

Recommendations for action of September 2024

Team of authors:

Christian Agi, MPDL Services gGmbH;
Michael Beurskens, University of Passau;
Marion von Francken-Welz, University Library Mannheim;
Judith Ludwig, TIB Hannover;
Bernhard Mittermaier, FZ Jülich;
Heinz Pampel, HU Berlin.

The online version of this publication can be found at https://doi.org/10.5281/zenodo.13837688
You can find the original German version at https://doi.org/10.5281/zenodo.13837664

This publication is a translation of “Regelungen zu künstlicher Intelligenz in Lizenzverträgen”, provided by TIB – Leibniz Information Centre for Science and Technology.

Recommendations for action:
Arrangements on artificial intelligence in licence agreements[1]

Contents

A. Introduction

Artificial intelligence (AI) methods open up a wide range of new possibilities for science in digital information production, processing and analysis. The Alliance of German Science Organisations has dedicated itself to shaping science-friendly framework conditions for digital research under the motto "Shaping digitality - promoting openness and sovereignty", considering legal, technological, financial and organisational aspects.[2] The potential of AI was discussed in the discussion paper "Digital Change in the Sciences and Humanities" by the German Research Foundation (DFG), among others.[3] The importance of scientific publications, including in the digital working environment, was emphasised by

[1] The recommendations for action were developed in the Task Force: Arrangements on Artificial Intelligence in Licence Agreements (TF: RKIL). The members of the task force were: Christian Agi, Prof Dr Michael Beurskens (spokesperson), Dr Marion v. Francken-Welz, Judith Ludwig (spokesperson), Dr Bernhard Mittermaier, Prof Dr Heinz Pampel. All internet sources cited were last accessed on 18/09/2024.

[2] https://doi.org/10.5281/zenodo.11189484.

[3] https://doi.org/10.5281/zenodo.4191345.

the scientific organisations in their 2015 "Positions on the creation of a science-adequate open access publication market"[4] and in their statement on the implementation of EU copyright directives[5] .

As technology advances, there is a need to give researchers the opportunity to work with scientific publications beyond their intellectual reception. This includes the use methods of so-called artificial intelligence, but also text and data mining (TDM) frequently associated therewith, which enables the processing of content by machines (→ C.II.3.c).[6] Shaping the legal structure of contracts with publication service providers in the interests of science is of central importance here. On the one hand, there is often an unequal **distribution of negotiating power** between publication service providers and research institutions and, on the other hand, there is still **uncertainty with regard to the mandatory legal framework**, particularly in the field of artificial intelligence.

The unclear legal situation and restrictive, but possibly inapplicable contractual provisions (→ C.II) can also lead to a "**chilling effect"**[7] in such a way as to cause the leeway expressly desired by the legislator in the interest of risk minimisation not to be used and research not to be carried out at all or possibly only under delayed and more difficult conditions. This does not only impair individual academic freedom (Art. 5 (3) first sentence Basic Law of the Federal Republic of Germany (GG)) and the digital sovereignty of researchers, but can also have a negative impact on Germany as a research location as a whole, as the innovative capacity of science is restricted. At the same time, overly restrictive clauses can be used as a reference in subsequent negotiations and thus become permanently entrenched as a general principle to the detriment of science. This development must be prevented so that innovative working practices in the digital world remain open to science.

The science-friendly design of applicable provisions is of great importance to the science community. Thus, in 2022, "Guidelines on text and data mining for research purposes in Germany"[8] were developed as part of NFDI4Ing, and in 2024 the International Coalition of Library Consortia (ICOLC)[9] and the Association of Research Libraries (ARL)[10] made demands regarding AI in licence agreements.

This handout provides a compact overview of the permitted and prohibited practices regarding TDM and AI under German law and makes recommendations for designing and shaping such contracts. It is primarily aimed at persons who negotiate licence agreements with publication service providers and who have to work with these agreements in libraries, but also indirectly at researchers who want to use licensed content for TDM and AI applications.

[4] https://doi.org/10.5283/epub.33586.

[5] https://www.allianz-der-wissenschaftsorganisationen.de/themen-stellungnahmen/stellungnahme-der-allianz-konsultation-zur-umsetzung-urheberrecht/.

[6] https://royalsociety.org/news-resources/projects/science-in-the-age-of-ai/.

[7] For a monograph on this, see J. Staben, Der Abschreckungseffekt auf die Grundrechtsausübung, 2016.

[8] https://oa.tib.eu/renate/handle/123456789/10352.

