From: [redacted]
Sent:
To: AI Call For Views
Subject: Fwd: AI submission

Follow Up Flag: Follow up
Flag Status: Completed

Introduction
The Alliance of Independent Authors is a London-based organization representing thousands of authors from all over the world who run micro book-publishing businesses. These authors use digital publishing tools and services to self-publish and licence short- medium and long-form works of fiction, non-fiction and poetry in audiobook, ebook, and print book format.

Our membership includes these right holders and also a number of approved publishing services e.g book editors, designers, producers, distributors, and marketers, including individual freelancers, sole traders and SMEs, as well as larger corporations within the self-publishing sector.

A list of our members can be found at www.allianceindependentauthor.org [ALLi] and we have a large subscriber following at www.selfpublishingadvice.org, ALLi's Self-Publishing Advice Center.

This response to the UK government's have been compiled by Orna Ross Director of ALLi and Joanna Penn, founder of The Creative Penn and ALLi's Enterprise Advisor and award-winning creative entrepreneur. As well as our work for the author community, we are both working author-publishers publishing fiction, non-fiction and in Orna's case poetry. Both have sold books in almost 100 countries around the world and both appear regularly on Amazon and other bestseller lists.

General
We welcome the fact that the government wants to make the UK a global centre for AI and data-driven innovation and that the UK protects computer-generated works which do not have a human creator (s178 CDPA) and protects "the person by whom the arrangements necessary for the creation of a non-human generated work are undertaken" (s9(3) CDPA).

We agree that artificial intelligence and machine learning systems ("AI" or "AI systems") today autonomously generate creative works which would, were they created by humans, be eligible for copyright protection. AI systems are creative and communicative, rational and analytical, independent and autonomous, efficient and accurate, and capable of assigning free choice from alternatives.

And we appreciate the collaborative, consultative approach of the UK government's IPO in seeking a shared understanding of the main questions relating to intellectual property policy and artificial intelligence across a wide range of applications, including publishing.

A one size fits all policy approach to such a general-purpose technology with widespread applications is not possible. We welcome the open consultation process and call for views on artificial intelligence (AI) and copyright and associated rights and aim to provide the perspective of the independent author-publisher.

Author-Publishers

The use of AI systems in the production of works grows ever more common in what is being called "the 3A era" of advanced technology and the Self-publishing 3.0 era of independent author-publishing. More and more authors, publishers and other creators now use, and want to expand their use, of AI tools to generate text, as working drafts and completed works.

This fosters creative expression (by humans and machines) and is to be encouraged.

As ALLi has argued in its Copyright Bill of Rights, to be effective a contemporary copyright policy must be framed around how author-publishers work, trade and negotiate in the 21st century, and how today's readers actually discover, buy and read books. Self-Published works now outsell the work of all the traditional publishers combined and authors are leading the charge in terms of publishing innovation.

The parameters of this discussion are wide ranging. We have limited our replies to these questions around copyright and text, audio and video generated by AI systems to the topics that we at the Alliance of Independent Authors consider to be most pressing.

Replies

1. Do you agree with the above description of how AI may use copyright works and databases, when infringement takes place and which exceptions apply? Are there other technical and legal aspects that need to be considered?
1A: Identification
How will consumers know if a book is created by AI or not? How will that impact copyright and infringements? There are already books for sale on Amazon that have been generated by AI. These include those that are labeled as such, for example, Booksby.ai lists the books published. It also no doubt includes books that have an author name on and do not declare how they were generated.

1B: Translated works
Translations are a particular issue as a translation of a book should also reward the original creator. AI translation tools, such as Deepl.com, can create a decent translated version of a full-length book within seconds. That can easily be human edited and published and given that digital publishing is now international, the owner of the original work may never even know about it, let alone benefit from it.

1C: Use of out of copyright works to train Natural Language Generation models results in bias and diversity issues in output
The use of copyright works section notes, "It is possible to avoid infringing copyright by using licensed or out-of-copyright works. For example, an AI could be trained using the works of Shakespeare, which are no longer protected by copyright. But unless a work is licensed, out of copyright, or used under a specific exception, an AI will infringe by making copies of it."
The lack of clear licensing results in current Natural Language Processing (NLP) and Generation (NLG) models like Open AI's GPT-3 being trained only on out of copyright works, mostly books published 50+ years ago. They are therefore trained on works by an exclusive group of authors — usually older white, male, Christian, wealthy, English-speaking authors — and chosen by a specific group of curators across the centuries. This has resulted in the known issue of bias and lack of diversity built into AI datasets which has resulted in racism, sexism, and other forms of exclusion by the NLG systems.

2. Is there a need for greater clarity about who is liable when an AI infringes copyright?
Yes. New licensing law is necessary to take account of the growing use of AI in publishing.

