CISAC's written statement for the 4th session of the WIPO Conversation meeting on IP and Frontier technologies

CISAC wishes to thank WIPO for the opportunity to express its thoughts and concerns during this 4th session of the WIPO Conversation on Intellectual Property and Frontier technologies focusing on "Data beyond AI in a fully interconnected world".

One of the overriding objectives of CISAC is to advocate for and protect the rights of creators in all cultural industries. WIPO and member states have addressed time and again over the decades to technological advances which have disrupted the application and enforcement of intellectual property rights.

While CISAC acknowledges the need for the advancement of technology and welcomes such advancement, in particular, with regard to such a potentially enormously beneficial technology like artificial intelligence ("AI"), CISAC stresses that this technological development cannot be permitted to diminish or erode the rights of creators. Over the decades, creators have supported and accompanied the development of new technologies provided they did not significantly hamper and diminish their fundamental intellectual property rights.

In this short paper, CISAC wishes to address some of its thoughts on the use of data in the framework of data mining and data training.

Despite the myriad forms of AI in existence today, resulting in the difficulty of defining AI, one aspect of the technology that applies universally is the requirement of AI to have access to data for the purposes of machine learning.

Typically, AI created to generate artistic works has been trained on data constituting existing artistic works. When it comes to music, these training materials often contains works protected by copyright. Feeding training materials to an AI requires an act of reproduction of these materials, and this is an act which is restricted by copyright.

In order to train an AI on particular music, the music file is made available to the AI system, and that AI system mines the music file for the relevant data.

AI creators training AI without proper authorisation from the copyright owners of the copyrighted works used in training are infringing copyright simply by virtue of training the AI, as the training process requires a reproduction of the original work, and indeed, any work subsequently generated by the AI could also constitute an adaptation, which is also an act restricted by copyright.

Furthermore, creators of AI which generate musical works may wish to commercialise the resulting AI works. This would mean that copyright owners' works are being exploited to create profit for third parties without any remuneration to the rightsholders. CISAC notes that the use of copyright works as training data in AI

systems without any remuneration for this use of the rightholders' works unfairly harms and prejudices the rights of the rightholder.

Text and data mining ("TDM") is another matter of significant concern. TDM is defined as any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations.

An important point to note is that with TDM, the end goal is the data generated by the TDM analysis. When TDM is applied to musical works as part of the process of creating an AI/AI-assisted work, TDM is merely part of a longer process. The data gleaned from TDM of a copyright work is then processed through an algorithm which creates a new work. It is arguable that TDM as defined and envisaged does not actually apply to the use of copyright works in "training" algorithms.

As an example, in the European Union, Article 3 of the Directive 790/2019 on Copyright in the Digital Single Market ("CDSM") provides a copyright exception for TDM for the purposes of scientific research. We often see universities spin out technologies developed in those institutions for commercial purposes. The unforeseen implication of Article 3 is that copyright material could be used to train an AI, and once the AI is sufficiently trained and the "scientific research" completed, the AI would constitute a commercially viable product. CISAC believes that in the cases where an AI system is trained on copyright works for scientific research, and then generates works of its own once that system is trained, and those works are introduced into the stream of commerce, such commercial availability should retroactively negate the scientific research exception.

More pertinently, Article 4 of the CDSM allows for an exception to the reproduction right in a work for the purposes of TDM, unless the use of the works for TDM has "been expressly reserved by their rightholders in an appropriate manner, such as machine-readable means in the case of content made publicly available online".

CISAC notes with concern that a very pro-AI (and arguably anti-creator) TDM exception has been introduced in Japan. On the 1st of January 2019, the Japanese legislature amended the Japan Copyright Act so that copyright works could be used to train AI without infringing copyright of rightsholders1.

Similarly, the Republic of Korea is planning to introduce a TDM of the same nature in the framework of the copyright act review. Korean copyright law only protects "creative production" which expresses human thoughts and emotions. For the purposes of Korean copyright law, data which has been mechanically abstracted from music file, is considered to have neither creativity nor expression, which means it is not a work (or derivative work) which is subject to copyright protection. Therefore, the implication is that, under Korean copyright law, if a work has been converted into a data form and used in AI training, the rights of the rightsholder of such original work vis-à-vis its subsequent use in AI are extinguished.

In CISAC's opinion, developments such as in Japan and Korea should be avoided.

Thank you very much for your attention.

---

[1] Japan Copyright Act. Article 30-4 (Utilisation not intended for the enjoyment of thoughts or emotions expressed in the work)