

A United States Ontology Right

Christopher J. Palermo
Hickman Palermo Becker Bingham, LLP, San Jose, California
cpalermo@h35g.com – www.hickmanpalermo.com

11 November 2019

Overview

The present debate over whether, or how, to use patents to protect innovations in artificial intelligence and machine learning has revealed that present patent systems are ill suited to many of these sorts of innovations. While exceptions exist, many AI-ML innovations involve relatively little technical invention. They rely on well-established code bases and software implementations of statistical algorithms that have been well explored in the literature. Instead, innovation often arises from the combination of better, different or larger data sources; human recognition of the most predictive attributes or features of records in those data sources, after repeated experimentation; and human formulation of a configuration file to drive the software engine using parameters appropriate to the problem. While these improvements involve difficult analysis and human conception of solutions to problems, the sum of the improvements does not model well against established norms for patent protection. In a time when appellate jurisprudence and USPTO policy—apparently ebbing and flowing as the Deep State responds to executive change—are aligned against patenting inventions allegedly founded on mathematical concepts or mental steps, innovators' attempts to use patents to protect AI-ML innovations appear to be a fool's game. This paper proposes that the United States adopt a *sui generis* form of protection for these innovations which blends selected attributes of familiar IP regimes in a manner that provides an incentive to continued innovation while balancing the level of protection against the relatively lower level of inventiveness represented in these innovations.

A Simplified View of Machine Learning Innovations

Improvements in artificial intelligence or machine learning can be broadly classified in several categories. Two categories are suitable for patent protection:

- Improvements in unsupervised learning techniques such as better clustering algorithms or better dimension reduction techniques.
- Improvements in supervised learning techniques such as better logistic regression or linear regression algorithms, improvements in random forest, neural network or gradient boosting tree algorithms or architecture.

Depending on how the claims are drafted, inventions in these categories may raise patent eligible subject matter issues. However, they are analogous to other kinds of software inventions that have been pursued using patents for many years. Techniques for protecting these inventions using existing patent systems is outside the scope of this paper.

For patents, other categories of AI-ML innovation are more problematic:

- Improvements in data sources, content, organization and relationships. The author has experience in many invention disclosure meetings in which inventors assert that the benefit of the invention is, “We have access to data that others don’t.” That is, the innovation consists of collecting, or obtaining from others, and/or blending, training datasets that are new or different with respect to others, or unobtainable by others. In many cases, innovation derives from defining the schema or ontology of one or more datasets, or a blended dataset. The schema or ontology defines the attributes of data represented in records in the dataset. A trivial example of a schema is that records define vehicles; each vehicle comprises a record identifier, maker identifier, model identifier, date produced, propulsion type, last date of major maintenance and so forth. A schema or ontology may have any number of attributes and any level of hierarchical complexity depending on the domain.
- Better feature extraction through human recognition, after repeated iterations of training with a finite training data set, that a particular *top-N* set of features is most predictive of the correct output when an evaluation record is received and evaluated using a trained classifier. That is, for a dataset having *M* features, where *M* usually is quite large, repeated training runs reveal that a lesser set of *N* features is sufficient to provide 95% or greater confidence that the output is correct for a particular evaluation record.
- New configuration files to drive established software engines. Configuration may specify the number of neural network layers, data pre-processing steps, training parameters such as dropout points and other values that have a direct effect on results or performance of the generic software.
- A combination of the foregoing, or the machine learning model that results from them.

This paper asserts that this form of innovation is unsuited to patent protection and a different IP regime is needed to balance the public interest in freedom of action against the social benefit of a legal scheme that creates a continuing incentive to innovation. Stated bluntly, this paper views the level of human ingenuity, for the AI-ML categories identified immediately above, as too low to justify the high transactional cost, long term of protection and expansive remedies offered by the patent system. Instead, these innovations involve repeated application of the same data science techniques to different problem areas, while undoubtedly producing better results for the specified problem arising from the hard work involved in assembling the training data, conducting feature extraction and feature analysis, and forming useful configuration files. This work is socially beneficial and consumes the resources of innovators, so some form of protection is appropriate.

A United States Ontology Right

This paper proposes the adoption of new legislation to create a United States ontology right having traits drawn from several different familiar IP regimes. The following table summarizes, for six (6) established types of IP, attributes typically applied to determine whether the right is granted, whether the right is infringed, remedies for infringement, and term of protection.

IP Type	Attributes for Grant					Attributes for Infringement					Remedies			
	Eligibility?	Novelty?	Inventive step?	Originality?	Work and Risk?	Similarity?	Access?	Copying?	Threat?	Actual Damages?	Statutory Damages?	Equitable Injunction?	Statutory Injunction?	Term
Utility Patent	X	X	X			X				X		X		Long
Design Patent	X	X	X			X				X		X		Medium
Trade Secret	X						X		X	X		X		Perpetual
Copyright	X			X			X	X				X		Long
Utility Model	X	X								X		X		Short
EU sg DB Right	X			X	X	X	X	X		X		X		Short
Ontology Right	X			X	X	X					X		X	Short

The last row of the table proposes the attributes of a United States ontology right. The scope of protection of the ontology right is to include:

- A schema or ontology as described by the applicant, including specific identification of data fields or attributes and their relationships as a hierarchy, graph or other structure.
- A set of features of the schema or ontology that the applicant wishes to protect, having determined that those features yield desired results for training data of the schema or ontology, when applied to a particular problem.
- A configuration file operable with specified AI-ML software that is available to others for use under license.

