

March 7, 2014

Via electronic mail to SoftwareRoundtable2013@uspto.gov

Attention: Seema Rao, Director, Technology Center 2100

IBM Corporation comments in response to “Request for Comments Regarding Prior Art Resources for Use in the Examination of Software-Related Patent Applications”.

IBM thanks the United States Patent and Trademark Office (“Office”) for the opportunity to comment on (i) publicly available prior art resources which may be used by the Office to improve the quality of examination of applications for patents related to software technologies, and (ii) concerns regarding the manner in which Office examiners formulate and implement search strategies to identify prior art for software related inventions and how those concerns may be addressed. Although IBM’s comments are offered in the context of examination of software-related patent applications, these comments are also generally applicable to many other technical fields of invention.

In response to the questions:

1. What specific databases, web sites, tools and other resources do you find useful in searching for software-related inventions? Please indicate strengths and limitations of each resource.

IBM suggests the following resources (not listed in any specific order) for consideration by the Office. The resources are believed to be useful in searching for software-related inventions.

- **IP.COM** (<http://ip.com/>) – This database provides access to the Software Patent Institute (“SPI”) database of software technologies which is available exclusively through this database. SPI focuses on finding hard-to-access software data. Furthermore, this database provides access to a wide range of technical documents from many third parties including IBM Technical Disclosure Bulletins and IBM Redbooks. This database also provides semantic search interface. Limitation: search for patent data is limited to U.S., EPO, and WIPO.

- **Derwent World Patent Index** (http://www.stn-international.com/dif_dwpi.html) – Derwent database provides access to patent documents from 48 patent issuing authorities as well as providing access to non-patent literature on a wide range of subject matters. Furthermore, the abstract of invention disclosures submitted to the Research Disclosure website (<http://www.researchdisclosure.com/>) is also made available through Derwent. Limitation: the database does not offer full text.

- **FAMPAT – Questel** (<https://www.questel.com/index.php/en/2012-11-20-10-09-15/fampat>) – FAMPAT has coverage of over 90 countries/authorities with bibliographic information, patent full text, legal status, and concept searching capabilities. Limitation: may require extensive training as the database has many advanced search features.

- **ACM Digital Library** (<http://dl.acm.org/>) – This database provides access to full text of every article ever published by the Association for Computing Machinery and bibliographic citations from major publishers in computing (journals, transactions, conference proceedings, etc). Limitation: Document delivery will be required for non-full text citations.

- **IEEE Xplore** (<http://ieeexplore.ieee.org/Xplore/home.jsp>) – This database provides full text of all publications by the Institute of Electrical and Electronics Engineers (journals, books, conference proceedings, transactions, standards, etc), as well as IEEE publishing partners. The site also provides over 300 educational courses. Limitation: none known.

- **The DBLP Computer Science Bibliography** (<http://www.informatik.uni-trier.de/~ley/db/>) – This database provides bibliographic information on major computer science publications. According to its website, DBLP currently indexes over 2.3 million publications from more than 1.2 million authors. The long list of computer related topics currently covered on this site can be found at:

<http://www.informatik.uni-trier.de/~ley/faq/What+is+the+scope+of+dblp.html>

Limitation: The information is bibliographic only. This information is best used to further analyze search results from other sites. Also, many full text articles require purchase through third parties. The search engine has limited features.

- **CiteSeerX** (<http://csxstatic.ist.psu.edu/about>) – This database is a library of computer literature as well as resources such as algorithms, data, metadata, services, techniques, and software that can be used to promote other digital libraries. CiteSeerX offers a long list of features including full-text indexing of the entire articles and citations. Limitation: None known.

- **Networked Computer Science Technical Reference Library** (<http://www.ncstrl.org/>) – This database houses a collection of computer science technical reports made available for non-commercial use by over 100 institutions. Limitation: Search engine has limited fields and features for searching.

- **The Collection of Computer Science Bibliographies** (<http://iinwww.ira.uka.de/bibliography/>) – This website is a collection of bibliographies of computer science literature. According to this website: (i) it contains more than 3 million references where more than 1 million of references contain URLs to an online version of the paper; (ii) abstracts are available for more than 800,000 entries; and (iii) more than 2,000 links to other sites carrying bibliographic information. Limitation: Search engine has limited fields and features for searching.

- **Google Scholar** (<http://scholar.google.com/intl/en/scholar/about.html>) – This is a search engine providing access to a large number of articles, theses, books, abstracts, patents, court opinions, academic publishers, professional societies, online repositories, universities and other web sites. Limitation: limited search features.

- **Innography** (<https://www.innography.com/>) – This website offers a wide range of tools for gathering and analyzing patent related data where according to its site, the data is constantly updated from 102 patent jurisdictions. The site offers semantic search features. Limitation: subscription-based access.

