COPYRIGHTAGENCY

Submission to ACCC Digital Platform Services Inquiry – September 2024 report revisiting general search services

May 2024

Introduction

This submission addresses licensing of content for AI-related developments in Australia, and why the Government should assist with its development. The Government is being asked by multinational companies that provide search services to water down Australia's copyright legislation to benefit them, at the expense of Australian society. These requests should be resisted.

This inquiry is not considering broad issues associated with Generative AI, but it is considering generative AI to the extent it relates to general search services.

Multinational companies that offer search services are continuing to develop Generative AI products and services based on the unlicensed use of other people's content. The issues paper refers to Google's Search Generative Experience (SGE) and to Microsoft's Copilot (powered by OpenAI's Chat GPT). Both Google and OpenAI have entered into some licensing arrangements for news media content, demonstrating that it is possible, but have not compensated all the other content creators whose content they have used to develop their commercial services, without consent or compensation.

In its report, the ACCC should recommend that the Government take steps to:

- insist upon and promote the use of legitimately-sourced content in connection with AI tools and services, including licensed content; and
- require transparency from AI developers, including attribution of third party content.

Local AI development

AI technologies are moving very fast, but it is helpful to look what types of developments are likely to occur in Australia.

The 'foundation' large language models (LLMs) that have been the focus of most mainstream attention, such as OpenAI and Google Gemini, have been developed in the US, where there is the sufficient venture capital and computing capabilities. They have been the focus of attention for failing to compensate creators of content that they have used to develop their commercial services, and the consequential legal actions (mostly in the US).

Local development of AI tools that use third party content are more likely to be based on smaller, curated datasets. This includes for 'fine-tuning' LLMs, and Retrieval Augmented Generation (RAG), to improve the relevance and quality of outputs.

The content for these developments can easily be legitimately sourced, including under copyright licensing arrangements. The Australian government can take steps to ensure that content is legitimately sourced.

Locally developed AI tools that use legitimately sourced content are likely to produce higher quality outputs, and will also provide collaboration opportunities for Australia's creative industries, including in the news media sector.

Transparency and attribution

For a range of reasons, AI developers must be required to disclose third party content that they have used in connection with development of an AI tool, including as part of training data for language models. Apart from providing content creators with necessary information, transparency is important for other reasons such as bias.

This should include attribution of third party content in Generative AI outputs.

Responses to arguments for watering down Australia's copyright system

Large multinational companies, including those that provide search services, are pushing for Australia's copyright legislation to be watered down to benefit them. Their proposals are not in the best interests of Australian society, and should be resisted. These same companies have demonstrated an unwillingness to pay for content that increases their profits, including news media content.

Below are responses to arguments that these companies have been making.

Claim	In fact
Australia's copyright system is a barrier to development of AI models	There is a vast range of content available for lawful use by AI developers, including under efficient and fair licensing arrangements.
	The cavalier approach to the use of other people's content by billion-dollar companies based in Silicon Valley should not be used as a blueprint, anywhere.
	The UK's recent rejection of a broad text and data mining (TDM) exception indicates that such an exception is not required for innovation in the UK, and is not compatible with support for creative industries.
Australia's copyright legislation is a 'grey area'	It is not. It sets out when someone needs permission (a licence) to use someone else's content, including in connection with AI. It also sets out the requirements for copyright protection, including where a creator is using a technological tool. AI is not a technological development that requires or warrants changes to Australia's copyright legislation.
	As the Business Council of Australia has said:
	In Australia, the development, use, and outcomes of AI are regulated by technology-neutral laws of general application [including copyright].
	...
	This remains appropriate: Australia has a mature legislative and regulatory environment that provides protections, guardrails, and avenues of recourse. This technology-neutral approach ensures regulatory focus remains – appropriately – on the harms, not the technology.[1]
AI training is like reading a book	This analogy (like similar analogies that have been made before by tech companies seeking to develop commercial products using other people's content) is inapt and inaccurate. Training for Generative AI uses the *expression* of facts and ideas in other people's content (e.g. the way that an author has written about a particular topic), which is protected by copyright.