Patent attorneys may have the immediate response that this definition of protection is too narrow. However, this author believes that most innovators and companies in the field would view it as sufficient to protect what they have actually contributed to the field of innovation and to protect against competitors who have not done the same work or investigation. In other words, sometimes good enough is good enough. This level of definition and protection is appropriate to balance relevant social factors.

As shown in the last row of the table, the United States ontology right would require a determination of eligibility, to ensure that the three (3) elements set forth above have been presented in an application. If the disclosure is eligible, then examination would only consider whether the disclosure is original to the applicant, in the copyright sense, and represents a degree of work and risk. The latter factor is adapted from the European database protection regime of 1996, under which databases are eligible for protection in Europe independent of originality if there has been substantial investment in obtaining, verifying or presenting the contents. (Several commentators have observed that machine-generated training datasets are not within the European database right as presently interpreted; training data resulting from investment should be; uncharacteristically, the European IP regime in this area appears “better” for rightsholders than the US regime.) If the answers to these questions are yes, then a United States ontology right registration is granted. There is no examination for novelty or inventive step in the patent sense.

At the infringement stage, courts would resolve whether an accused infringer is using a schema or ontology, set of features and configuration that are similar to those in the ontology right registration. “Similar,” in this sense, means functionally and structurally the same with an allowance for differences in terminology. The intent is that a standard akin to “substantial similarity” in copyright would be used. Questions of access, copying or theft are irrelevant and therefore innocent infringement is possible.

Remedies are limited to statutory damages, of an amount subject to discussion and debate, and a statutory injunction. The term “statutory injunction” means that courts do not have discretion whether to grant or withhold a court order against infringement. If infringement is found, an injunction against use of the subject matter of the registration is mandatory.

A short term, on the order of five (5) years, is envisioned. Technology in this field is fast-moving, and given the impact of an injunction right, a short term is appropriate.

The author recognizes that this proposal may be viewed as radical, incomplete or even flawed. The intent of this paper is to begin a conversation about the appropriateness of *sui generis* protection for one class of AI-ML innovations. Although others such as IEEE have argued to the USPTO that new forms of IP protection are not needed for AI, the author believes that this perspective dooms a large number of meritorious AI-ML innovations to the costly cycle of application, rejection, appeal and decision at the USPTO or the courts, for these innovations will be deemed either lacking in inventive step, or ineligible for patents. Stakeholders should recognize that trying to patent the kinds of AI-ML innovations described above will continue to be problematic simply because examiners, courts and competitors will not view these innovations as “inventions.”

Given the powerful forces presently aligned against patents for a broad swath of technology, including technology that is subject to massive investment and other risk, alternative modes of protection deserve serious consideration. The United States ontology right proposed here recognizes that contemporary AI-ML innovations involve a level of investment which deserves protection, yet likely exceeds the level of technical conception or invention appropriate for the patent system.

*

REFERENCES

Potter, JM, *Computer Software Patents and Computer Implemented Inventions at the EPO and the UKIPO—a primer*, WP Thompson IP, UK, 2018.

IEEE-USA Response to USPTO Request for Comments on Patenting Artificial Intelligence Inventions, 16 October 2019.

Kappos, D, *Pressing Issues at the Intersection of AI and IP*, draft paper of 14 October 2019.

Lavagnini, S., *Artificial Intelligence and Copyright Protection*, available online at Lexology, <https://www.lexology.com/library/detail.aspx?g=4f5fb5fa-b968-4049-8297-0cff617917b5>.

European Union, Directive 96/9/EC of the European Parliament and of the Council, 1996.

Norton Rose Fulbright, *The Database Directive, AI and the data economy*, 12 November 2018, available online at Lex Universal, <https://lexuniversal.com/en/news/20629>

Gervais, D, *Exploring the Interfaces Between Big Data and Intellectual Property Law*, 1 October 2019, available online at JIPITEC Journal of IP, IT and E-Commerce Law, <https://www.jipitec.eu/issues/jipitec-10-1-2019/4875>

González Otero, B, *Before the Singularity: Copyright and the Challenges of Artificial Intelligence*, 25 September 2018, Kluwer Copyright Blog, <http://copyrightblog.kluweriplaw.com/2018/09/25/singularity-copyright-challenges-artificial-intelligence/>

Chavannes, R, *IP protection of deep learning systems*, 22 October 2018, Fair Balance blog, <https://blog.chavannes.net/2018/10/ip-protection-of-deep-learning-systems/>

Liyanage, Y, *INSIGHT: Intellectual Property Challenges During an AI Boom*, 29 October 2019, Bloomberg Law Big Law Business, <https://biglawbusiness.com/insight-intellectual-property-challenges-during-an-ai-boom>