IMPORTANT COPYRIGHT CONSIDERATIONS RELATED TO

TEXT-TO-X GENERATORS

AI models that create images, music, videos or other content based on a few keywords or a short description are gaining momentum. Although these AI tools are increasingly being used, there are still many unanswered legal questions including many issues under copyright law.

In this brAInfood, the Knowledge Centre Data & Society provides a brief introduction to text-to-x generators and explains some copyright considerations related to the training of generative AI systems, the role of human contribution in the creative process and the output.

AI FLANDERS
BUILDING OUR DIGITAL FUTURE

This brainfood was prepared in consultation with Prof. J. Vanherpe and I. Emanuilov, IP scholars at CiTiP (KU Leuven)

Knowledge Centre Data & Society (2022). Important copyright considerations related to text-to-x generators. Brussels.

What are text-to-x generators?

Text-to-x generators fall under the broad label of generative AI or foundation models. In essence, these generators are able to produce images, music, video or code based on text input (key words or descriptive phrases, so-called "prompts"). This technology has seen some significant breakthroughs during the last few years which led to the release of powerful generators (e.g. Stable Diffusion, AIVA, Dall E, MuseNet, Midjourney or CoPilot).
The software relies on complex machine learning models that are trained on unlabeled data (incl. images, video, audio,…). So far, generative AI innovation has been driven both by Big Tech-supported consortia as well as by start-ups.
These AI models create a lot of opportunities for companies or professionals working within the content creation field, but will also challenge their business models.

when training an AI-system

When training an AI-system, existing materials or content are used. This requires:

Permission. This means entering into a contractual license (incl. open licenses: creative commons, copyleft,…).

Applicable exception. In Belgium/ EU there is a long but closed list of exceptions. Relevant exceptions may be:
- Educational and scientific purposes (but no subsequent commercialization of output)
- Text- and Data Mining (TDM) but limited scope of application:
  - (i) Specific exception for research organisations and cultural heritage institutions; or
  - (ii) General exception but this only applies insofar the use of works has not been expressly reserved by their rightholders (e.g. through metadata or Terms of Use)
- Temporary acts of reproduction which are transient or incidental and an integral and essential part of a technological process.

However, these exceptions are interpreted restrictively, so relevance may be limited

Public domain. This means that the materials are not/no longer protected under copyright (i.e. after 70 years from the death of the author). Here, however, there is a risk of outdated/ unrepresentative material.
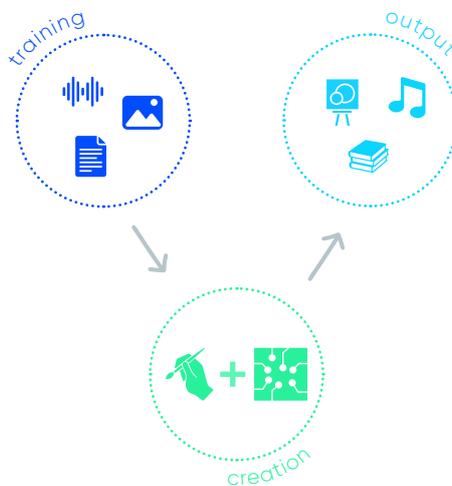
If absent: infringement of the copyright of the authors of the content featured in the training data.

during the creative process

During the creative process the role of the human is important: human intervention needs to contribute decisively to the originality of the output (see next step).

There are three relevant phases:
- **Conception/preparation**: e.g. design of AI system, initial setting of parameters or weights, additional training (data),
- **Execution:** e.g. modifying parameters or weights, used prompts,…
- **Redaction of output:** e.g. selection of output and post-production editing/modification,…

In order to avoid discussion: relevant human activities should be logged or recorded.

regarding the output of an AI-system

Regarding the output generated or assisted by an AI-system the following issues may arise:

Infringement. If original elements of an existing work can be found in the AI-output, it is an illegal copy. If the AI-output is clearly inspired by and resembles existing content, it is an illegal adaptation. A style or genre is, however, not protected under copyright.

If there is no legal ground (e.g. permission): both situations result in copyright infringement.

Liability lays with the actors who contributed to the creation or divulgation of infringing content.

Eligibility for copyright protection. In order to enjoy copyright protection, a work needs to be original. This means:
- Free and creative choices
- Intellectual creation
- Reflect personality of author.

Consequently, entirely AI-generated output is unprotected (by lack of meaningful human intervention), while AI-assisted output is protected if a decisive human contribution is present (see creative process).

Property rights in output. Many parties may stake a claim in AI-assisted output: AI-developer/-producer, AI-operator, owner,… but these claims are fact-specific and depend on (the extent of) the actual involvement of actors.

Due to legal uncertainty contractual licenses, such as DALL-E license*, are being used.