[9] ICOLC Statement on AI in Licensing (22.03.2024), https://icolc.net/statements/icolc-statement-ai-licensing. An ICOLC Negotiation Strategy is available for ICOLC members in the internal area; alternatively, it can be requested by authorised institutions, e.g. via the MPDL Services gGmbH.

[10] https://www.arl.org/news/association-of-research-libraries-releases-guiding-principles-for-artificial-intelligence.

B. Status quo: licence agreement negotiations

Since the adoption of the Directive on Copyright in the Digital Single Market (DSM Directive)[11] , many licence agreements contain provisions on TDM, which often stipulate high requirements for its corresponding use. Normally, at first stipulations are made which permit TDM and then reference is made to dedicated web pages of the publication service provider describing the procedure by which one may perform TDM. Some publication service providers provide programming interfaces (APIs) for TDM. They enable more efficient mining than what can be achieved by web crawling, but also allow the publication service provider to control what is downloaded. A cross-provider option for performing TDM using an API is offered by Crossref. One example of the restrictive business behaviour of publication service providers is the charging of a fee, although the TDM use in question is required to be free of charge for the purposes of scientific research by mandatory law (Art. 3 DSM Directive = section 60d German Copyright Act (UrhG)). The requirement to have an individual agreement with the individual user for TDM and, where applicable, the requirement of personal registration, which is associated with unnecessary data collection, are also questionable.

In addition to these TDM clauses designed as supposedly necessary permissions, the use of AI methods is often very extensively restricted or prohibited (including for research purposes) by independent clauses in licence agreements.[12] For example, the use of the licensed content as training data for AI systems and its input into AI systems (including automated translation, summarisation, concordance formation) is supposed to be contractually restricted or prohibited. However, there is an overlap with TDM, which can and may also use AI methods (→ C.II.3.c). In this respect and also in other scenarios, contractual clauses prohibit AI methods that are, however, permitted by permissions stipulated under EU law (→ C.II).[13]

C. Legal classification of provisions on "artificial intelligence" in licence agreement negotiations

C.I. Definition

Licence agreements often refer to the term "artificial intelligence" without clearly defining it. The definition in Art. 3 No. 1 AI Regulation[14] could currently serve as an aid to interpretation: An "AI system *" means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments".*

However, these and similar abstract definitions go far beyond the text-related generative AI[15] that is often the focus: they already include trivial algorithms such as text completion ("autocomplete"), spelling

[11] 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.

[12] For example, the DEAL contracts with Wiley, Springer Nature and Elsevier, https://deal-konsortium.de/vertraege.

[13] See also ECJ, judgement of. 27. 6. 2013 - C-457/11, C-458/11, C-459/11, C-460/11 (VG Wort/Kyocera) para. 37 on the precedence of limitations over a licence (what is permitted by law can no longer be permitted by a licence agreement).

[14] Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Regulation).

[15] OpenAI ChatGPT, Google Gemini, Anthropic Claude, Mistral Le Chat, Aleph Alpha Luminous, etc.

and grammar checking and even Excel tables that use formulas. Even outside the statutory definition, AI is used as a blanket generic term that encompasses all technical designs that come close to (individual) human abilities such as logical thinking, learning, planning and even creativity.[16]

Many of the conceivable application scenarios for AI, such as the use of locally installed systems for translation, do not constitute copyright-relevant acts or fall under statutory rules providing for permissions (limitations), so that they cannot be the subject of contractual prohibitions to the detriment of users who are not themselves parties to the contract. In addition, the use of AI tools by users cannot be effectively controlled, let alone prevented. Therefore, when considering the law, the focus should not be on "artificial intelligence" as such, but instead on specific user acts that are to be permitted or prohibited by the contract.

A blanket reference in the contract to the term "artificial intelligence" is unsuitable due to its broad scope as described above; instead, one should specify concrete user acts that are to be permitted or prohibited by the contract.

C.II. Legal categorisation of the acts of use

C.II.1 Lex loci protectionis

In order to be able to make a legal assessment, it must first be determined which law applies to the respective act of use. According to the so-called lex loci protectionis (Art. 8 (1) Rome II Regulation[17]; Art. 5 (2) second sentence Rome Convention[18]), the law of the country in which the alleged act of infringement was committed is mandatorily applicable to prohibitions or permissions regarding the use of copyright content. Any agreement on the applicability of a different copyright law is (necessarily) invalid (Art. 8 (3) Rome II Regulation).[19]

According to the so-called lex loci protectionis principle, German copyright law is mandatorily applicable to the use of copyrighted works if the relevant acts (e.g. reproduction, distribution) take place in Germany.

