



Guidance on copyright restrictions on the use of chatbots in educational institutions

Need for increased attention

There is an urgent need to raise awareness of the copyright restrictions on the use of chatbots, such as ChatGPT, by institutions, teachers and students at Danish educational institutions.

The Rights Alliance has found that copyrighted material in the form of textbooks are being uploaded to generative AI tools, such as ChatGPT, without the necessary prior permission from the publisher who has published the books.

This guide describes the copyright restrictions on the use of chatbots in educational institutions. The guidance is based on the Rights Alliance's practical experience in protecting copyrighted material from illegal use with the help of AI tools. At the same time, the guide is based on principled positions on the current regulation in the area.

First, we answer frequently asked questions that may arise in connection with the use of textbooks and chatbots. Then we describe the relevant copyright legislation and how it regulates the use of protected content in connection with chatbots.

Please contact [the Rights Alliance](#) if you are unsure whether a given use of textbooks in connection with chatbots is legal or not.

This guide does not describe the rules for training generative AI models. Read the Rights Alliance's guidance on training generative AI [here](#).

Q&A om anvendelse af lærebøger mv. i forbindelse med chatbots

RettighedsAlliancen giver her svar på spørgsmål, der kan opstå i forbindelse med brug af lærebøger og andet ophavsretligt beskyttet materiale og chatbots:

Can I upload textbooks or textbook excerpts to ChatGPT or other chatbots?

Answer: No, it is not allowed to upload a textbook or excerpts from it in ChatGPT or other chatbots, unless separate permission has been obtained from the publisher of the textbook.

Can I upload textbooks or excerpts of textbooks to a customized chatbot such as OpenAI's "GPTs" and then share the chatbot with others?

Answer: No, it is not allowed to upload a textbook or textbook excerpts to a customized chatbot and share it with others, unless separate permission has been obtained from the publisher of the textbook.

Can I upload textbooks or excerpts of textbooks to a customized chatbot such as OpenAI's "GPTs" that are not shared with others and only used for my own use?

Answer: No, it is not allowed to upload a textbook or textbook excerpts to a customized chatbot, even if it is not shared with others and only used for private use, unless separate permission has been obtained from the publisher of the textbook.

Is it allowed to upload a textbook or excerpt to a generative AI tool for the purpose of generating an audiobook version of the textbook?

Answer: No, it is not allowed to upload a textbook or excerpt thereof to an AI tool to generate an audiobook version of a textbook, unless separate permission has been obtained from the publisher of the textbook.

Do the agreements with Copydan Text & Node cover the use of textbooks in relation to ChatGPT or other generative AI tools?

Answer: No, the agreements with Copydan Text & Node do not currently cover the use of textbooks in relation to ChatGPT or other generative AI tools.

How copyright law regulates the use of protected content in chatbots

Copyright exclusivity and exceptions

The fundamental principle of copyright law is that the rights holders of copyrighted content, such as a publisher, must give permission for their content, such as a textbook, to be used in various contexts, including making copies (i.e. copying) and making it available to the public (i.e. making it available to others, for example online). This is also known as the exclusive right, and it follows from section 2 of the Copyright Act.

There are a number of exceptions to the exclusive right that may be relevant for the use of works in connection with chatbots and teaching, including the text and data mining exceptions in section 11 c and b, the private use exception in section 12 of the Copyright Act and section 13 a on display in teaching activities. Read more about this from page 5.

How do chatbots work?

With the advent of ChatGPT and similar tools, new ways have emerged for the misuse of protected works, including textbooks. Users can upload digital copies of textbooks to ChatGPT and other chatbots, after which the works are processed, enabling the chatbot to answer questions about the content or generate text that is similar or identical to the uploaded material in the form of books and articles, among other things. As described in the section below, this is illegal unless explicit permission has been granted.

It has also become possible and easy with a number of AI services to fine-tune chatbots to answer questions in a specific way and based on a limited knowledge base consisting of text documents, such as books or articles. The latter functionality is copyright relevant as the creator of the customized chatbot must upload digital copies of the text documents to the AI service provider.

