4th Session of the WIPO Conversation on IP and Frontier Technologies

Intervention by Dr Rita Matulionyte, Macquarie Law School, Macquarie University

Thank you very much for the possibility to intervene. The discussion so far has shown that different IP rights can both enable and hinder innovation related to data. I would like to draw your attention to two additional problems with relation to trade secrets' role in ethical AI context

One problem that has been discussed to a limited extent is the effect of trade secrets on transparency and explainability of algorithms. Transparency and explainability of AI are among the core principles of ethical AI. In order understand how algorithms make their decisions or whether their decisions are fair and privacy is protected, interested groups might need access to parameters of algorithms and training data, which are often protected by trade secrets. It means that trade secrets might inhibit the transparency of data used and the explainability of algorithms.

In some situations, this might be especially problematic. For instance, the US courts have been using algorithm based COMPAS technology to estimate the risk of recidivism. When one of the defendants against whom this technology was used asked for the access to the parameters of the algorithm so that he challenge the decision made about him, the court denied request based on the trade secret privilege.

In this case trade secret privilege prevented a person whose rights were affected by algorithmic decision making to get an explanation how the algorithm made a decision, what parameters it used or on which data it was trained. As a result, trade secret protection over algorithms arguably violated due process rules.

The second problem to which I would like to draw your attention is that trade secrets over training data might preclude the enforcement of copyright and related rights in machine learning context.

Currently, most jurisdictions do not have a specific exception which would cover the use of copyrighted works as training data in machine learning projects. In the EU, the text and data mining exception is limited too. Therefore, unauthorised use of copyrighted content in machine learning projects might constitute copyright infringement.

At the same time, training data is often kept confidential and thus copyright holders normally do now know about the fact that their works were used in machine learning projects. When Google used 11,000 novels scraped from the Smashwords website to train its Google Assistant algorithm, right holders were lucky enough to learn about such use and many were not very happy. However, in many cases right holders do not know about their works being used for various machine learning projects and thus they cannot challenge such uses and cannot enforce their economic and moral rights.

I would therefore encourage a more in-depth discussion on how we could address this conflict between trade secrets and the demand for more explainability of algorithms and more transparency around data that was used to train the algorithms.