C.II.2 Mandatory permission to use

The permission to use TDM cannot be restricted by contract; for scientific purposes (Art. 3 DSM Directive = section 60d UrhG, → C.II.3.c), for acts of reproduction by libraries and other cultural heritage institutions (Art. 5 (2) lit. c InfoSoc Directive[20] = section 60e (1) UrhG, → C.II.3.e) as well as for the conversion of texts into an accessible format for the benefit of people with visual or reading disabilities (Art. 3 f. Marrakesh Directive[21] = section 45d UrhG). Furthermore, the customary use of databases and

[16] https://www.europarl.europa.eu/topics/en/article/20200827STO85804/what-is-artificial-intelligence-and-how-is-it-used.

[17] Regulation (EC) No 864/2007 of the European Parliament and of the Council of 11 July 2007 on the law applicable to non-contractual obligations ("Rome II").

[18] Berne Convention for the Protection of Literary and Artistic Works revised in Paris on 24 July 1971.

[19] See ECJ judgement of. 27 June 2013 - C-457/11, C-458/11, C-459/11, C-460/11 (VG Wort/Kyocera) para. 37; ECJ judgement of 5 March 2015 - C-463/12 para. 66 (Copydan/Nokia). 5.3.2015 - C-463/12 para. 66 (Copydan/Nokia) on the precedence of limitations over a licence (what is permitted by law can no longer be permitted by a licence agreement).

[20] Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, OJ L 167, 22.6.2001, p. 10.

[21] Directive (EU) 2017/1564 on the use of certain works protected by copyright and related rights and amending Directive 2001/29/EC, OJ L 2017 L 242, p. 6.

software must not be impaired by contracts (section 69g (2) UrhG; section 55a second sentence UrhG; section 87e UrhG).

Outside of these mandatory provisions, the prevailing - albeit controversial - view is that restrictions or separate contractual arrangements are possible (→ C.II.3.d.).

C.II.3 Acts of use in connection with AI systems

Acts of use are only relevant under copyright law if the specific act of exploitation - i.e. reproductions in particular - is reserved by law to the authors (Art. 2 f. InfoSoC Directive=sections 15 et seqq. UrhG). Thus, as a general rule, under German law, the creation of adaptations and rearrangements is generally permitted and only their publication and exploitation is reserved to the authors (→ C.II.3.a). However, regardless of any contractual agreements, even an act of use that is generally reserved to the author may be justified by a statutory permission (limitation, Art. 5 InfoSoc Directive=sections 44a et seqq. UrhG).

The AI Regulation does not contain any explicit permissions for the use of copyright-protected content in this sense. In connection with the provisions of the Copyright Act, a distinction must be made between two cases of copyright-relevant use of works in connection with AI systems: On the one hand, copyright-protected content can be used as mere **input data**[22] for processing by AI systems without the AI system itself being changed as a result. This is usually the case, for example, when editing a text with the help of local summarisation or translation tools.

On the other hand, copyrighted content can be used as **training data**[23] for the development or enhancement of existing AI systems. Anyone who trains their own (specific) AI system falls into this group. Query data (prompts and transmitted content) and generated results may also be reused when using third-party services if no other arrangement was made for his or if this was not excluded prior to use.[24] In this respect, many commercial AI providers require assurances from their users that they are authorised to grant rights of further use of all input data (i.e. including any third-party content) - however, users generally are not so authorised.

C.II.3.a. Adaptations and transformations (section 23 UrhG)

The production (as opposed to the publication and exploitation) of "adaptations and transformations" is permitted by law and is not dependent on a contractual grant of rights (section 23 (1) first sentence UrhG). Adaptations serve the purpose of adapting the original work to specific circumstances, e.g. translating it into another language or another art form or of adapting it for other means of expression.[25] When transformed, the work is used in a modified form, but without a function serving the original work.[26]

Thus, translating or summarising texts with the help of AI systems is also permitted without the rightholders' consent and without remuneration, provided that it is ensured that the works are used exclusively as input data. This can be done by operating a purely local system or by applying the

[22] See Art. 3 No. 33 AI Regulation.

[23] See Art. 3 No. 29 AI Regulation.

[24] e.g. https://openai.com/policies/eu-terms-of-use/: "If you do not want us to use your Content to train our models, you have the option to opt out by updating your account settings." - So, if you don't want your entries (e.g. in the summary) to be used as training data, you must explicitly switch this off in the settings beforehand.