The technology that enables ChatGPT to answer questions or generate text based on uploaded books, articles, images, etc. is the same technology that enables customized chatbots. The technique is called "RAG" (Retrieval Augmented Generation).

The purpose of RAG is to enable language models to use material that the model is not pre-trained on. When a ChatGPT user or a creator of a GPT uploads a digital copy of, for example, a textbook to ChatGPT or a customized chatbot, the text in the book is broken up into smaller pieces and embeddings are created, which are mathematical representations of the text.

When users subsequently interact with their version of ChatGPT or the custom chatbot, ChatGPT/the chatbot first searches for relevant text in the RAG database of text and embeddings of the previously

uploaded material. If relevant text is found in the database, this text is inserted in identical or summarized form into the ChatGPT/chatbot output. Along with the text from the RAG database, the output also consists of text generated from the chatbot's (i.e. language model's) pre-training data.

Currently, OpenAI's GPT service is the most popular way to create a customized chatbot. There are many AI services based on OpenAI's models, such as Chatbase.co.

OpenAI allows users to publish their GPTs on a "GPT Store" where other users can access the customized GPTs. With other AI services, such as Chatbase.co, it is possible to publish your chatbot on a website or via an online platform, such as Instagram.

The Rights Alliance has tested chatbots using RAG databases and found that it is possible to recreate parts of and entire works that have been uploaded to a chatbot using RAG technology.

What copyright relevant actions occur when uploading works to ChatGPT and customized chatbots?

When uploading digital copies of literary works to ChatGPT etc. or OpenAI's GPT service etc. using the RAG technique, digital copies of the works are created. Copy creation generally requires permission from those who own the copyrights to the work.

The copies have also been shared with a third party, such as OpenAI, as the work copies are now stored in a database on OpenAI's servers. This is an exploitation of the work that generally also requires the permission of the rights holders.

There are subscription forms at OpenAI, for example, where the works are not uploaded directly to OpenAI, but instead stored in a vector database at the uploader. However, when the chatbot takes text from the vector database and inserts it into an output, this text is also shared with OpenAI or the third party that runs the AI model. This means that in this context, copy is also shared with a third party, which requires permission from the rights holders.

If the GPT is simultaneously published on the GPT Store or in other ways, for example on online platforms, the works uploaded to the GPT are made available to the general public. This is because the GPT has access to a database with copies of the works and because it is possible for users of the GPT to recreate the works. Such provision requires permission from the copyright holder of the uploaded works, for example a publisher if a textbook is uploaded.

There is also a risk that users of GPTs will generate copies of all or parts of uploaded works to the GPT. The same applies to users of ChatGPT, where users have uploaded a copy of a protected work. This copying also requires permission from the copyright holder.

Can the text and data mining exceptions apply when uploading content to ChatGPT or creating a custom chatbot?

In June 2023, exceptions to the exclusive right were introduced for some specific situations with the rules on text and data mining (TDM) in sections 11 b (general access) and 11 c (special access for research) of the Copyright Act. The rules make it possible in specific cases to make copies of protected works if it is done for the purpose of TDM. TDM is described in the provision as an automated, analytical process that aims to analyze text and data in digital form in order to generate information, including patterns, trends and correlations.

The concept of TDM has not yet been addressed and delineated by the courts. We assess that uploading copies of works to chatbots goes beyond the copying that may be permitted by the TDM exception in some specific situations. This is partly due to the history surrounding the creation of the TDM exception, where generative AI and the disruptive and socio-economic consequences of the technology were in no way discussed. Partly because the use of works uploaded to chatbots goes beyond the mere analysis of text and instead copies copyrighted elements of the work.

We consider the use of the TDM exception for uploading works to chatbots to be inconsistent with basic copyright principles, including the three-step test derived from the Berne Convention. Under the three-step test, exceptions to copyright and related rights may only be used in certain special cases that do not conflict with the normal exploitation of the work or other subject matter and do not unreasonably prejudice the legitimate interests of the rightholder. If the TDM exception allows works to be copied for use in chatbots, this would be contrary to the normal exploitation of the work and would unreasonably prejudice the legitimate interests of the rightholder.

