



*Written submissions on aspects of the Copyright
Amendment Bill [B13B-2017]*

*Submissions requested on aspects of the Copyright Amendment Bill
[B13B-2017] and the Performers' Protection Amendment Bill [B24B-2016]
as requested by the Select Committee on Trade and Industry, Economic
Development, Small Business Development, Tourism, Employment and
Labour of the Council of Provinces*

To: The Select Committee on Trade and Industry, Economic Development, Small Business
Development, Tourism, Employment and Labour, Council of Provinces

Ms Noziphiwo Dinizulu via ndinizulu@parliament.gov.za

and Mahdiyah Koff via mkoff@parliament.gov.za.

For further information please contact:

Dr. Andrew Rens, Senior Associate,
Research ICT Africa

arens@researchictafrica.net

Samantha Msipa, Research Fellow
Research ICT Africa
smsipa@researchictafrica.net

About Research ICT Africa

Research ICT Africa (RIA) is an African digital policy, regulation and governance think tank based in Cape Town. It conducts research on digital economy and society that facilitates evidence-based and informed policymaking for improved access, use and application of information and communication technologies (ICTs) for social development and economic growth. RIA seeks to support policy and governance that will reduce the uneven distribution of opportunities and harms associated with the intensifying processes of digitalisation and datafication. Through active participation in international, continental and national processes of digital governance RIA provides evidence-based alternative strategies in the areas of intellectual property, internet governance, data governance, cybersecurity, algorithmic governance, innovation that will produce more equitable and just outcomes. Understanding the digital economy, and how it can be the basis for innovation and entrepreneurship that serves the needs and challenges of marginalised communities – including women, youth, children, the elderly, and people in rural areas, for example – is an integral part of RIA's work.

Acknowledgements and declaration

This submission was drafted and prepared by Dr. Andrew Rens, and Ms Samantha Msipa, with valuable inputs, contributions and revisions from other RIA staff members. Any errors or omissions remain the authors' own.

The research informing this submission is made possible by a grant from Canada's International Development Research Centre (IDRC) to develop AI policy capacity in Africa

Availability

We confirm our availability to make oral representations based on this submission.

Overview

Research ICT Africa (RIA) welcomes this opportunity to submit written comments on proposed changes to the Copyright Amendment Bill [B13B-2017].

We make this public interest submission to help ensure that the intensifying global processes of AI and datafication can be harnessed to contribute to the national project of reducing poverty, unemployment and inequality and ensuring that the benefits of advanced technologies and opportunities to innovate, improve lives and livelihoods by deploying them are more evenly spread. We acknowledge the extensive work on the Copyright Amendment Bill [B13B-2017] (CAB) and the Performers' Protection Amendment Bill [B24B-2016] (PPAB) by the Portfolio Committee on Trade and Industry and the work of both the National Assembly and the National Council of Provinces in multiple rounds of public consultation on the two Bills.

Innovation Friendly Copyright

This submission is made in the knowledge that unequal opportunities to innovate reflect the structural inequalities that characterise our economy and the need in policy formulation to actively redress historical injustices reflected in inequality today. In previous written and oral submissions RIA has supported a number of provisions already included, and suggested changes to some proposed provisions. RIA continues to support the inclusion of an open ended provision that enables innovative uses without the necessity of amending legislation.

In this submission RIA focuses on issues raised by the rapid technological advancements in computerised techniques that are popularly called Artificial Intelligence or AI. RIA remains ready to offer our expertise on any aspect of the bills that affects the digital economy and information and communications technologies.

Specific comments

Artificial Intelligence, Data and Copyright

Artificial Intelligence is an emerging general purpose technology. AI has many beneficial uses, for example AI technologies have proven to be powerful research tools. One example is that an AI model can accurately predict the structure of proteins which will accelerate medical and biological research. AI raises copyright questions about the inputs that are used to create an AI model and also the outputs. The questions about inputs are analysed under informational analysis, and the question about outputs is analysed in the section on authorship and AI outputs.

S12A, informational analysis and AI

In its current form S12A represents an appropriate balance between the exclusive rights given by copyright and the flexibility needed for uses in essence unpredictable because they are innovative.

Section 12A is the only section of the Act that explicitly protects uses for a research purpose. Explicitly permitting research is crucial to encourage the kinds of research essential to artificial intelligence - a crucial technology of the 4th Industrial Revolution. Research in the age of the Fourth Industrial Revolution and Artificial Intelligence can make progress only if computational methods such as text and data mining can be carried out with regard to copyright works without fear of legal liability. Therefore computational analysis of copyright work requires clear legal authorisation by copyright legislation. The express inclusion of research and specifically informational analysis such as text and data mining as conceivable legitimate purposes makes sense for this reason.

Artificial Intelligence

Machine Learning is an important AI technique. Machine Learning requires vast data sets. This data is often not subject to copyright. For example factual weather data is not copyright, nor should it be because everyone should be able to use the same facts. But for some AI systems the 'input data' is text, images, video or computer code which are copyright works. Machine learning systems often require reproduction and adaptation of vast numbers of copyright works, often tens or hundreds of thousands of works. However these works are used by the system rather than a human, and for the novel purpose of machine learning. Even when human intervention is required, for example to label images, it is as part of the preparation and not for a traditional copyright use. Some machine learning systems appear to make use of copyright works in ways that require permission of the copyright holder of each work, however given the vast numbers of works involved this is simply impossible and requiring permission be granted by even a 1000 rights holder would be the same as prohibiting South Africans from developing machine learning systems in important areas. Therefore, it is essential that copyright law permits use for machine learning, and other AI technologies. The European Union has attempted to develop a stand alone exception, however this exception is regarded as unworkable (Geiger, 2021) while at the same time there is widespread agreement that most machine learning in the United States is authorised by §107 but that the provision is sufficiently flexible to prohibit exploitive practises (USPTO Report 2020).

