Technology Strategy Branch
Department of Industry, Science and Resources
GPO Box 594 | Canberra ACT 2601

digitaleconomy@industry.gov.au

26 July 2023

Submission from the Australian Society of Authors on Supporting Responsible AI

About the ASA

The Australian Society of Authors (**ASA**) is grateful for the opportunity to make a submission to the Department of Industry, Science and Resources in relation to artificial intelligence (**AI**) and welcomes the government's intention to mitigate the risks from AI.

The ASA is the peak body, professional association, community, and voice of Australia's writers and illustrators. We have almost 4,000 members drawn from every sector of the writing and illustrating world, including: novelists, non-fiction writers, biographers, illustrators, academics, cartoonists, comic artists, scientists, historians, graphic novelists, educational writers, children's writers, crime writers, science-fiction writers, romance writers, editors, bloggers, journalists, poets and more.

To prepare this submission, the ASA has conducted a survey of our members, held focus group meetings with professional authors about the ways in which AI is impacting their careers, and participated in discussions with equivalent author organisations around the world.

Executive Summary

Artificial intelligence offers opportunities and efficiencies, but also poses significant threats, particularly to the creative industries. In order to minimise the risks and embrace the upsides, we believe that AI ought to be carefully regulated, and we support a risk-based approach. We also support *mandated regulation,* in addition to voluntary regulation, given the speed and far-reaching consequences of AI development.

The creative industries are particularly vulnerable to disruption from AI in ways which will diminish skills, supplant talent, and have long-term consequences on quality cultural material. In a rapidly changing and complex landscape, there are so many issues to canvas including serious privacy concerns, but, due to our limited resources, we have confined our comments to issues within the ASA's core expertise. This submission will explore the risks Generative AI poses to the professional lives of creators and the social, cultural, and economic value of their creative work.

Our concerns are:

- the risk of copyright infringement and degradation of author rights
- the risk to incentives to create
- the risk to integrity in publishing.

Our recommendations are:

- *Mandate transparency around inputs*: oblige AI companies to be transparent about their training datasets
- *Support opt-in licensing*: copyright owners should be in a position to prohibit or authorise the exploitation of their works and be compensated for such uses
- *Mandate transparency around outputs*: AI-generated products should be labelled as such
- *Maintain copyright*: do not introduce new exceptions into the *Copyright Act* which would permit copying or mining of copyright works for AI-training purposes
- *Protect creator income*: consider a new scheme for the remuneration of creators, either by way of a cultural levy on AI products and services to ensure creators are paid for their work, or a universal basic income
- *Protect consumers*: ensure AI companies are liable for inaccuracy, bias, and the harm of misinformation

- *Mandate human oversight*: embed a requirement for human oversight in AI policy development, particularly in publishing, given our shared reliance on truth and integrity in the world's information channels
- *Establish a special expert group:* the particular vulnerability of creators across the arts requires a sector-specific response. An expert panel should consider copyright issues, appropriate regulation, and the impact of Generative AI on the cultural, social and economic life of Australia, to ensure the vision of *Revive* is supported and not thwarted by AI
- *Slow down AI development to conduct safety checks:* heed the calls of tech experts for regulatory sandboxes for AI - the training of AI systems more powerful than GPT-4 should be paused in order to put appropriate safeguards in place. Just as new drugs cannot be released to the public before their side-effects are tested, new AI tools should not be made publicly available until they are safe.
- *Work internationally*: Coordinate with as many other countries as possible to reduce the ability of OpenAI, Google, Microsoft and others to shop for permissive forums in which to set up practices.

Parallel to advocating for regulation, the ASA is pursuing an industry code of conduct to put guardrails in place in the book industry.

As we are a very small non-profit organisation, we are not sufficiently resourced to answer every question in the Discussion Paper. Our submission will primarily address:

- ***Question 2: What potential risks from AI are not covered by Australia's existing regulatory approaches? Do you have suggestions for possible regulatory action to mitigate these risks?***

- ***Question 3: Are there any further non-regulatory initiatives the Australian Government could implement to support responsible AI practices in Australia?***

We agree with and have adopted the definitions set out in the Discussion Paper.

Risk Analysis

1. Risk of copyright infringement and degradation of author rights

The Discussion Paper issued by the Department of Industry, Science and Resources states that the paper *does not seek to consider intellectual property as it relates to AI*. In our view, this results in a fundamental gap in the Department's inquiry, as **the appropriate use and regulation of AI cannot be debated until the harmful ways in which it has been developed are resolved**.