Even if the actual copying of the work by uploading to ChatGPT or GPTs is found to be covered by the TDM concept, an illegal use of the work is subsequently made.

Firstly, we have learned that OpenAI will generally use copies of uploaded works and/or dialogs with ChatGPT/GPTs for training their AI models. The rights holders of the works have not given permission for this, and OpenAI cannot use the TDM exceptions as they have not collected the works themselves and had legal access to them. The user who uploads works to ChatGPT/GPTs is contributing to this copyright infringement.

Secondly, it would be an illegal act to publish a GPTs/chatbot to which protected works have been uploaded, as it would also be an act of making the work itself available, cf. the possibility of recreating text from the work with GPTs.

Copying under the TDM exception also requires that the copying is based on a copy where there has been legal access and a legal model, cf. section 11(3) of the Copyright Act. This means that you must have obtained access to a legal copy of a work before you can potentially make further copies for use in TDM. This is related to the following section on the private use exception.

Can the private use exception apply when uploading content to ChatGPT or creating a custom chatbot?

Section 12, subsection 2, no. 5 of the Danish Copyright Act makes it legal in special situations to produce individual copies of literary works in digital form. It is a requirement that the copies are made solely for the personal use of the creator or their household. Personal use does not cover professional or educational purposes.

Even if the copies are made for personal use, it is not permitted to copy borrowed or rented copies of the work (paragraph 3). It is also not permitted to use external assistance for copying literary works if the assistance is for professional purposes (subsection 3, no. 3). It is also not permitted to use technical equipment made available for commercial purposes in the case of literary works (paragraph 5, 2nd indent). Outside assistance refers to people to whom you are not connected by personal ties, e.g. you can get help from a friend or close colleague.

Copying under the private use exception also requires that the copying is based on a legal model, cf. section 11, subsection 3 of the Copyright Act.

Our assessment is that the rules on private copying cannot be used to legalize the uploading of textbooks to ChatGPT. This applies to both teachers and students, as the private use exception only applies to "personal use". According to Rosenmeier et al., *Immaterielret* (2024) page 221, the actions students take during education are not "personal". Therefore, there must be a Copydan Tekst & Node agreement with the educational institution before copying by students is permitted. There is currently no Copydan Tekst & Node agreement covering copying in connection with uploading to chatbots.

Teachers or educational institutions cannot upload textbooks to chatbots either, as they will be making copies in a work-related context that is not covered by "personal use".

In addition, the relationship would not be covered by the exception because the maker loses control of the copy when it is sent to OpenAI or another AI provider.

An additional factor is that OpenAI or another AI provider is involved in the making of the copy and makes its technical equipment available for commercial purposes. This applies even to the free version of ChatGPT, as OpenAI has a commercial interest in users using their AI tools. Der vil også være mange situationer, hvor kopieringen sker uden et lovligt forlæg eller på baggrund af et lånt eller en lejet kopi, hvorfor undtagelsen ikke kan finde anvendelse.

Can the educational use exception apply when uploading content to ChatGPT or in connection with the creation of a customized chatbot?

With the amendment of the Copyright Act in 2023, a new provision on the use of copyright material for educational purposes was introduced, cf. section 13 a. If there are no suitable licenses in an area,

copies may be made and made available to the general public for purposes that serve exclusively for educational purposes, provided that it is not for commercial purposes. The provision has the character of a compulsory license provision, and it only applies to the extent that appropriate licenses are not available. In Denmark, there are collective contractual licenses that enable the use of books for educational purposes. Section 13a does not apply outside of these collective contractual licenses, even if they only enable a certain type of use. For example, for the use of copying a certain amount of pages or percentages of a book. "Demonstration" under section 13a would be situations where the use supports, enriches or supplements teaching.

It follows from section 13a, subsection 2 that it is a prerequisite that the use is made under the responsibility of the educational institution on its premises or elsewhere or via a secure electronic environment accessible only to the pupils, students and teaching staff of the educational institution.

/ RettighedsAlliancen