# IP-Explorer: A Semantic Web Based Intellectual Property Knowledge Base and Trading Platform

Lei Zhang[1], Hugh Glaser[2], Hai-Tao Zheng[1] and Ian Millard[2]

[1] Tsinghua-Southampton Web Science Laboratory at Shenzhen, Graduate School at Shenzhen, Tsinghua University
{zhanglei, zheng.haitao}@sz.tsinghua.edu.cn

[2] School of Electronics and Computer Science, University of Southampton
{hg, icm}@ecs.soton.ac.uk

**Abstract.** In this paper, we demonstrate IP-Explorer, a semantic web based IP knowledge base and trading platform with the following characteristics: First, it is based on the semantic web technology; Second, it is built by mining the China High-Tech Fair(CHTF) database and the State Intellectual Property Office of PRC (SIPO) database; Third, "Smart Pushing" proactively pushes useful information based on user profiles, browsing and searching records, and activities of peers; Fourth, "IP family" collects IP's from different industry sectors and integrate them as a "family", helping the user to broaden his horizon; Fifth, "Technology Roadmap" provides a visualization tool for decision makers to observe technology history and trends.

**Keywords:** Semantic Web, Intellectual Property, Data Mining.

## 1 Introduction

Intellectual Property (IP) is one of the key incentives that help to keep industry innovations active by granting exclusive rights to the owners. Acquisition of new IP's is considered to be a feasible way to survive the financial crisis for Small and Medium size Enterprises (SMEs). There have been a number of IP trading platforms supporting browsing and searching of IP's. However, most of them require good background knowledge when selecting searching keywords in order to get acceptable results.

We are facing a dilemma now. On the one hand, there have been a huge number of IP's available. On the other hand, owners of SME's find it difficult to use current IP systems. Therefore, we need to build a new IP knowledge base and trading platform to bridge IP providers and consumers. The employment of the Semantic Web technology makes it possible.

## 2 The System

The platform builds a semantic web database for IP's from the High-Tech Fair [1] and the SIPO [3][1].

[1] To access the data, visit http://www.weblab.sz.tsinghua.edu.cn/.

The China High-Tech Fair is the largest and most influential exhibition on science and technology in China, regarded as the No.1 Chinese exhibition in this industry. Each year, CHTF would receive more than 500,000 people, and accomplish a total of 13 billion US dollars in terms of product and technical transactions. The database of CHTF contains information about 16726 projects along with their kernel technology IP numbers.

SIPO is affiliated to the State Council of China with the responsibility to organize, administrate, and coordinate IPR protection work nationwide. The database of SIPO contains more than 7'000'000 entries of granted IP's classified with the International Patent Classification (IPC)[2].

Our IP-Explorer system is based on the following two steps of data processing:

1. Semantic Web representation of IP-related information including IP number, author, affiliation, IPC category, abstract, and full-text. This is implemented using 3store [5] RDF repository and the core RKBExplorer [4] as user interface.
2. An ontology design for the IPC tree structure and a cross-link discovery algorithm to establish links among semantic-related categories. For original tree links, the ontology is simply "isA" relation. For cross-links connecting semantic-related IPC categories, they often take forms of "useMaterial" and "useDesign". The cross-links are discovered by abstracting keywords for IP descriptions, clustering co-authors, and then mapping back to the IPC structure.

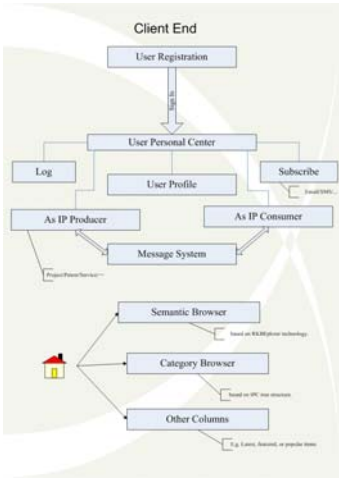


**Fig. 1.** Client-side Module.

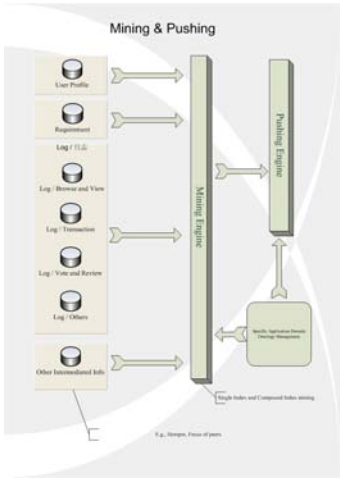


**Fig. 2.** Mining and Pushing Engines.

Based on the semantic web technology and data mining, IP-Explorer is different from traditional IP platforms in the following three phases of transaction:

- **Pre-sale Smart Pushing**: Instead of asking the consumers to "pull" information by searching, the new platform should proactively "push" related

[2] The International Patent Classification (IPC), established by the Strasbourg Agreement 1971, provides for a hierarchical system of language independent symbols for the classification of patents and utility models according to the different areas of technology to which they pertain.

IP's based on the user profile, browsing and searching history, and peer activities. This feature helps users without computer skills to get exposed to many useful IP's quickly and is supported by client-end multiple-panel browser and server-end mining and pushing engines, as shown in Fig. 1 and Fig. 2.

- **In-sale IP Family Integration and Technology Roadmap**: The user may start with one query on chips without realizing the importance of other devices, packaging, design, or materials. By integrating IP's from different industry sectors, the system would help the user to broaden his target area. Technology Roadmap provides a visualization tool for decision makers by observing technology development over time and indicating future trends. Analysis of IP location distribution (see Fig. 3) is provided to evaluate industry-chain maturity.

**Fig. 3.** Geographic proximity analysis of China photovoltaic industry technology roadmap.

- **After-sale Service and Support**: The system helps to identify people with matching expertise to provide after-sale service and support. The feature is similar to that of ExpertFinder [1], but the module is seamlessly integrated as part of our platform instead of another separate system.

## 3 Conclusion

IP-Explorer, to the best of our knowledge, is the first Semantic Web based IP knowledge base and trading platform. Ontology design and main features are original and beneficial.

## References


1. ExpertFinder Initiative, http://expertfinder.info/.
2. The High-Tech Fair, http://www.chtf.com/.
3. State Intellectual Property Office of China, http://www.sipo.gov.cn/.
4. Glaser, H., Millard, I. and Carr, L., RKBExplorer: Repositories, Linked Data and Research Support. In: *Eprints User Group, Open Repositories 2009*, 20/05/2009, Atlanta, GA, USA.
5. Harris, S. and Gibbins, N., 3store: Efficient Bulk RDF Storage. http://eprints.ecs.soton.ac.uk/7970/.