---

[1] https://consult.industry.gov.au/supporting-responsible-ai/submission/view/483

Claim	In fact
	Generative AI tools are designed to produce new content that expresses facts and information (e.g. in writing) in ways that mimic human-created content. The quality of the outputs (e.g. well-written text) is increased by the quality of the material that the models are 'trained' on. Language models are dependent upon well-written content (by human professional writers) to 'learn' how to write well.
Australia is out of step with international developments	Australia's copyright system is something to be proud of. It includes world-leading provisions, such as those enabling efficient and fairly compensated use of content in the education and government sectors. Its importance to Australia's creative industries was recently reiterated in the National Cultural Policy. In the UK, the Government has recently rejected calls by the tech sector for a broad 'text and data mining' exception, and has made clear that AI development should not be subsidised by the creative industries. In Japan, the Government is considering scaling back the very broad exception for AI, given concerns for the creative industries.[2] In the US, billion-dollar companies used other people's content without their permission or payment to develop commercial products, arguing that their actions were allowed in the US by the notoriously uncertain 'fair use' provision. It is by no means clear that the exception does apply, as indicated by more than 24 court cases in train, and AI developers are now looking to licensing.
Copyright isn't the right model for creators' interests	Copyright is one mechanism for creators to be properly compensated, where AI developers are operating ethically and fairly to acquire lawfully-sourced content. There are others, which may apply in other circumstances. A mechanism similar to the news media bargaining code could apply, but it is not a replacement for creators' copyright rights.
Licensing is not viable	Licensing for AI is already happening. For example, there have now been several deals between OpenAI and major news media companies. Scholarly publisher Wiley recently announced a $23m deal for licensing content for training by a large AI developer. Licensing of large volumes of content, including via not-for-profit copyright management organisations, has been a feature of Australia's copyright landscape for more than 100 years. There are efficient large-volume licensing mechanisms in place for a large range of activities, such as media monitoring and music streaming.
Rightsholders wouldn't get much money from a licensing regime	Solutions for uses of masses of content have been in place for a very long time, including collective licensing by copyright management organisations. The market for licensing content for AI development is just starting to emerge. The value of third-party content to AI developers is clear. AI development will occur at many levels: not just the foundation models, but tools and applications built with smaller, curated datasets. Collective licensing means that the aggregate of payments from multiple sources can make a material difference to people working in the creative industries.

---

[2] https://japannews.yomiuri.co.jp/politics/politics-government/20230610-115423/;
https://japannews.yomiuri.co.jp/politics/politics-government/20231122-150910/

Claim	In fact
Licensing schemes will impede small developers	AI developers expect to, and do, pay for all other inputs. They do not ask hardware manufacturers, software providers or data centres to subsidise their activities. It is supremely unfair that the creative industries be asked to subsidise AI development.
	Licensing solutions are available to AI developers, big and small.
	Some local AI development will receive government assistance, for example as direct funding or taxation arrangements. Australian governments should remind recipients of their obligations to only use legitimately sourced content, and provide information about how to ensure this (including via licensing).
	The organisations making this claim are huge multinational companies. Their market power is not created by Australia's copyright system. It needs to be addressed by measures in Australia's competition regulation framework, such as the news media bargaining code.

Copyright and AI reference group

As the ACCC is aware, copyright-related AI issues are being reviewed by the Attorney General's Copyright and AI Reference Group (CAIRG). This group includes representatives from more than 60 stakeholder organisations, including from the creative industries, technology companies, education sector and library sector. We have found the process to be thorough, and well-coordinated with AI-related activities in other portfolios.

About Copyright Agency

We are a not-for-profit organisation with 40,000 members, who are primarily writers (including journalists), artists (including photographers, illustrators and cartoonists) and publishers (including in the educational, academic and news media publishing sectors).

We are appointed by the Australian government to manage special schemes in the Copyright Act that allow uses of content for education and government purposes, as they apply to text and images. These schemes apply to any content, from anywhere in the world, in any format (including print and online).

We also manage licensing arrangements for the business sector that allow uses of content from Australia and from other countries, under our arrangements with more than 70 similar organisations in other countries. These licensing arrangements include collective licences for media monitoring organisations, and collective licences for businesses that include news media content.

We collect and distribute around $100m in copyright fees each year, contributing to support for Australia's creative industries, including news media companies.