Generative AI Models have been built unlawfully

In the last six months, there have been rapid advancements in Generative AI software, particularly Large Language Models (LLMs) and Multimodal Foundation Models (MFMs). OpenAI has developed ChatGPT and DALL-E; Google has developed Bard; Microsoft has developed its Bing AI chatbot, in addition to significantly investing in OpenAI; Midjourney, Inc. has developed AI software that creates imagery and art from word prompts; and Stability AI has developed Stable Diffusion, software capable of generating photo-realistic images given a word prompt.

These Generative AI models are trained by ingesting vast amounts of text (the inputs or training dataset) to produce outputs.[1]

The quality of the outputs of Generative AI is reliant on the quality of the training dataset. To date, LLMs and MFMs have acquired their training datasets by scraping the internet, including digitised books, news articles, Wikipedia, blogs, search queries, Twitter and Reddit posts, YouTube videos, Flickr images, and more. While we know that books, journals, essays and articles have been included in the training datasets, there is little to no transparency over the details.[2]

According to OpenAI, their training dataset includes "two internet-based books corpora (Books1 and Books2) and English-language Wikipedia".[3] We do not know which books comprise "Books1" and "Books2" but suspect they have been substantially sourced from pirate sites.

We can also assume the datasets would include copyright works created by ***Australian creators***. When we visit large-scale piracy websites, novels by Australian authors are routinely found. Further, most of the training dataset used by OpenAI is in the English language,[4] putting English-language creators at the most risk of having had their work used.

---

[1] For example, Midjourney's founder David Holtz has referenced his company's scraping of a hundred million images from the Internet to build Midjourney.
[2] Narasimhan, K., Radford, A., Salimans, T., Sutskever, I. (2018). Improving Language Understanding by Generative Pre-Training.
[3] Brown, T. et al. (2020). Language Models are Few-Shot Learners
[4] Brown, T. et al. (2020). Language Models are Few-Shot Learners

Authors' rights have been ignored

Authors have not given their permission for their copyright works to be used in this way. In fact, most authors have no idea whether their works have been used or not; the circumstances are entirely opaque. At the moment, authors and artists cannot opt out of being included in the training datasets and cannot opt out of being named in the prompts.[5]

This is even more problematic where Indigenous Cultural and Intellectual Property (ICIP) is concerned. At a time when the National Cultural Policy has put 'First Nations first' and is working on stand-alone legislation to acknowledge and protect ICIP, Generative AI tools may be used to produce and perpetuate inauthentic and fake art, and appropriate Aboriginal and Torres Strait Islanders' art, design, stories and culture without reference to Traditional cultural protocols.

If the copying of the training datasets occurred in Australia, we believe it would be unlawful, but it is unlikely Australian law applies to the development of LLMs and MFMs due to jurisdictional issues. In the United States, where many AI companies are based, copyright law provides for a more permissive doctrine of 'fair use'. Nevertheless, we consider it unlikely that a court would find the appropriation of copyright works that have been illegally sourced is 'fair', particularly given the highly commercial nature of AI companies and the fact that the authors and publishers of those pirated 'inputs' never received any payment at all.

LLMs and MFMs will continue to scale and be monetised. For example, subscriptions to ChatGPT-4 currently cost approximately AU$30 per month. Accordingly, OpenAI is reaping revenue based on their software, which was built on the back of the intellectual and creative labour of creators without their permission, without acknowledgement, and without any of that profit being returned to the creators whose work enabled their technology in the first place.

Creator objections around the world

Globally, authors and artists are objecting to this unfair appropriation of their work. There are organised, outraged, large-scale creator protests occurring around the world.

The US Authors Guild has sent an Open Letter to the CEOs of OpenAI, Alphabet, Meta, Stability AI, IBM, and Microsoft objecting to large-scale copyright infringement. More than 10,000 authors from around the world, including Australia, have signed that letter. The ASA endorses this letter.

In the United States, the Writers Guild of America (**WGA**) has been on strike since 2 May 2023, concerned about the impact of both streaming and AI on the jobs and livelihoods of screenwriters. They were recently joined by the Screen Actors Guild – American Federation of Television and Radio Artists (**SAG-AFTRA**) who are concerned about, among other things, film studios using actors' likeness without consent or payment, supplanting actors with AI-derived content.

---

[5] Salkowitz, R. (2022). Midjourney Founder David Holz on the Impact of AI on Art, Imagination and the Creative Economy.

CISAC – the International Confederation of Societies of Authors and Composers – is the world's leading network of authors' societies, with 225 member societies in 116 countries, representing more than 6 million creators across music, audiovisual, drama, literature and visual arts. CISAC has penned an open letter calling upon governments to commit to developing and adopting policies and legislation that are consistent with upholding creator rights, licensing, transparency, credit, and legal responsibility for AI operators. The ASA endorses this letter.

