NEWSLETTER

知产快报

- The challenges we are facing in the patent industry include the sheer volume of information, the diversified languages of the information, and the increasing complexity and obscurity of the information. Artificial intelligence (AI) can help us in managing the challenges with search and retrieval of information where intelligent semantic technology has progressed to become a useful tool. Will it be the right-hand man of patent search experts or will it terminate the jobs of all patent search experts? Based on the rich experience using intelligent semantic search in patent examination, the author here explains the future of technology and lays out its different aspects in details including the causes, basic principles and the application methods of intelligent semantics in patent search.

AI Semantic Technology in Patent Search

The challenges we are facing in the patent industry include the sheer volume of information, the diversified languages of the information, and the increasing complexity and obscurity of the information. Artificial intelligence (AI) can help us in managing the challenges with search and retrieval of information where intelligent semantic technology has progressed to become a useful tool. Will it be the right-hand man of patent search experts or will it terminate the jobs of all patent search experts? Based on the rich experience using intelligent semantic search in patent examination, the author here explains the future of technology and lays out its different aspects in details including the causes, basic principles and the application methods of intelligent semantics in patent search.

Patent search methodologies have constantly in development for more than hundred years. As early as the end of 19th century, the United States government has applied the patent search technology in its consulting work, and it used a manual search on paper patent filings. Since 1970s, the computer-based patent search has become more popular and the efficiency of patent search has been greatly improved. However, for the public at that time, it was still very difficult to obtain patent information. It wasn't until 1997, when IBM began offering patent information services on the internet, that it became relatively easy for the public to access patent information. Throughout the development history of patent search, along with the change of information technology, the way of obtaining and using patent information advanced as well.

Since the beginning of the 21st country, the innovation of information technology has been accelerating, and the development of AI technology took a frog leap. New technologies such as shopping recommendation algorithm, face scanning payment and voice assistant have been integrated into our daily life. In the patent information processing, a new technology called intelligent semantics has also entered into an important period of development. In May of 2018, WIPO held a conference "Intellectual Property Office on ICT strategy and AI". According to the related materials, IP offices in many countries and regions have started to apply intelligent semantic technology in patent examination. In the fields of commercial patent database, several tools have equipped with intelligent semantic search function. What changes have new technologies brought to patent search?

The Difficulties of Traditional Patent Search

The difficulties not only lies in the field of patent search, but in all traditional search fields. Information explosion is the main source of difficulties. With the exponential growth of human knowledge, it is not easy to search for the required information quickly and accurately. The field of patent also faces the same problem of data growth. Taken China for example, the total number of patent filings including the patent applications for invention and utility model published jumped from less than 2.5 million at the end of 2008 to more than 17 million at the end of 2018. This is a sevenfold increase in 10 years and it brings huge challenges to the patent search.

Based on the traditional Boolean search method, search experts are committed to finding a balance between comprehensiveness and efficiency. However, the complexity and obscurity of technical concepts, let alone the diversified languages, have complicated all the searches as the search experts have to list ways of expressions as many as possible.

The Rise of Intelligent Semantic Search

In view of the difficulties, a new semantic technology that automatically expands the search content has emerged. Specifically, a search expert can build a large thesaurus and search the contents entered by users in the thesaurus in order to improve the success rate by extending the relevant keywords. However, although the technology improves the comprehensiveness of patent search, the application of the algorithm hits more patents in the searching process and involves more unnecessary file reviews.

Another smart semantic technique is to sort information based on semantic relevance. This technology is based on the patent text data to carry out machine learning and to automatically calculate a correlation between words and

documents and then sort information based on this correlation. The goal of this technique is to rank the most relevant technologies at the top and directly improve the efficiency of the search while reviewing files. The patent search tool based on this technology allows the users to enter any length of text or enter a patent number and the system will automatically recommend the most relevant literature.

Application of Intelligent Semantic Technology in the Field of Patent Search

As for the semantic search technology based on thesaurus, since the computer only expands the content input by users, the Boolean search is still performed in the specific patent search, so there is not much difference between this semantic search tool and the traditional Boolean search tool, and thus it is easy to understand and use.

However, the search technology based on latent semantic index is completely separated from the traditional Boolean search, which is quite different from our understanding in the past years. This search technique completely abandons the process of defining a subset from a document set, and instead sorts a document set according to the relevant content of the text input by the user. The following is a brief introduction to the two methods of patent search using the said technology.

