

Intellectual Property Protection for Artificial Intelligence

Position by Wikimedia's Free Knowledge Advocacy Group

Wikimedia's mission and projects intersect with artificial intelligence and machine learning in two major ways:

First, Wikimedia itself develops AI/ML tools to assist editors in improving the projects, such as finding gaps in Wikipedia's coverage, identifying poor quality articles or edits or categorising uploaded images to make them searchable.

Second, Wikimedia projects are designed to be used, reused, and repurposed. As they are freely licensed they are frequently used in training AI/ML systems.

Our experience with these systems highlights an important consideration when discussing authorship in relation to AI/ML systems: **such systems as they exist today are tools developed and deployed by humans**. While the term "artificial intelligence" has long been associated with the concept of artificial sentience, as it is used in research and industry today, it most typically means a system that performs tasks that previously required human cognitive processing, such as translating complex and varied audio recordings of speech into text. The most salient aspects of these systems — that they are complex; that their internal decision-making processes can be difficult to predict; or can run to some extent autonomously — do not mean that they constitute a cognisable legal entity or an author.

Complex algorithms that govern ordinary computer systems do not raise novel questions of authorship; unpredictable processes, such as generating random numbers or simulating chaotic phenomena, are still the tools of those who use them; and autonomous action by machines is at least as old as clockwork. The appropriate denomination for much of what is called AI today would be: Automated decision-making, ADM. There is no basic difference to be made just because some ADM systems produce outputs that look like works created by humans.

The Question of Authorship

The EU would likely be best served by avoiding to determine when a system is sufficiently autonomous and "creative" and instead focus on the one relevant real-life scenario, which is: **Automated systems designed to perform specific tasks by one party, and deployed in particular applications of those tasks by another.**

Within that scope, there are a number of questions relevant to IP development, such as whether and how authorship might be attributed between the tool's creator and its user. This hinges on whether the use of the tool maintains a sufficient connection between the user and resulting content for the latter to be called a creation in the meaning of authors rights. Such questions already exist and have been dealt with within IP law and we would advise the EU

to treat AI/ML as any other tool used for the production of potentially protected content, be it software or hardware. New layers of protection, on the other hand, would be inappropriate.

IP law affect the Development of AI/ML

We believe it is of critical importance to not only look at how AI/ML developments affect IP law, but how IP law may affect the impact of AI/ML systems on individuals' human and civic rights.

For instance, copyright policy affects what data sets are used to train AI/ML, frequently leading to skewed and biased systems. **Copyright policy should not be so restrictive as to create incentives for AI/ML systems to be trained exclusively on public data, or works from the public domain.**

Furthermore, when AI/ML systems involve the personal information of individuals, assertions of **trade secrets, copyright and related rights should not prevent affected individuals from exercising their privacy rights.**

Also, when AI/ML systems are used to make determinations that affect people's lives, such as setting credit ratings, bail amounts, or potential for future violations, **IP rights should not prevent inspection, auditing, and analysis of systems.** Intellectual property laws should not be usable as pretexts to evade accountability.

The sui generis database rights and its effect on AI/ML

The European Union recognises a sui generis database right. According to the European Commission's own [evaluation](#), Directive 96/9/EC does little to foster investment in database maintenance in Europe. On the other hand it caters for plenty of open questions and confusion when it comes to data re-use. AI/ML would likely benefit from more versatile and easily accessible training data if this right was to be repealed or at least made optional (which is permissible under international law).

Nota Bene: We supported the work on Creative Commons' paper on IPR in AI and recommend it for further reference.

Contact:

Dimitar Dimitrov, EU Policy Director, Wikimedia, dimi@wikimedia.be, +32497720374