3. Is there a need to clarify existing exceptions, to create new ones, or to promote licensing, in order to support the use of copyright works by AI systems? Please provide any evidence to justify this.

Yes. A financial model for authors and publishers to license copyright works to AI platforms depends on a robust licensing system.

As evidenced by the prevalence of pirated media in markets where it is not available to consume legally or easily by other means, there is a likelihood of using copyright work without permission increases if it is hard to access legally.

Currently, publisher contracts and self-publishing agreements do not include a reference or license to permit the use of copyright works in AI systems. Such a license should enable the original author to benefit from the Natural Language Generation models used to generate new works. For example, a publisher of popular romance books could train a model like GPT-3 to write romance in the style of one or more authors. Those original authors should be paid to provide their works to the model, or should benefit from the sale/license of the subsequent work (even though the copyright to the AI-generated work will reside with the publisher).

4. Is there a need to provide additional protection for copyright or database owners whose works are used by AI systems? Please provide any evidence to justify this.

Yes, as per 1C above. If the need to train these AI models on diverse and contemporary books is acknowledged, then there is a need for a license to use existing copyright works in these models.

Since NLG systems like GPT-3 are creating 'original' work, it will be impossible to know what has been infringed. Therefore, the easy availability of a licensing model for 'in copyright works' would encourage legal usage.

6. If so, what form should this protection take, who should benefit from it, and how long should it last?

There are three routes that copyright legislators can take when considering creative works generated by an AI system

- deny copyright protection for works that have been generated by an AI system
- attribute copyright of such works to the creator of the AI system
- attribute copyright of such works to the creator of the AI system and the (already copyrighted) creators of the works used to feed the system

We favour option 3, while acknowledging the challenges involved. The time period currently in place--50 years from the date the work is made (s12(7) CDPA)--seems optimal.

7. Do other issues need to be considered in relation to content produced by AI systems?

7.1: The Alliance of Independent Authors' guide to blockchain *Authors and the Blockchain: Towards a Creator Centered Business Model* explores how the new technology of blockchain allows a work to be tracked through the chain of use and creator acknowledgement and compensation for copyright holders to be given precedence. Services are coming into the publishing and self-publishing space with the intention of providing such solutions e.g. Access copyright, StreetLib.

The use of blockchain to monitor, attribute and recompense copyright is worth considering in relation to AI systems. Blockchain provides the ability to track micro-percentages so that, for example, 100,000 books were read into an AI system could each have their copyright acknowledged and a micro-percentage assigned for use of the text.

7.2: AI-generated voice technologies impact on audiobooks, video and multi-media content

The section on Other copyright protection for computer-generated works notes that "So-called "entrepreneurial works" – sound recordings, films, broadcasts and typographical arrangements – do not have an originality requirement. These belong to their producers, makers and publishers, regardless of their creative input. This protection would appear to apply to AI-generated material, without need for specific provision.""

This is not the case for independent authors, who retain their copyright while producing works on self-publishing platforms, works that may include an audio-visual component, from film poetry to dramatised fiction.

This is currently most notable in relation to audiobooks, which is driving significant profits in the publishing and self-publishing sectors at present. Current audiobook licensing rights are either too restrictive e.g. one narrator in one language in one territory, or too wide e.g. 'story rights' which could mean anything to effectively deal with the challenges of AI as outlined above.

We anticipate that multi-media works may prove equally popular with consumers in the near future.

It is suggested that audio, video and multi-media copyright for individual creators rights be included and specifically considered in any AI-licensing considerations.

8. Does copyright provide adequate protection for software which implements AI?
Yes. The current copyright regime and exceptions provide the scope for necessary new licensing but clarification is needed to take account of artificial and augmented intelligence. It is unclear how the concepts of "free and creative choices", "originality" and "personal touch" apply to AI works and clarification and a revised/ separate definition of originality and "joint ownership" in relation to AI is needed.

9. Does copyright or copyright licensing create any unreasonable obstacles to the use of AI software?

See 1 C above. Without adequate licensing laws, yes. The IPO consultation document suggests that "It is possible to avoid infringing copyright by using licensed or out-of-copyright works. For example, an AI could be trained using the works of Shakespeare, which are no longer protected by copyright" but, as pointed out by Diversity.ai, using out of copyright datasets perpetrates "dead white male" syndrome and language bias, entrenching outdated racism, sexism etc and failing to allow for diverse voices.

By withholding works in copyright, the publishing industry could be preventing the "development of AI and the free flow of data needed to improve innovation."

This could be mitigated with new licensing models.

We believe that AI-generated works should be protected by a new rights license but regardless of whether these types of works are protected by copyright, it is already difficult to determine if a work has been generated by a human, a machine, or both. Further technological solutions are needed to determine the authorship of works, and ensure the right type of licence and copyright protection is applied.

[redacted]

and

[redacted]