The first approach is to use intelligent semantic search completely independently. At this time, searchers only need to input a piece of text that they want to put into the semantic search system. The computer can sort the tens of millions or even hundreds of millions of patents stored in the database according to the text input by users. Searchers may obtain the required patents by browsing in turn. Because this search method does not need any search strategy, people who have no experience in patent search could also get started quickly. For those people who have rich experience in patent searching, this method can also be used for exploratory searches at an early stage and there will be a certain probability and could have a satisfactory result. In addition, more keywords or related patent classification numbers can be sorted out through a statistical analysis.

Using intelligent semantic search independently could improve the efficiency of the patent search but it can't replace human search experts and Boolean search. One of the

main reasons for this is that the intelligent semantic technology is unable to understand the technical contexts. In other words, the intelligent semantic technology is not able to extract the core technical features of the technical contexts to carry out an accurate search like the patent search experts. For the following text, for example, patent search experts will initiate the patent search based on the understanding of the descripted technical solution and the prior art, the character "gap" and "bending" will be deemed as the core feature to search, but the current intelligent semantic algorithm could not simulate the abstracted thinking process from text to a technical solution.

A one-piece autodyne device, including telescopic rod and clamping device for clamping filming equipment, clamping device including objective table and described in above the object stage can stretch clamping mechanism, whose character is: the clamping device described one-piece rotation mentioned in connection to the top of the telescopic rod, mentioned on the slide has a gap, is equipped with a clamping mechanism and described gap position corresponding to the bending part, after folding telescopic rod can let into the gap and bending.

The best way to use intelligent semantics is in conjunction with Boolean search. To be specific, after obtaining the searching results by Boolean search, then utilize a sorting factor which is independent of Boolean search to have a new sort. Boolean search is used to accurately express the core features of the technical scheme. Intelligent semantics express other non-core features in terms of relevance and present them to search experts in a rank. The advantage of such a search strategy is to utilize human thinking, i.e., search experts' thinking in defining core feature, and thus ensure the accuracy of search. Meanwhile, it augments human thinking via intelligent semantic algorithm search to ingest, digest, understand, and analyze the tremendous amount of data to provide insights.

According to the above-mentioned example, if Boolean search is used solely, then the features such as: autodyne device, clamping device, gap, etc. should be limited at the time of searching. But if Boolean search combines intelligence semantics, the searcher could only use Boolean search to search core feature such as: gap, bending etc. and then use the full text as a factor to sort the search results, then he will get a new set of hits, i.e., the relevant patent applications concerning selfie devices and clamping could be

ranked on the top of the result. This search method not only improves the accuracy of the search, but also avoids missing important patents.

Thinking of AI, many people worry about jobs currently done by humans being replaced by machines. However, I believe that in patent search it shouldn't be used to replace human thinking as the search requires a lot of subjective thinking, judgement and comprehensive analysis.

Search experts do not need to regard intelligent semantics as competitors at all, instead, take it as an import tool in the future to empower and assist themselves. It is believed that the combination of semantic technology and expert Boolean search will be the mainstream in the future, and the semantic ordering based on massive data model and the precise hits based on Boolean will play a perfect movement of the integration of humans and AI.

The newsletter is not intended to constitute legal advice. Special legal advice should be taken before acting on any of the topics addressed here.

For further information, please contact the attorney listed below. General e-mail messages may be sent using LTBJ@lungtin.com which also can be found at www.lungtin.com

ZHOU, Jun, Patent analyst, Former patent examiner: LTBJ@lungtin.com

ZHOU, Jun
(Patent analyst, Former patent examiner)

Mr. Zhou has rich knowledge and experience in the aspects of Patent Landscaping and patent portfolio strategy. Prior to joining in CNPAT LungTin IP Value Inc.,Mr Zhou served as a patent examiner of China National Intellectual Property Administration, and has one year work experience in the Intellectual Property Development and Research Center of CNIPA. He has participated in a number of patent early warning and patent navigation projects which assist CNIPA and governments in making IP or industrial decisions, and presided several competitive intelligence analysis projects in the field of new energy vehicles and Intelligent Connected Vehicle.

18th Floor, Tower B, Grand Place, No 5, Huizhong Road, Chaoyang District, Beijing 100101,P.R. China
Tel: 0086-10-84891188   Fax: 0086-10-84891189
LTBJ@lungtin.com   www.lungtin.com