

Inquiry into artificial intelligence (AI) in New South Wales: Submission in relation to copyright issues

2 July 2023

The Campaign for AI Safety welcomes the opportunity to lodge a submission into the Inquiry into artificial intelligence (AI) in New South Wales. In addition to this submission, the Campaign intends to make a broader submission.

This submission is focused on the issue of **copyright** in relation to term “b” (*“the social, economic and technical opportunities, risks and challenges presented by AI to the New South Wales community, government, economy and environment”*) in the Terms of Reference.

We recognise that copyright law in Australia is a federal responsibility, but one with profound effect on businesses and individuals in New South Wales, and therefore should be considered by the inquiry.

Our policy recommendations

Generative AI arrived largely unanticipated by the public, creative professionals, publishers and relevant stakeholders. The vast majority of creators never imagined that their works would be used for machine learning. Copyright law never considered this novel use and the ease and speed with which content could be generated with AI. This has created a climate of uncertainty for NSW creative professionals as well as businesses using AI-generated content.

Therefore, we recommend that Australian parliaments clarify and reinforce the copyright regime in relation to training of AI models on copyrighted materials. Specifically:

1. Third parties who wish to use copyrighted materials in training AI models must obtain **specific consent** from the owners of materials.
2. Specific consent implies that third parties **must not be allowed to coerce copyright holders** to provide such consent. For example:
 - a. Consent must not be part of terms and conditions of unrelated services (such as video or gaming distribution platforms). This way social media companies will not be able to deny services or features of their platforms to those who do not consent to such use of copyrighted materials.
 - b. Specificity of consent means that it needs to be a separate agreement, ideally a contract with consideration, which copyright holders have an opportunity to take time to consider, review, and negotiate. It must not be a checkbox underneath a registration form on a website.

3. **Consent must be granular.** For example, copyright owners should be able to specify if they allow third parties to have technical means to generate works “in the style of” the author of the materials.
4. Copyright holders must have effective means of **negotiating and receiving compensation** for the use of their copyrighted materials (such as opt-in collective bargaining mechanisms).
5. **Data used in training AI models must be fully referenced.** References to the works used in training must be made publicly available.
6. Training models using “synthetic data” generated by other models trained on copyrighted materials must be considered equivalent to training on those copyrighted materials.
7. Legal liability in case of infringement of copyright must apply both to parties that train models and parties that use those models to generate content.
8. “Fair use” provisions in law must not apply to AI model training.
9. Materials generated completely or substantially by AI models must not be copyrightable.

These clarifications should be legislated, without waiting for courts to provide them. Appropriate legislation, if promptly enacted, will give certainty in relation to these important questions to the AI industry, the NSW business community, and copyright holders, including artists, writers, publishers, and the wider public.

We clarify the terms used in the recommendations above:

- **AI models** refer to generative AI models and systems (including constellations of models), including large language models (e.g. OpenAI’s GPT-3), multimodal models (e.g. Google Deepmind Gemini), large diffusion models (e.g. Stable Diffusion).
- **Training** includes fine-tuning.
- **Copyright** includes other substantially similar intellectual property rights.
- **Copyrighted materials** include confidential information and similarly protected content.

Case study 1: Chat-GPT powered by GPT

Chat-GPT is a famous AI-powered chatbot that uses a variety of models in the GPT series (including GPT-3.5 and GPT-4) to generate answers (or “completions”) to user inputs (“prompts”). GPT models were trained on various undisclosed sources of information.

It has been alleged that many of those sources of data are copyrighted or privacy-protected¹. Thanks to the training data, Chat-GPT is able to reproduce (albeit with some error) portions of copyrighted works and summarise them². Authors face diminished demand for their works when imperfect reproductions and regurgitations are accessible from OpenAI and Microsoft.

Because of the secrecy around the training data used by OpenAI³, copyright owners might not realise that their rights could be violated.

AI companies like OpenAI stand to unfairly gain from the use of works, whose creators never considered the possibility that they would be used for machine learning, do not consent to such use, are not compensated, and do not enjoy attribution of moral rights.

Case study 2: Google Deepmind’s Gemini

Google Deepmind is working on a system called Gemini, which they claim will have a number of advanced capabilities⁴. It has been reported⁵ that Gemini may be trained using materials uploaded onto Google-owned YouTube platform.

