

Turbocharging the IP System to keep pace with the data-driven world

Aman Sinha

adroit.aman@gmail.com || aman.sinha@in.bosch.com

<https://www.linkedin.com/in/adroitaman>

Intellectual Property Counsel
Robert Bosch Engineering and Business Solutions, Bangalore, India

Text of the Intervention

In my humble opinion, as an IP enthusiast, I would make four specific points to address the concerns regarding the sufficiency of the current IP system for data.

1. I submit that the valuation of data as a commodity has increased manifolds in the Digital Age. Thus, a global framework for a sui generis Intellectual Property right for Datasets appears to be necessary. In the absence of such a framework, contractual obligations may be created that would restrict or even prevent users from analyzing text or data of the datasets. However, in cases where significant commercial exploitation is happening through such datasets, some form of benefit-sharing needs to be made mandatory. This would ensure that the contributors in developing, arranging, manipulating data sets are also fairly rewarded.
2. In most countries, text data mining for non-commercial use is considered an exception to copyright infringement. I believe that the intent behind this is progressive and reasonable. However, the premise for this appears to be unnecessary. Text data mining attempts to unearth ideas and correlations contained in the data sets. The informational value that is derived is not explicitly expressed in the dataset, thus, not protected by copyright. Therefore, there is no need for a specific exception for text data mining, and the same should not hinder the progress of science, technology, and development.
3. Machine-Learned models and Machine Learning As A Service (MLaaS) utilities have become high-value assets in this new age of innovation. Therefore, there is an imminent requirement to address the threat of Machine Learning methods and predicted APIs working around Artificial Intelligence systems and Machine Learning models. The developers of such AI systems and ML methods need to be adequately protected from Model Extraction attacks that use adversarial clients. In addition, the usage of training datasets derived through unscrupulous means like Model Inversion attacks needs to be addressed.
4. There is an increasing concern over classifying data as a “public good”, even in my country, i.e., India. Through the introduction of a Data Protection framework, it is being deliberated to make commercial data (anonymized non-personal data) available to other competitors as a means to improve competition. This assumption of an anti-trust problem concerning the accessibility of data is deeply troubling. Given the lack of clarity in the existing Privacy and Competition laws in dealing with data, an attempt needs to be made to recognize new rights in relation to data.

As we enter a brave new world with connected, automated and intelligent devices, we'll have to figure out the rules to govern it.