ChatGPT, DeepL and DALL-E 2: what do the T&Cs say about IP ownership, data retention and infringement?

The speed of AI tools has rendered them valuable assistants to many employees, with or without their employer's knowledge. At first glance, this appears to be a win-win situation; the employee completes more tasks in less time, allowing the employer to increase the employee's workload and, subsequently, productivity. However, the rise of AI tools' popularity comes with risks, from inaccuracy to socio-economic and legal issues.

This article examines the potential benefits and IP risks of three AI tools that are commonly used in business:

- ChatGPT;
- DALL-E 2; and
- DeepL Translator.

Scrutinising the terms and conditions of the free versions of these tools should help users and their employers navigate the IP legal implications relating to:

- ownership of content;
- data retention; and
- IP infringement.

ChatGPT and DALL-E

ChatGPT is the product of AI research laboratory OpenAI, which has commercial and corporate ties to Microsoft. ChatGPT is based on the GPT-4 large language model. Accordingly, a user can provide a text, audio or visual prompt to the model, and request a specific result. The model then offers the best probabilistic output based on its training.

Although ChatGPT can produce visual outputs, its equivalent for image generation is DALLE. DALL-E is also owned by OpenAI, and is based on the same model as ChatGPT.

It is worth noting that OpenAI's generic terms of use differ slightly from the new Europe terms of use. The latter entered into force on 15 February 2024 and apply to users residing in the European Economic Area (EEA), Switzerland and the United Kingdom. The most noteworthy change for the free versions focuses on a more user-friendly approach to OpenAI's provision of services and the potential liability in case of damage (see "Our Commitments to You"). In addition, consumers based in the EEA are offered a legal guarantee on the OpenAI services.

Content ownership

According to OpenAI's generic terms of use, the user's input (ie, whatever information the user gives to the model as a prompt) is owned by the user. Additionally, OpenAI "assign[s]" the user – which agrees and complies with the terms – all its rights on the output. The output is the 'answer' given by the model that corresponds to the given input. Such assignment implies that OpenAI is the original owner of the output (ie, the copyright holder of the AI-generated reply).

Whether OpenAI (or any other entity using the model in its platform) initially owns the generated output is at best dubious, as there is uncertainty around the copyrightability of AI-generated material. Multiple jurisdictions (eg, the United States) have established human authorship as a prerequisite of copyrightability. Therefore, an entirely AI-generated work might not fulfil this condition. On the other hand, it could be argued that the user is the owner of the content they request from AI programs. The reason is that said program is merely a tool used by a human to create an original work that expresses their creative choices. However, copyrightability and authorship are delicate legal issues that are presently the subject of heated debate, and might not impact the average user to a great extent.

Especially for DALL-E, OpenAI claims that it can create original images and works of art based on the user's input, the rights of which OpenAI then assigns to the user. This legal scheme currently seems to lack legal justification. In any case, it does not ostensibly affect the user's legal rights and obligations.

It should be also noted that since September 2022, OpenAI has had a content policy in place for DALL-E that advises users not to conceal that the image has been created by AI, rather than a human.

Data retention

As in many AI tools, ChatGPT requires the retention and use of the data made available in the ChatGPT website or mobile applications. In reality, the data that users provide in their prompt are collected, maintained for a certain period and used by the AI developer to train the AI. As a result, in case the prompt includes confidential information or data protected by IP rights (eg, copyright and trade secrets), these will become part of the data set that trains the model. This may compromise the employer's IP rights and harm their business interests.

In the case of DALL-E, the image used in the prompt will be retained for training if the user requests the AI model to:

- expand an existing image (inpainting);
- edit an existing image by adding or removing elements (outpainting); or
- create a variation of an existing image.

This might have unwanted consequences (eg, leaking proprietary images (company data)) in case an employee uploads an image under its employer's copyright or trademark.

On the bright side, OpenAI allows users to opt out from the use of input to improve the AI model. Users in the EEA, Switzerland and the United Kingdom can easily opt out by changing their account settings. Users in the rest of the world must either:

- opt out through OpenAI's privacy portal; or
- switch off the training option during ChatGPT conversations.

IP infringement

Even though the nature of generative AI is to 'create' a relatively autonomous 'best possible' output that derives from the training data set, there might be instances of potential IP violation. This might be the case if part of the output is identical or – most probably – similar to part of the input.

Whether mere similarity is sufficient to establish copyright infringement depends on the applicable law of each country. However, under several national copyright legislations (eg, in the United States), an output identified as being similar to a work included in the training data set might be considered infringing.

Despite this difficulty in establishing a link between 'similar' input and output, IP infringement might be easier to detect when the AI-generated output is an image. OpenAI expects that trademark or copyright-protected content might be generated by the AI model in the output. In addition, copyright infringement disputes have commenced in different jurisdictions in the context of image generation through AI tools using IP-protected content as input. Even though these legal actions are not addressed towards AI users, users could still be asked to indemnify.

Regardless of whether infringement in the output is rare, the business risk of using potentially infringing output might be considered too severe to ignore. Moreover, it is important to examine the relevant clauses – if any – regarding warranty, indemnification and liability in the terms of service of AI tools.