[25] Official statement of reasons, Bundestag Printed Paper IV/270, 51.

[26] Spindler/Schuster/Wiebe, 4th ed. 2019, UrhG section 23 para. 4-7, with further references.

applicable terms of use accordingly when using third-party systems. In contrast, adaptation does not justify the use of the works as training data.

Contracts cannot effectively prohibit end users from using texts or images made available from databases as input data for AI for adaptation and transformation.

C.II.3.b. Authorised temporary reproductions (Art. 2 InfoSoc Directive = Section 16 UrhG)

Reproductions generally require legal permission (limitation) or consent (Art. 2 InfoSoc Directive = Section 16 UrhG), so that, as a rule, contractual provisions for retrieval, storage and any transfer are crucial.[27] Reproductions are all physical representations (including on a data carrier) that are capable of making the work directly or indirectly perceptible to the human senses in any way (e.g. by displaying it on a screen).[28] The type of material and production process - e.g. analogue or digital, manual or mechanical - is irrelevant, so that downloading and uploading texts or images are covered, just as are sending them by e-mail or printing them out.[29]

However, if a merely **temporary** reproduction of the original is required for an otherwise permitted purpose, the EU-wide standardised statutory rule (limitation) of Art. 5 (1) InfoSoc Directive (implemented in Section 44a UrhG) providing for permission comes into consideration. It permits temporary acts of reproduction to the extent that they are only transient or incidental, constitute an integral and essential part of a technical process, serve the lawful use of a work or other protected subject matter and have no independent economic significance. Such temporary reproductions are required when used as input data for the creation of adaptations and transformations (→ C.II.3.a). This statutory exception is mandatory and permitted without requiring the rightholders' consent and without remuneration. Contracts cannot provide for any restrictions or prohibitions to the detriment of end users.

Only temporary reproduction is always permitted, provided that 1. it is only transient or incidental and 2. it has no commercial significance of its own and 3. it is necessary for a subsequent lawful use of the work (in particular an adaptation or transformation).

C.II.3.c. TDM for scientific purposes (Art. 3 DSM Directive = Section 60d UrhG)

Reproductions and extractions for TDM for scientific research purposes are permitted for research organisations and cultural heritage institutions as well as individual researchers for all works or other protected subject matter to which they have lawful access (Art. 3 DSM Directive = Section 60d (1) UrhG). Such reproductions may not be subject to any contractual restrictions (remuneration, prohibition, etc.). However, pursuant to sections 60g (1) and 137o UrhG, in Germany, this only applies to contracts concluded from 1 March 2018 onwards; old contracts may therefore provide for restrictions.

[27] However, it is important to note here that this relates exclusively to "reproduction"; as the act of exploitation; any subsequent publication or exploitation based thereon must be considered separately and regularly requires an additional legal basis.

[28] Official statement of reasons, Bundestag Printed Paper IV/270, 47.

[29] Dreier/Schulze/Schulze, 7th ed. 2022, UrhG section 16 para. 6-15.

TDM includes any " any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations" (Art. 2 No. 2 DSM Directive, largely the same as section 44b (1) UrhG). The definition of TDM is explicitly linked to "automated analysis", which goes beyond the mere copying of content (crawling). Pure archiving (without analysis) is not covered by this permission. Whether input into external AI systems is permitted has not yet been conclusively clarified in legal terms. Some argue that the authorised acts of use may also be carried out by commissioned third parties.[30] But this is contradicted by the fact that section 60d (4) UrhG does not authorise making reproductions available to third parties except within narrow limits. However, the TDM provision covers the training of an own AI model (but not the support of third parties, in particular of commercial providers in training their systems): Recital 105 AI Regulation explicitly clarifies that the use of copyright-protected works as part of TDM for the development and training of general-purpose AI models is also permitted without requiring the authors' consent if it is done for the purposes of scientific research. This means that the development of search tools or purely statistical analysis tools, for example, cannot be prohibited by contract. The creation of purely local, unpublished AI systems cannot be excluded by contract either.

Explicitly not covered by the TDM-permission is the downstream use of the reproductions by third parties, for example in cross-institutional research groups for review purposes or general publication (Recital 15 at the end DSM Directive). In section 60d (4) UrhG, the German legislator has made a special arrangement for making reproductions available to certain third parties who are only authorised to use the work within this framework.