- **Corporate and university technical report archives** – Many companies and universities make their technical reports available on their websites. These documents can be useful sources of technical information. An examples of such sites include: **MIT** - <http://libguides.mit.edu/content.php?pid=360757&sid=3326963>; and **School of Computer Science at Carnegie-Mellon** <http://reports-archive.adm.cs.cmu.edu/>. Limitations: Additional time required searching several specialized sites by examiners; some reports may not be available in full text and will require document delivery services; some sites have limited search capabilities.

2.What are your concerns regarding the manner in which USPTO examiners formulate and implement search strategies to identify prior art for software related inventions? How should these concerns be addressed?

IBM applauds the availability of many search tools to the examiners on the examiners' desktop computers including Examiner's Automated Search Tool ("EAST"), Web-Based Examiner Search Tool ("WEST"), and the Foreign Patent Access System (FPAS), as well as the availability of trained professionals at the Technology Centers (Information Technology Resource Person ("ITRP")) and in the Scientific and Technical Information Center ("STIC") for searching non-patent literature ("NPL"). Having said that, IBM suggests careful review of: (1) the manner in which USPTO examiners formulate their search strategy; (2) the types of literature they search; and (3) which databases or other resources they use to identify prior art.

Regarding the manner in which USPTO examiners formulate their search strategy, it is important to realize that in the field of software technologies many different terminologies are employed to describe the same or similar concepts. Therefore, in order to formulate a sound search strategy, a good understanding of the alternative terminologies employed to describe software inventions is essential. Such understanding requires ongoing training of the examiners through a variety of channels, including subject matter experts on search strategies, and software technology experts from industry and academia. Examiners training should also focus on the importance of employing different search strategies depending on the source. In other words, it is important to be mindful that the search strategy employed for patent searching often has to be modified for non-patent literature databases, especially those databases which provide information limited to bibliographic, abstract, and indexed data. In this regard, the use of semantic searching should be strongly encouraged and be made part of the training. Examiners should also be encouraged to consider use of a more iterative approach to

develop search strategies, in which they analyze search results and use the best identified prior art references to modify search terms. Furthermore, examiners should be encouraged to reach out to applicants during the search strategy phase if they conclude such a conversation could help them to better formulate their search strategy.

Regarding the types of literature examiners search, under section 904.02 of MPEP, an examiner is expected to conduct a thorough search by, among other things: (A) *identifying the field of search*. This section furthermore states that in determining the field of search, “three reference sources must be considered - domestic patents (including patent application publications), foreign patent documents, and non-patent literature (NPL).” According to section 904.02, “[n]one of these sources can be eliminated from the search unless the examiner has and can justify a reasonable certainty that no references, more pertinent than those already identified, are likely to be found in the source(s) eliminated.” (Emphasis ours).

IBM is concerned that the exception in this section (i.e., “unless the examiner ...”) may in practice result in no search (or at best a limited search) of non-patent literature due to a number of reasons that include a lack of robust training on how to search non-patent literature, and a lack of sufficient training on semantic searching. Indeed, it is not clear to IBM how an examiner can conclude with reasonable certainty that searching non-patent literature will not identify a more pertinent reference without actually searching the non-patent literature. In our experience, it is often the case that the best prior art reference, particularly pertaining to software inventions, is to be found in the NPL. Furthermore, as Office also realizes in certain fields (such as high technology) patent documents may not be the best source of prior art. Therefore, IBM suggests eliminating this exception and making the search of the non-patent literature mandatory in every search, at least for software patent applications.

Finally, as to which databases or other resources examiners should use to identify prior art, the determination of which database to use is at least as critical (if not more critical, to improving search results) as proper training and mandatory search of non-patent literature. IBM’s reply to question one makes it amply clear that there are significant on line resources available, many of which are primarily focused on software technologies and include non-patent literature, which should be carefully considered by the Office when conducting patentability searches of software inventions. Many of these resources provide semantic search capability which is critical to employ in the field of software technologies, in which (as stated earlier) many different terminologies are employed to describe the same or similar concepts.

Conclusion

In conclusion, IBM believes (i) increasing the level of training, both in terms of software technology itself as well as search techniques (such as semantic searching) for non-patent literature, and reaching out to applicants as necessary to formulate the search strategy; (ii) mandatory search of non-patent literature; and (iii) careful selection of which databases to use during the patentability search will provide the examiners with a higher quality software patentability search results, which should ultimately result in better examinations and higher quality issued patents.

Respectfully submitted,

Manny W. Schechter
Chief Patent Counsel
Intellectual Property Law
IBM Corporation
schechter@us.ibm.com
(914) 765-4260

Paik Saber
Senior Corporate Counsel,
Intellectual Property Law
IBM Corporation
saber@us.ibm.com
(408) 927-3390