Litigation has commenced around the world, with class-action lawsuits launched against Stability AI, DeviantArt and Midjourney,[6] Getty Images commencing legal proceedings against Stability AI for using its images without payment,[7] bestselling authors filing a class action lawsuit against OpenAI,[8] and comedian and writer Sarah Silverman and others commencing action against OpenAI and Meta alleging that their copyright has been infringed in the training of the companies' AI systems.[9]

Authors and narrators, supported by SAG-AFTRA, were outraged to discover that Findaway Voices, an audiobook distributor owned by Spotify, was providing Apple with access to some of its audiobook files to train their AI-narration tools - tools which threaten their livelihoods. Findaway Voices and Apple have recently agreed to halt any/all use of their files for machine learning purposes, however, we understand that authors are still finding machine learning clauses in the Findaway distribution agreements.[10]

This is a threshold issue. If the Australian public and private sectors proceed to adopt Generative AI, without firstly tackling the copyright infringement issue explained above and demanding appropriate licensing solutions, we will have tacitly approved the unremunerated appropriation of decades of authors and artists' work, disregarded the premise of copyright law, and decentivised future creative and intellectual labour.

---

[6] United States District Court Northern District of California San Francisco Division (2023). Class Action Complaint against Stability AI, Midjourney and DeviantArt.
[7] United States District Court for the District of Delaware (2023). Getty Images Complaint against Stability AI.
[8] United States District Court Northern District of California San Francisco Division (2023). Class Action Complaint Against Open AI.
[9] Milmo, D. (2023). Sarah Silverman sues OpenAI and Meta Claiming AI Training Infringed Copyright.
[10] Strauss, V. (2023). Findaway Voices, Machine Learning, and the New Rights Frontier.

Recommendations

Mandate transparency on inputs:

- Oblige all companies which offer Generative AI to be transparent about their training datasets. The copyright works they have "ingested" must be disclosed. Creators are entitled to understand how and to what extent their works have been, and will be, used, and for what purpose.

  For example, in Europe, the proposed ***Artificial Intelligence Act*** (AI Act) classifies the level of risk an AI technology poses to the health and safety or rights of a person. Some practices are so high risk they are banned outright, whereas other practices are regulated. Relevant to this submission, the AI Act requires creators of LLMs and MFMs to:

  - be transparent to end users about AI-generated content,
  - help identify deep fake images and
  - ensure that details of copyrighted data used to train their AI systems are publicly available.[11]

  In our view, these same requirements should be legislated in Australia.

Uphold moral rights under the Copyright Act

- Authors and illustrators are entitled to recognition and credit when their works have been exploited by AI systems. The obligation to be transparent about training datasets and the right to prevent their works from being treated in a way that would harm their honour or reputation is consistent with the moral rights of authors under the Copyright Act.

[11] Greene, L., Timon, V. (2023). Artificial Intelligence Act Passed by the European Parliament.

Opt-in licensing:

- To support licensing solutions, the Government could play an educative role and help inform users of the underlying copyright issues in order to increase pressure on Meta, Alphabet, Microsoft, OpenAI, Midjourney, Stability AI and others to negotiate fair licences with creators.

- Following the precedent set by the *News Media and Digital Platforms Mandatory Bargaining Code*,[12] the Government could legislate a code to incentivise commercial arrangements. In the same way that Google and Facebook would not have licensed news content from proprietors without the threat of designation under the code, similarly AI companies could be pressured to negotiate with writers and illustrators to enter into fair collective licensing.

- Licensing requires traceability and record keeping. To the extent that ISPs/OSPs are not already required to preserve metadata, new obligations could be introduced to maintain providence and authorship records of copyright works.

Maintain copyright protection for creators

- Do not amend the *Copyright Act* to introduce new exceptions that would permit copying for AI training purposes (such as a new text and data mining exception). Resist such calls from the tech sector.

[12] Australian Competition and Consumer Commission (2022). News Media Bargaining Code.

2. Risks to incentives to create

We are deeply concerned about the impact of Generative AI on the viability of author and illustrator careers in Australia. Margins in the book industry are thin and author earnings are, in the main, precariously low.[13] Many writers and illustrators rely on supplementary sources of income to afford to write books and make complex art, including copywriting, freelance writing, and concepting art jobs. It's likely that these jobs - jobs which help writers and illustrators pay the bills, but also teach them the necessary skills to become professional authors - will be taken over by Generative AI.