Contracts concluded from 1 March 2018 onwards cannot effectively restrict the use of copyrighted works by TDM for scientific purposes, including the creation of internal scientific AI systems. Making reproductions available to third parties without contractual permission is only permitted under the conditions set out in section 60d (4) UrhG.

C.II.3.d. TDM for non-scientific purposes (Art. 4 DSM Directive = section 44b UrhG)

Reproductions and extractions for TDM may be made for other persons and/or for other purposes if they concern lawfully accessible works and other protected subject matter and these are not subject to an appropriate reservation of use (Art. 4 DSM Directive implemented in: section 44b UrhG). Thus, uses are only permitted if the rightholder has not reserved them. A reservation of use for works accessible online is only effective if it is made in machine-readable form. According to recital 18 DSM Directive, metadata and terms and conditions of a website or service are also sufficient to comply with machine-readability. This means that publication service providers can in principle limit the use of TDM to non-commercial scientific purposes.

Contracts can effectively make provision for the use of copyrighted works for TDM for non-scientific purposes.

[30] BeckOK UrhR/Bomhard, 42nd ed. 15.2.2024, UrhG section 60d para. 23; Dreier/Schulze/Dreier, 7th ed. 2022, UrhG section 60d para. 9; Raue, Benjamin: Die Freistellung von Datenanalysen durch die neuen Text- und Data Mining-Schranken (sections 44b, 60d UrhG), ZUM 2021, 793 (801).

C.II.3.e. Reproduction acts by libraries and other cultural heritage institutions (Art. 5 (2) lit. c InfoSoc Directive = section 60e (1) UrhG)

Section 60e (1) UrhG permits public libraries and cultural heritage institutions (in conjunction with section 60f (1) UrhG) to reproduce or have reproduced a work from their holdings or exhibition for the purposes of making available, indexing, cataloguing, preservation and restauration, including more than once and with technically necessary alterations[31] . The classic example of such an act of use would be machine-aided subject indexing in libraries. These acts of use are also permitted with the help of AI systems (in particular OCR) without the rightholder's consent and without remuneration. In contrast, the training of a model outside of the aforementioned purposes (making available, indexing, cataloguing, preservation and restoration) is not covered by the limitation. However, pursuant to sections 60g (1), 137o UrhG, in Germany this only applies in for contracts concluded from 1 March 2018 onwards; old contracts could therefore provide for restrictions.

Contracts concluded from 1 March 2018 onwards cannot effectively restrict the reproduction of copyrighted works for the purposes of making available, indexing, cataloguing, indexing, preservation and restoration.

C.II.3.f. Contractual permission, in particular Creative Commons licences

Use for training generative AI and even more so the use of open access content as input data for AI systems may be permitted by the respective separate terms of use of the individual work. In particular, content licensed under CC-BY, CC-BY-SA or CC 0 (regardless of the version) can be assumed to be permitted for use in AI training under copyright law.[32] With a CC-BY-ND licence, on the other hand, the (automated) conversion as part of AI training is not covered by the licence; a CC-BY-NC licence is also problematic if the trained model is to be reused in return for payment. Contracts with publication service providers cannot restrict these arrangements. To the extent that use is not free (Section 23 (1) second sentence UrhG), the sources must also be indicated in accordance with the CC-BY requirements (see also Art. 53 (1) lit. c, lit. d AI-Regulation).

Contracts cannot restrict the use of copyrighted works under a CC-BY, CC-BY-SA or CC-0 licence to the detriment of the user. In the case of CC-BY-NC, CC-BY-ND and CC-BY-NC-ND, there are restrictions by virtue of the Creative Commons licence. Further restrictions are not possible in contracts with publication service providers.

C.III. Options provided for by law for TDM arrangements

However, licence agreements can contain **specific requirements** for TDM or the re-use of content. Insofar, despite Art. 3 (4) DSM Directive, no consensual definition of best practices has yet been reached between rightholders, research organisations and cultural heritage institutions in Germany, so that contractual specification can bring added legal certainty here:

Rightholders may apply measures "to **ensure** the **security and integrity of the networks and databases** in which the works or other protected subject-matter are hosted". (Art. 3 (3) DSM Directive = section 60d (6) UrhG). This applies, for example, to arrangements that address a "large number of

[31] Unlike the TDM permissions (see above) and the privileged treatment of temporary reproductions, this provision is not mandatory throughout Europe, but applies specifically under German law.