Once enacted S12A will authorise use of copyright works for informational analysis including machine learning, provided that the machine learning in question is considered fair in view of the balancing exercise required by S12A(2). Where machine learning is used to produce competing works that would have an adverse effect on the market for the original then it likely would not be regarded as fair, but where it produces something other output, such as an algorithm that could recognise a cow as a cow then it likely would be. The internal balance in section 12A enables appropriate outcomes even when applied to novel

technologies.

S12A is the most appropriate way to deal with computational analysis and machine learning, since these are general purpose techniques that are constantly evolving and developing - thus an open ended approach that balances the rights and interests involved is the best way to keep pace with technological changes. In a S12A balancing inquiry some machine learning uses are not fair, for example, using only the works by a particular artist in order to imitate her style would not be permitted.¹

Recommendation:

The following purpose be added to those explicitly protected by the fair use clause:

after “research” insert 1 “,including informational analysis,”

The clause would then read:

“(i) Research, including informational analysis, private study or personal use...”

Authorship and AI outputs

When the Bill was first drafted AI technology was not able to produce outputs that closely mimicked human work such as text and images. However in the meantime some AI techniques have been developed that can produce text and images that resemble those produced by humans. For example Chat GPT can produce text in response to a prompt. Another example is Stability AI that uses

¹ That is the conclusion reached in the Opinion of the Office of Legal Counsel and Legislative Affairs of the Israel Ministry of Justice on “whether ML enterprises can make unauthorized use of copyrighted materials to train Artificial Intelligence (AI) system” 18 December 2022 available at <https://www.gov.il/BlobFolder/legalinfo/machine-learning/he/machine-learning.pdf> . The original opinion is in Hebrew but there is an English summary on page 3.

[stable diffusion](#); a “technique the AI system uses to generate output images that are similar to those found in its training data”.² These outputs are produced by models trained using very large collections of images or texts. Unlike preceding technologies the image or text output is not directed by a human other than by the prompt instead it is assembled by the model based on associations it has formed during training. Neither the user nor the engineers that trained the model can predict what an output will be like. The question is whether the resulting output should be treated as a copyright work.

There are a number of reasons why machine learning outputs should not be treated as copyright works. Copyright is intended to act as an incentive to human creators. Even when copyright vests in a juristic person or the State it does so only when a human creates a copyright work. The relationship of the author to the juristic person as an employee or someone commissioned to produce the work is the basis for ownership vesting in the juristic person. The producers of the AI model do not need the incentive of copyright. Millions of AI outputs are being produced without any clear copyright incentive. Many producers make money from their AI models using a freemium model, users can have a limited number of users free, but pay for more. That model does not require copyright to exist.

Copyright has historically extended only to human activities, and efforts to extend copyright to outputs from non-humans have been rejected. Extending copyright to AI outputs is a far reaching extension of copyright. The consequences are hard to predict. If however AI outputs are awarded copyright status now it will become close to impossible to undo since the grant of exclusive rights will give rise to an entrenched interest which will resist withdrawal of copyright. Extending copyright to AI outputs shouldn't be permitted until the costs and benefits are clear,

² Some outputs are so similar that some artists have alleged infringement. For example in [Andersen v. Stability AI](#) the plaintiff alleges that the AI generated artwork results from copies of the images used in the training dataset. However these kinds of disputes can be resolved under existing copyright law.

If AI outputs are treated as copyright works the result will be a very unequal competition between human creators and AI models that are able to produce outputs at a much greater speed and scale than is possible for any human authors. The result will be devaluing of human creativity.

Parliament has an opportunity to deal with this issue in the Copyright Act. If it does not do so then it leaves courts without the guidance of democratically elected representatives and may be persuaded to extend copyright to AI outputs. Parliament should clarify that only human creative outputs are copyright. Attempting to define AI or AI outputs in order to explicitly exclude them from copyright creates a risk that as technology changes that the definitions will become outdated. Instead it is suggested that the CAB makes it clear that copyright attaches only to human creativity.

Recommendations

Insert a sub-section that states:

2A (3) (a) Copyright extends only to the products of a natural person's skill, effort and creativity.

(b) In any dispute concerning whether copyright or authors rights apply to a product the author or their successor shall bear the onus of proving that a work or aspect of a work is the product of the skill, effort and creativity of the author.

References:

Geiger, Christophe *The Missing Goal-Scorers in the Artificial Intelligence Team: Of Big Data, the Fundamental Right to Research and the failed Text and Data Mining limitations in the CSDM Directive* (2021) PIJIP/TLS Research Paper Series no. 66.
<https://digitalcommons.wcl.american.edu/research/66>

United States Patent and Trademark Office *Public Views on Artificial Intelligence and Intellectual Property Policy*, October 6 2020,
https://www.uspto.gov/sites/default/files/documents/USPTO_AI-Report_2020-10-07.pdf