Within the publishing industry, it's likely that publishers will seek cost savings by reducing employment of editors, illustrators, graphic designers, marketers, and literary translators as machine learning tools are increasingly embraced. This risk is not very different to the risk faced in virtually every industry; that AI will replace human labour. Such losses must be balanced against the opportunities and efficiencies that AI will bring. While job losses are an economy-wide issue, we make two observations particular to authors:

(a) earnings are already so low that even a small disruption may mean the permanent loss of many professional writers, resulting in a contraction of insights, of voices, of unique Australian perspectives;

(b) humans actively want to make art; unlike the outsourcing of menial, dangerous jobs, the arts derive meaning from the very fact that humans create it. All art is ultimately about the shared human experience, which is antithetical to AI-generated content.

It's entirely possible to prompt Generative AI tools to write in the style of Helen Garner or Michael Robotham or Kylie Scott - any well known author - and generate knock-offs. What will stop anyone from generating entire books or artworks in the style of popular authors and illustrators and making that work available to buy online? Not only will this diminish the value of their creative work, established and prolific authors will be damaged by impersonations and mimicry. Like all public figures, authors are at risk of deep fakes, but also plagiarism which will be difficult to police in what is already a very crowded, noisy market.

A consequence of the ease of AI-generated books and art will be a corresponding "junkification"[14] and flooding of the market.[15] Several authors have created "a book in a day" with the help of AI tools and managed to upload and sell these books on Amazon.[16] An increase in AI-generated books will make the challenges of discoverability and dilution of audiences even tougher for professional writers.

---

[13] The latest Macquarie University research into author income confirms that the average earnings from writers' creative practice is $18,200 per annum.
[14] Herrman, J. (2023). The Junkification of Amazon: Why Does it Feel Like the Company is Making Itself Worse?
[15] Baker-Whitelaw, G. (2023). Amazon's Ebook Charts are Full of AI-Generated Spam.
[16] Bensinger, G. (2023). Focus: ChatGPT Launches Boom in AI-Written E-Books on Amazon.

The longer term risks, therefore, are:

- a future skills gap in professional writing, editing, design and artwork;
- a future where it is even harder to earn a living as a writer or artist. The reality in Australia is it is ***already*** impossible to earn a living wage from writing, other than for a few outliers. Even if artisanal pockets of creativity spring up, we may be left with hollowed out and diminished creative industries where only the privileged few make art.

It is beyond the ASA's submission to explore the impact of AI on literacy and education, however we note that a loss of writing skills means a loss of thinking skills as, from an early age, writing is how we interrogate problems, learn critical analysis, express ourselves and share thoughts.

As author, Ted Chiang, writes in *The New Yorker*, "The hours spent choosing the right word and rearranging sentences to better follow one another are what teach you how meaning is conveyed by prose. Having students write essays isn't merely a way to test their grasp of the material; it gives them experience in articulating their thoughts. If students never have to write essays that we have all read before, they will never gain the skills needed to write something that we have never read."[17]

Recommendations:

Protect creator income:

- Do not amend the *Copyright Act* to allow for copyright protection to be given to works solely or substantially created by AI, which would further decentivise human creativity.

- Introduce a cultural levy on AI products and services in Australia and use that levy to properly pay creators. There are precedents around the world for a levy on disruptive technology, such as the levy on the purchase of blank tapes and other recording media in recognition of the losses to creators. Calls for a universal basic income for creators have been strengthening as incomes from creators have diminished around the world over the last two decades. If drastic change is coming and we wish to embrace AI's upsides, we must think equally radically about new forms of remuneration so that the rise of digital platforms doesn't mean the decimation of creative careers.

[17] Chiang, T. (2023). ChatGPT is a Blurry Jpeg of the Web.

3. Risk to integrity in publishing

Outputs of Generative AI are biased and inaccurate

LLM software is a step-change in technology; it can both parse and produce convincing, coherent, human-like language. The significance of this is far reaching. As historian, philosopher and author Yuval Noah Harari has said, “AI has gained some remarkable abilities to manipulate and generate language, whether with words, sounds or images. AI has thereby hacked the operating system of our civilisation.”[18]

The superficial plausibility of LLMs obscures inherent biases and inaccuracies. OpenAI has admitted to bias on the grounds of race, gender and religion: “Broadly, our analysis indicates that internet-trained models have internet-scale biases; models tend to reflect stereotypes present in their training data.” [19]

It’s relevant to note that AI researchers are overwhelmingly white, male and American. Margaret Mitchell, a Microsoft employee, described AI researchers as a “sea of dudes” in 2016.[20] Australian Professor Toby Walsh has made a similar observation: “Not only are four-fifths of AI researchers male, they are also mostly white males. Black, Hispanic and other groups are poorly represented within AI, both in academia and in industry.”[21] Researcher, computer scientist, and author Dr Joy Buolamwini[22] has written extensively about the profound social implications of AI and the ways in which its inherent bias may wind back the advancements of the civil rights movement and has founded the Algorithmic Justice League.[23]

Microsoft, Google and OpenAI have also acknowledged that Generative AI Models “hallucinate” or make things up. “Facts” may be baldly stated, including with supporting citations which have been entirely fabricated.