[32] https://creativecommons.org/2023/08/18/understanding-cc-licenses-and-generative-ai/.

access and download requests" (i.e. a possible server overload) (Recital 16 DSM Directive). However, such arrangements must not prevent the effective use of the TDM permission. Nor can individual users be required to conclude an individual supplementary agreement in order to be able to make use of the TDM permissions that already apply by the operation of law. For example, publication service providers can provide separate access interfaces (APIs) for TDM purposes, but arrangements restricting the use of content within the scope of TDM permissions would be invalid.

Pursuant to Art. 3 (2) DSM Directive (section 60d (5) UrhG), the institutions (or researchers) are obliged to apply **"appropriate security measures" in the event of any hosting** - retention must also be expressly limited to scientific research, including the review of scientific findings. In this way one can set guidelines for access security.

Contracts can provide for measures that ensure the security and integrity of the networks and databases through appropriate security precautions and set guidelines for the copies made in the context of TDM.

D. Recommendations for action

D.I. Contract negotiations

D.I.1 Principle

AI clauses are not mandatory in licence agreements. If a contract does not contain an AI clause, the use of licensed content is permitted within the scope of the statutory permissions (→ C.II.C.II.3). However, any uses beyond this are not possible if they are not contractually permitted.

However, licence agreements should not restrict the use of licensed content for the purposes of scientific research and teaching in connection with AI methods. A comprehensive ban on the use of AI methods cannot be accepted under any circumstances[33]. A ban on pure data analysis with AI is also unacceptable, as it is covered by the statutory permission for TDM.

D.I.2 Possible arrangements

If AI clauses are included in licence agreements, one must specify the individual acts of use to be covered thereby, as a clear-cut definition of AI is not possible (→ C.I). If certain uses are prohibited or restricted, it should also be made clear which acts of use are not covered by the AI clause and are therefore permitted. Where an AI clause is included, the interests of all parties involved must be explicitly considered in the text of the contract - this means that the interests of the scientific institution and the scientists concerned must be given appropriate consideration alongside the interests of the publication service providers.

D.I.2.a. AI clauses

In licence agreements there are provisions conceivable according to which

- non-commercial uses (including direct and indirect training of an AI tool) for research and teaching purposes as well as for purely internal purposes of the institution are expressly permitted in order to avoid legal uncertainties for institutions, researchers and teachers.

[33] See ICOLC Statement on AI in Licensing (22 March 2024), https://icolc.net/statements/icolc-statement-ai-licensing.

- the training of generative AI for commercial purposes is excluded (→ C.II.3.d).
- the licensed content may not be uploaded to third-party AI systems if this constitutes consent to the use of the content for training the provider systems (→ C.II.3.b).
- no AI model trained with the licensed content (→ C.II.3.c) and no output of an AI model that contains recognisable, protected works or parts of works may be made available to the public[34] or replace the licensing of the contractual materials[35] .
- necessary measures are applied to ensure the security and integrity of the networks and databases, such as an appropriate limit on the number of permitted downloads per minute (→ C.III). However, arrangements must be made for (alternative) procedures that enable or facilitate TDM for the institution and its researchers, such as downloading via an application programming interface (API).

D.I.2.b. Clarification

In the event that an AI clause is agreed and in particular if an attempt is made to define AI in general terms, it should be made clear that uses in the internal systems of an institution, non-public uses by authorised users for scientific research purposes and the publication of analysis results[36] are permitted without restriction. It should also be clarified that purely automated data analysis or input into browsers or standard software for word processing, translation, spreadsheets or similar purposes is also permitted (→ C.II.3.a).

D.I.2.c. Liability

In the context of AI clauses, institutions should not assume any further liability, in particular for the behaviour of their users[37] . Rather, the usual general liability rule should apply, according to which institutions are not liable for breaches of contract by their users unless the institution has caused, deliberately encouraged or tolerated them[38] .

D.I.2.d. Supplementary provisions

In addition, a general, clarifying clause is recommended, according to which any uses the organisation and the users are permitted to make under the applicable law remain unaffected.

In addition, the applicability of German law to the licence agreement should be agreed if possible. At the very least, reference should be made to the application of the lex loci protectionis principle or to the wording of Art. 8 (3) Rome II Regulation.

[34] See Baumann, Malte: Generative KI und Urheberrecht - Urheber und Anwender im Spannungsfeld, NJW 2023, 3673 (3677); Mamaar, Niklas: Urheberrechtliche Fragen beim Einsatz von generativen KI-Systemen, 481 (489).

[35] Licence agreements already contain similar clarifying clauses.

