

MEETING OF INTELLECTUAL PROPERTY OFFICES (IPOS) ON ICT STRATEGIES AND ARTIFICIAL INTELLIGENCE (AI) FOR IP ADMINISTRATION

Geneva, May 23 to 25, 2018

ADDENDUM TO DOCUMENT WIPO/IP/ITAI/GE/18/1 REGARDING SUMMARY OF THE REPLIES TO THE NOTE ON APPLICATIONS OF AI TO IPO ADMINISTRATION

Prepared by the International Bureau of WIPO

1. The International Bureau received communications from the Intellectual Property Office of the Philippines (IPOP HL) and the European Patent Office (EPO) regarding applications of Artificial Intelligence to Intellectual Property Office administration. The following paragraphs should be inserted between paragraphs 13 and 14 of document WIPO/IP/ITAI/GE/18/1:

“14. IPOP HL is currently using a third-party search engine called DTSearch for its patent search operations. Similar to all other search engines, the system has the capability to perform incremental index, fuzzy search, and other functions. Although the said system is a low-end of AI, it is more powerful than the traditional database search. The ASEAN TMView, ASEAN DesignView, WIPO ASEAN Patentscope, and IPOP HL eTMFile platforms also make use of free search engine systems such as SOLR.

15. IPOP HL also uses COGNOS, a commercially available Business Intelligence Software, to support the management reporting requirements of the Office. In using this system, IPOP HL undertakes an ETL (extract-transfer-and load) process from the IPAS database into COGNOS readable packages.

16. Based on IPOP HL’s experience, the reliability of any search system would depend on the accuracy of the raw data to be processed. The IPOP HL has also noted that an image search system for trademarks would be a good enhancement of the IP administration system, while “data mining solutions” would provide the foundation for a knowledge management system that would be useful to address the inconsistencies in IP examination results.

17. The EPO has been active in developing business solutions using Machine Learning and AI in the following areas at various degrees of implementation:

- Automatic Pre-classification of incoming patent applications for allocation to corresponding units in charge of search and examination;
- Automatic Classification of patent documents according to CPC scheme;
- Automatic Re-classification of patent documents according to changes in CPC scheme;
- Automatic Search of prior art for incoming patent applications;
- Automatic generation of queries;
- Automatic annotation of patent literature;
- Automatic detection of problem/solution in patent document;
- Automatic detection of Exclusion from patentability;
- Automatic translation of patent document;
- Identification of migration/penetration trends of specific technologies (Computer Implemented Invention) in other technology sectors;
- Automatic Figure and Image search for patent drawings.

18. Through its DataScience team, the EPO is mainly developing its own Artificial Intelligence systems (respectively Machine Learning models) based on open source software libraries that are fit for purpose. Therefore the EPO is in the unique position to combine its DataScience team`s expertise with an unmatched business understanding through our examiners and a most valuable collection of data; i.e. historical saved search data and, of course, the EPO prior art corpus. The EPO also makes use of commercial products in the automatic annotation area through software providers in different projects. The EPO uses Patent Translate in the area of Machine Translation but is also developing its own machine learned translation. The EPO has generated its own reference data (gold standards) and system for measuring the performance of automated search tools. More specifically, this is supported by a benchmarking and evaluation framework to measure the benefit of automation improvements in search and a data science environment to analyze and prototype machine learning and data processing solutions. The EPO has developed a Patent Document Model (PDM) and its implementation in the Knowledge and Information Management Environment (KIME). Together they enable an enrichment oriented management of patent and other data for Machine Learning purposes.

19. The EPO does not consider personal anecdotal evidence and uses curated gold standards. Citation, classification, and categorization data was generated by thousands of highly skilled experts in decades of work. Frequently, generating the ground truth is the most labor intensive step when using machine learning. Development necessitates specialists to avoid obvious mistakes. The EPO is in a position to share its expertise on how to evaluate tools that are either in-house developments or external buy-ins. For both scenarios a deep understanding of how the evaluation is done is crucial for the success of the tool. There are many promises and expectations in the area of Artificial Intelligence and Machine Learning but small errors in the training and subsequently the evaluation can have a disastrous impact once a badly trained system goes live.”

2. Consequently, in the document, paragraph 14 onwards should be renumbered accordingly.

[End of document]