While it is not clear what data is being used for training, if Google indeed is using videos hosted on the YouTube platform, that would be a violation of users’ trust and needs to be clarified in legislation as illegal because no specific consent was sought from the users who created and uploaded their videos, and no mechanism of compensation was agreed to or established for such use.

¹ “ChatGPT Language Model Litigation”, Joseph Savier Law Firm, <https://www.saverilawfirm.com/chatgpt-language-model-litigation>.

² “OpenAI’s ChatGPT may face a copyright quagmire after ‘memorizing’ these books”, The Register, https://www.theregister.com/2023/05/03/openai_chatgpt_copyright/.

³ Only cryptic names for datasets are given in OpenAI’s GPT-3 paper, such as “Books1”, “Books2”, etc. “Language Models are Few-Shot Learners”, <https://arxiv.org/abs/2005.14165>.

⁴ “Google DeepMind’s CEO Says Its Next Algorithm Will Eclipse ChatGPT”, Wired, <https://www.wired.com/story/google-deepmind-demis-hassabis-chatgpt/>

⁵ “Why YouTube Could Give Google an Edge in AI”, The Information, <https://www.theinformation.com/articles/why-youtube-could-give-google-an-edge-in-ai>

Why copyright matters for AI safety

There are three reasons why enforcement of copyright as described in our recommendations matters for AI safety:

1. There is ongoing debate on the issue of “emergent abilities”⁶ of large language models, including abstract thinking and dangerous capabilities (such as ability to suggest dangerous chemical compounds). Because they are seen as unpredictable, they can be hard to prevent, detect, limit or moderate.

Alternative explanations of these abilities have been proposed, including training data size⁷. Because the lists of references to training data of the most advanced models have not been released to the public, researchers have a difficulty in resolving the important issue of to what extent, which specific abilities are emergent or are regurgitation of their training data.

Our recommendation #5 aims to rectify this situation.

2. According to the Campaign’s recent poll⁸, 66% of American adults agree that “artists and content creators should be paid if their work is used in training artificial intelligence (AI) models”, which AI companies are not currently doing at scale.

At the same time, AI labs often make reference to advancing AI “benefits all of humanity” and “avoid enabling uses of AI or AGI that harm humanity or unduly concentrate power”⁹

There is an apparent disconnect between the behaviour of AI labs and their stated intentions, which should make one suspicious of labs’ real motivations for building very powerful AI systems.

Enforcing copyright will help reduce concentration of power and resources in the hands of makers of AI systems.

3. Enforcement of copyright in regards to training data would require some time and regrouping, which can help cool down competitive race dynamics in the AI industry, which are detrimental to ensuring robustness and safety of AI¹⁰.

⁶ Ability is defined as “emergent if it is not present in smaller models but is present in larger models”, Jason Wei, et al., “Emergent Abilities of Large Language Models”, <https://arxiv.org/abs/2206.07682>

⁷ “Are Emergent Abilities of Large Language Models a Mirage?”, Rylan Schaeffer, et al., <https://arxiv.org/abs/2304.15004>

⁸ “USA AI x-risk perception tracker”, Campaign for AI Safety, <https://www.campaignforaisafety.org/usa-ai-x-risk-perception-tracker/>

⁹ OpenAI Charter, <https://openai.com/charter>

¹⁰ “AI Is Not an Arms Race”, Katja Grace, Time, <https://time.com/6283609/artificial-intelligence-race-existential-threat/>

Next steps

We recognise that copyright law in Australia is a federal responsibility. Therefore, we request that the NSW Parliament, as an urgent priority, work together with other parliaments in Australia to enact new legislation clarifying copyright in relation to training data, ensure that Australia works on similar clarifications in its international agreements with other countries, and that mechanisms of negotiation and distribution of compensation be established in Australian jurisdictions.

About the Campaign for AI Safety

The Campaign for AI Safety is a not-for-profit association with members from around the world. We are an association of people who are concerned about the dangers AI poses to humanity and advocate for a stop on the advancement of AI capabilities and stronger regulation that prioritises safe and responsible AI. We are not affiliated with any political group. For more information, and to read our policy recommendations, please visit www.campaignforaisafety.org.