[36] See Wandtke/Bullinger/Bullinger, 6th ed. 2022, UrhG section 60d para. 29.

[37] See ICOLC Statement on AI in Licensing (22 March 2024), https://icolc.net/statements/icolc-statement-ai-licensing.

[38] See Commission on Electronic Resources in the Bavarian Library Network and the Legal Affairs Commission of the German Library Association (2015): Checklist for licence agreements, p. 35.

In the case of multi-year contracts containing AI clauses, one can also provide for a procedure to take account of technical and legal developments and to review and adapt the clauses accordingly.[39]

D.II. Contracts on open access publications

In the case of works under so-called open licences, use for training AI systems may be permitted (→ C.II.3.f).

Contractual agreements that (also) cover open access publications, such as transformation contracts, publish-and-read contracts or framework agreements, must provide for unrestricted, free access for researchers to works published under CC licences on the publication service provider's platform - preferably via an application programming interface (API) for easy technical implementation.[40] The licence should be indicated in the title metadata in machine-readable form.[41]

D.III. Uses within the scope of the statutory permission

Where a contract does not contain an AI clause, the licensed content may be used within the scope of the statutory permissions (→ C.II.C.II.3). Particularly in the case of multi-year contracts, it is therefore advisable to continuously monitor the development of case law - especially that of the European Court of Justice (ECJ) - as a change in the interpretation of the statutory permissions without a contractual provision can change the scope of the acts covered by the statutory permissions (limitation).

In contract negotiations, the statutory permissions (limitations) should be used as powerful arguments to ensure that the licence agreement does not fall short of this minimum standard. If this is not successful and, invoking a statutory permission, licensed content is used in a way that is not contractually permitted, the users are left with uncertainties (chilling effect) and risks.

D.IV. Consequences of breaches of contract

In principle, rightholders cannot invoke agreements that restrict or prohibit permitted uses under sections 60a to 60f UrhG (Section 60g (1) UrhG). Contracts concluded before 1 March 2018 (Section 137o UrhG) are excluded therefrom. If contracts concluded after 1 March 2018 restrict AI applications that are permitted under Sections 60d or 60e UrhG, the institution may instead invoke these statutory permissions.

Users can also rely on section 60d UrhG if the licence agreement is governed by the law of a country outside the EU, provided that the relevant acts are performed in Germany (Art. 8 Rome II Regulation, → C.II.1); according to Art. 7 (1) DSM Directive (Section 60g UrhG), deviating agreements are not enforceable against the organisations themselves either.[42] Claims for damages arising from a breach of

[39] See ICOLC Statement on AI in Licensing (22 March 2024), https://icolc.net/statements/icolc-statement-ai-licensing.

[40] See Bruch, C., Deinzer, G., Geschuhn, K., Hätscher, P., Hillenkötter, K., Kreß, U., Pampel, H., Schäffler, H., Stanek, U., Timm, A., Wagner, A. (2015): Positionen zur Schaffung eines wissenschaftschaftsadäquaten Open-Access-Publikationsmarktes: Positionspapier der Ad-hoc-AG Open-Access-Gold im Rahmen der Schwerpunktinitiative "Digitale Information" der Allianz der deutschen Wissenschaftsorganisationen. Ad-hoc-AG Open Access Gold in the priority initiative "Digital Information" of the Alliance of German Science Organisations, p. 27, https://doi.org/10.2312/allianzoa.008.

[41] See above. p. 25 f.

[42] See also Brehm, Elke: Guidelines on text and data mining for research purposes in Germany (2022), https://doi.org/10.34657/9388.

contract based on acts of use of the institution which are permitted by law do insofar not materialise either.[43] A contractual claim for damages should only come into consideration insofar as the institution does not pass on contractual restrictions to its users.[44]

Regardless of whose legal position would ultimately prevail in a legal dispute, the institution may find itself subject to a temporary blocking of access for individual users or for the entire institution if the use exceeds the extent expressly permitted by the contract. Blocking is justified in any case where the publication service provider's appropriate measures to protect the security and integrity of its networks and databases (→ C.III) have been violated.

D.V. Consequences of copyright infringements

If the organisation actually exceeds the permitted scope of use under sections 60d or 60e UrhG through its own services and infringes copyrights, a cease and desist letter (section 97a UrhG) and claims for elimination of the infringement, for the destruction of unlawful reproductions and, in the case of intent or negligence, for damages (Sections 97, 98 UrhG) may come into consideration. Arrangements may be made by contract providing for the extent to which the organisation is also liable for infringements by its users.