In its generic terms of use, OpenAI has disclaimed any warranty of – among other things – non-infringement for both ChatGPT and DALL-E. This means that the output generated by the AI model may infringe third-party IP rights and that OpenAI should not be held liable for any damages caused by the use of its services. The user is also obligated to defend OpenAI in case of legal actions against the latter (eg, in case of breach of contract or IP infringement) and indemnify OpenAI for any expenses, caused by the use of the services.

The above might occur due to breach of contract (in case of violation of the terms of use) or, for example, IP infringement when using the output.

However, this is not the case for the users in the EEA, Switzerland and the United Kingdom where stricter consumer protection laws kick in.

In these countries, according to the terms of use, some AI developer-friendly provisions (eg, the "disclaimer of warranty", "limitation of liability" and "indemnity") have been removed from the terms of use. In addition, OpenAI commits to provide its services with "reasonable skill and care". Lastly, regarding dispute resolution, both OpenAI and the user can bring an action before national courts.

Considering the above, it might appear that users in these countries have an advantage while using ChatGPT and DALL-E. Still, the provisions include vague legal terms (eg, "reasonable skill and care") open to interpretation by the national courts.

For this reason, all users – irrespective of their location – should exercise caution when deploying the AI model, especially in the course of their employment.

DeepL Translator

With 1 billion users, DeepL is one of the most popular generative AI services worldwide.

DeepL is a machine translation service utilising neural networks or improving writing. Its translator service is the most commonly used – its relatively high accuracy and usefulness in a global economy making it popular with a range of users (eg, students and employees).

Content ownership

DeepL's [terms and conditions](#) make no specific statement regarding ownership of the input. Presumably, the text submitted for translation might be either the user's own creation (ie, the user is the copyright owner) or the work of a third party.

The copyrightability of the output (ie, the translated text) is contestable. In general, [translations might be copyrightable](#) if they fulfil the copyrightability requirements (ie, are an original work of authorship). This might be the case for the translation of a poem or a novel – with the original work having a high modicum of creativity – but it is unlikely to apply for machine-translated texts.

Data retention

DeepL's terms and conditions expressly state that the submitted texts, the translation and any correction provided by users will be used to improve the translation algorithm. DeepL does not provide the possibility for users to opt-out from the use of their data to improve the service. However, users are advised not to submit any personal data for translation to the free DeepL Translation tool, so as to avoid its retention in the company's servers.

If personal data must be translated, the user should submit it to DeepL Pro, which requires a paid account. This suggestion is reiterated in both the terms and conditions and the [privacy policy](#).

Since there is no possibility to opt out of using data to improve the service, an employee using DeepL is at risk of giving away sensitive information and IP-protected data for training purposes.

IP infringement

Unlike OpenAI, DeepL SE, the company behind DeepL, does not include extensive, multi-layered provisions in its terms regarding potential violations, including breach of contract or IP infringement.

The only relevant clause concerns exclusion of liability regarding the accuracy of the AI-generated outcomes and the availability of the services.

At first glance, it is hard to construe the word 'accuracy' to include possible copyright infringement claims. There are two points that might explain this. First is the nature of the expected use of this DeepL version (ie, translation of texts for use in the course of daily life, without a high risk of copyright infringement). This means when the text inserted for translation is the original text of the AI user and does not contain other third-party copyright material. Second is the fact that DeepL SE explicitly [forgoes any right](#) that it may have or is considered to have on the output for providing this service.

The fact that it may be unlikely for copyright infringement claims to be raised against the translated content does not mean that the user has a 'free pass'. Particularly in a professional context, employees should exercise caution regarding the information that they provide as an input to the platform, as well as how they will utilise the output.

For the input, the data included in the prompt could be used by DeepL SE – as stated above – in light of the lack of an opt-out mechanism.

Regarding the output, there is a low chance that a translation of a copyright-protected work not authorised by the copyright owner might be infringing. This is because only the original author can give permission for such transformation of the work.

	Input ownership by User	Output ownership by user	Data retention	Platform liability for IPR infringement
ChatGPT /DALL-E	X	(Assignment to user)	X	X (only for Europe Terms of Use and only under specific circumstances)
DeepL	X	X	X	(unclear)

Adopt a firm position

Although actions are completed faster and work processes appear to be streamlined, the potential dangers of generative AI tools for businesses should not be taken lightly.

Companies should adopt a firm position towards the use of generative AI tools for employees' working activities. In some cases, this may mean banning the use of such tools, while in other cases, this may involve establishing guidelines to mitigate potential exposure to risks. For example, the European Commission advised its staff to tread carefully when using AI tools by not sharing non-public, sensitive information and being aware of IP infringement.

Nevertheless, given the relative novelty of AI tools in the workplace and their unforeseeable impact, we may benefit from learning how to live with them. Therefore, even though risk-free use is unattainable, the education of employees on how to use AI tools safely – to the greatest extend possible – is crucial.

The views expressed in this article are those of the author and do not necessarily reflect the opinions of Ericsson

Eleftheria Stefanaki
Senior IP rights and open-source researcher
Ericsson