No matter how coherent or plausible the outputs of LLMs seem, LLMs are not actually “thinking”. As Noam Chomsky says, “Whereas humans are limited in the kinds of explanations we can rationally conjecture, machine learning systems can learn both that the earth is flat and that the earth is round. They trade merely in probabilities that change over time.”[24]

---

[18] Harari, Y. (2023). Yuval Noah Harari Argues that AI Has Hacked the Operating System of Human Civilisation.
[19] Brown, T. et al. (2020). Language Models are Few-Shot Learners
[20] Walsh, T. (2022). *Machines Behaving Badly.*
[21] Walsh, T. (2022). *Machines Behaving Badly.*
[22] Buolamwini, J. Poet of Code.
[23] Algorithmic Justice League.
[24] Chomsky, N. (2023). Noam Chomsky: The False Promise of ChatGPT.

Generative AI can be harnessed for propaganda and harm

ChatGPT can be used to answer general knowledge questions, perform search functions, complete sentences or paragraphs, and translate between languages. One of OpenAI's claims is that: "... for news articles that are around 500 words long, GPT-3 continues to produce articles that ***humans find difficult to distinguish from human written news articles***."[25] [emphasis added] This should alarm governments everywhere; news is the fourth estate, a critical check on power, and AI-generated content lacks journalistic independence, fact-checking, and ethical decision-making.

According to OpenAI: "Any socially harmful activity that relies on generating text could be augmented by powerful language models. Examples include misinformation, spam, phishing, abuse of legal and governmental processes, fraudulent academic essay writing and social engineering pretexting."[26]

Generative AI currently has no mechanism for filtering bias and preventing misinformation.

Recommendations

Transparency around outputs:

- AI-generated products should be labelled as such to give consumers transparency and choice. A classification system could be introduced which makes clear which goods or services are:
    - AI-free
    - AI-augmented or assisted
    - AI-generated.

  Transparency requirements are likely to be sector specific so mandatory industry codes could be developed with industry input, with penalties for non-compliance.

Leverage existing consumer protection laws:

- Do not allow AI companies to use their Terms of Services to disclaim responsibility for inaccuracy, or bias, or harm caused by their products.

Human oversight:

- Given our shared reliance on truth and integrity in the world's information channels, embed into AI policies a requirement for human oversight in publishing.

[25] Brown, T. et al. (2020). Language Models are Few-Shot Learners
[26] Brown, T. et al. (2020). Language Models are Few-Shot Learners

Non-Regulatory Recommendations:

Slow down

The Future of Life Institute has published an open letter calling on AI labs to pause the training of AI systems more powerful than GPT-4 due to concerns about humans' ability to control, predict and implement safeguards around AI.[27] We acknowledge that AI may achieve profound improvements in humanity but those gains are only meaningful if they also avoid corresponding social, economic, and democratic catastrophes.

Signatory, Yuval Noah Harari writes, "Just as a pharmaceutical company cannot release new drugs before testing both their short-term and long-term side-effects, so tech companies shouldn't release new AI tools before they are made safe."[28]

Given the risks, and the overwhelming task of global regulation, we urge the government to slow down and join other governments in calling for a pause in development or, failing that, a moratorium.

Work internationally

Coordinate with as many other countries as possible to reduce the ability of OpenAI, Google, Microsoft and others to shop for permissive forums in which to set up practices.

Protect creators

In the wake of a highly welcomed National Cultural Policy, we request the government establish a special expert group to consider copyright issues, appropriate regulation and the impact of Generative AI on the cultural, social, and economic life of Australia, to ensure the vision of *Revive* is supported and not thwarted by AI.

[27] Future of Life (2023). Pause Giant AI Experiments: An Open Letter.
[28] Harari, Y. (2023). Yuval Noah Harari Argues that AI Has Hacked the Operating System of Human Civilisation.

Contact:

Olivia Lanchester, CEO
olivia@asauthors.org

Lucy Hayward, Marketing & Communications Manager
lucy@asauthors.org

AUSTRALIAN SOCIETY *of* AUTHORS

Gadigal Country,
Suite C1.06 22-36 Mountain St, Ultimo NSW 2007
T 02 9211 1004
www.asauthors.org