INFORMATION, COMMUNICATION & COMPUTING

Fields of Expertise TU Graz

Source: istockphoto.com

Kay Uwe Römer
Source: Lunghammer – TU Graz

Following up on my previous editorial (TU Graz research #34 02/2025) on the implications of the AI Act for researchers, we meanwhile arranged a training course on the topic held by AI Factory Austria.

Interestingly, not the AI Act itself but copyright law in the context of AI had the most surprising ramifications. According to Austrian copyright law the author of the content holds the copyright for this content until 70 years after their death, including the exclusive right of reproduction, distribution, public performance and the right of publishing the authored content online. This also applies to copyrighted material published on the internet that is used for training large language models (LLMs), such as ChatGPT. If a user of ChatGPT later receives (a modified) copy of that copyrighted training data in response to a prompt, the copyright of the original author may still apply to the LLM output. If the LLM output is not a fair copy of copyrighted material used for training, but rather a “creative” variant or mix of different copyrighted training data that is sufficiently different from any input training data, then the output may devolve to the public domain as an AI cannot hold a copyright of its own output. Only if the user of the LLM made creative choices about the final form of their input (prompt) or shaped the output sufficiently, may they hold a copyright of the LLM output.

This observation may pose serious problems for researchers, specifically if they use LLMs to create output (e.g., program code) that is later used in a dissertation or in contract research. In a dissertation (doctoral thesis), the author of the dissertation, if they have used an LLM, may not actually hold the copyright of the research output published in the dissertation. In contract research, the copyright of the research output is often transferred from the university to the company paying for the research, but if an LLM was used to create the output, the university may not actually hold the copyright in the transferred output. Here, problems will occur if an author of copyrighted material finds out that an LLM has been trained using their copyrighted material and that material that was used for a dissertation or product somebody else claims to hold the copyright of was reproduced in LLM output. In such cases, the original author may sue for copyright infringement. A company may get suspicious that the research output it got from the university is actually in the public domain and sue the university for violation of contract. In any case it is highly recommended to carefully document every step taken and prompt used when creating material with an LLM if it is to be used in a dissertation, contract research or similar.

In this edition of TU Graz research, Matthias Neumann, of the Institute of Statistics at the Faculty of Mathematics, Physics and Geodesy, gives us some insights on his research. Enjoy reading!

Information, Communication & Computing