Moreover, the German Copyright Act contains provisions on penalties and fines (sections 106 et seqq. UrhG).

D.VI. Liability of the user

The users are not parties to a licence agreement between the institution and the publication service provider. The lex loci protectionis principle therefore applies to them. They can therefore base their uses on the statutory permissions, provided that all acts of use take place in Germany or in another EU Member State (→ C.II.1). However, users are liable if they infringe copyrights; in institutions organised under public law, however, the employer is primarily liable in relation to third parties (Art. 34 GG, Section 7 TV-L), who can only take recourse in the event of intent or gross negligence.

[43] See above p. 97.

[44] See above. p. 98.

E. Brief summary

- A blanket reference in the contract to the term "artificial intelligence" is unsuitable due to its broad scope as described above; instead, on should specify concrete user acts s that are to be permitted or prohibited by the contract. (→ C.I)
- According to the so-called lex loci protectionis principle, German copyright law is mandatorily applicable to the use of copyrighted works if the relevant acts (e.g. reproduction, distribution) take place in Germany. (→ C.II.1)
- Contracts cannot effectively prohibit end users from using texts or images made available from databases as input data for AI for adaptation and transformation. (→ C.II.3.a)
- Only temporary reproduction is always permitted, provided that 1. it is only transient or incidental and 2. it has no commercial significance of its own and 3. it is necessary for a subsequent lawful use of the work (in particular an adaptation or transformation). (→ C.II.3.b)
- Contracts concluded from 1 March 2018 onwards cannot effectively restrict the use of copyrighted works by TDM for scientific purposes, including the creation of internal scientific AI systems. Making reproductions available to third parties without contractual permission is only permitted under the conditions of section 60d (4) UrhG. (→ C.II.3.c)
- Contracts can effectively make provision for the use of copyrighted works for TDM for non-scientific purposes. (→ C.II.3.d)
- Contracts concluded from 1 March 2018 onwards cannot effectively restrict the reproduction of copyrighted works for the purposes of making available, indexing, cataloguing, subject indexing, preservation and restoration. (→ C.II.3.e)
- Contracts cannot restrict the use of copyrighted works under a CC-BY, CC-BY-SA or CC-0 licence to the detriment of the user. In the case of CC-BY-NC, CC-BY-ND and CC-BY-NC-ND, there are restrictions by virtue of the Creative Commons licence. Further restrictions are not possible in contracts with publication service providers. (→ C.II.3.f)
- Contracts can provide for measures that ensure the security and integrity of the networks and databases through appropriate security precautions and set guidelines for the copies made in the context of TDM. (→ C.III)

E.I. Overview: Acts of reproduction of licensed works

Act	Legal assessment
Use as input data	mandatory permission, section 44a UrhG; also: section 23 UrhG
Use as training data	Generally subject to permission, section 16 UrhG
As part of TDM for scientific purposes including AI processing activities with non-public use	mandatory permission, section 60d UrhG But: restriction for retrieval permitted for integrity and security of the database and specifications for hosting
As part of TDM for scientific purposes including AI processing activities with public use (incl. open source and free of charge)	permitted, section 44b UrhG But: reservation possible (machine-readable)
As part of TDM for other purposes	permitted, section 44b UrhG But: reservation possible (machine-readable)
Machine-assisted cataloguing and/or indexing (e.g. subject/formal indexing)	Mandatory permission in favour of cultural heritage institutions, section 60e (1) UrhG
Use in favour of the blind	Mandatory permission in favour of users sections 45a, 45b UrhG

E.II. Requirements for AI clauses

- No comprehensive or blanket prohibition of the use of AI methods (→ D.I.2.a).
- No prohibition of pure data analysis with AI methods (→ D.I.2.a).
- Clarification that uses in the internal systems of an institution, non-public uses by authorised users for scientific research purposes and the publication of analysis results are permitted without restriction (→ D.I.2.b).
- No special liability rule: Organisations should not be liable for breaches of contract by their users unless the organisation has caused, deliberately encouraged or tolerated them (→ D.I.2.c).

E.III. Requirements for supplementary arrangements

- The uses permitted to the institution and the users under applicable law remain unaffected (including the direct and indirect training of AI tools) (→ D.I.2.d).
- Choice of law in favour of German law for the licence agreement or reference to the application of the lex loci protectionis principle or reference to the wording of Art. 8 (3) Rome II Regulation (→ D.I.2.d).