

Technology helps users get more from patent information



Pierre Buffet, executive vice president of Questel, shares his thoughts on what patent information professionals want and how their needs can be met

What do patent information users need?

A patent is the link between R&D and the legal processes. This defines the players: researchers, IP professionals or patent information professionals, and the pure legal people who need patent information to use in court. We started by serving the IP experts but we have now broadened in both directions.

For researchers, patent information is a big repository of technical information that is often not available elsewhere. They use patent information for ideas, to avoid 'reinventing the wheel' and for current awareness. Alerting services are very important for them and they want to access it in a simple, Google-like way.

IP professionals need comprehensiveness. As key patent information professionals retire there are fewer and fewer in the 'top expert' category. We need to find ways to assist newcomers to do high-quality jobs.

What trends have you seen in the ways that patent information is searched and used?

Users now have access to the real information, which is the patent document. In the past, people tried to describe the document. Thanks to the arrival of full text, people are trying to get more information from their searches and asking detailed questions that no key word search could address. Sometimes people think that searching the full text is the only solution but that is not true. Searching full text without being flooded by many false drops and unreadable text remains a challenge.

Professionals ask for more searchable information and more added value, especially through integration and post processing. Time

to handle information is critical, as everyone looks for an ever-better return on investment (ROI). Sometimes I say that ROI is king, a French joke as 'roi' is French for 'king'.

Information providers can assist users in this by, for example, linking images and text and providing patents viewers. There is still room for improvement, though, with developments such as assisted reading, natural language processing and better ontologies. Information providers can also help by providing non-patent literature (including industrial designs), easier access and navigation and tailored solutions for different types of users. There is also the growing importance on searching patents in non Latin-based languages.

Which languages are users most interested in?

The priorities in language are not secret. People really want to access patents in Asian languages. German is still important too. Spanish-language patents are a fairly small portion of the total but they are of political interest, because of Mexico and the USA in addition to Spain. Latin American countries are not yet at the same stage of patenting activity as China, but Brazil could make Portuguese an important language for patents too. Russian is also becoming important again after a long period where patenting activity was very quiet after the collapse of the USSR.

There are particular challenges with translating Asian patents. In Chinese, for example, the concept of the sentence does not exist. We have to watch and see how Boolean searches will handle this. Trademarks in Chinese are very difficult

because both the Chinese characters and the Latin text need to be protected.

Machine-assisted translation is still quite poor but it helps, especially with Japanese patents where the machine-assisted systems are now very sophisticated. Two years ago almost nothing was translated from Chinese patents and now there are English-language abstracts of patents and utility models. Machine translation is not the answer but it is a way to the answer. Statistics-based machine-translation tools now lead the way and are much cheaper than previous approaches. Google helped make machine translation mass market by using such tools to translate Chinese information during the Olympic Games.

People say that cross-language searching is not good but I am not totally in agreement with this if you take care about how you search. People will have to use a combination of both routes, machine translation and cross-language searching. We've implemented this on our platform. It's not perfect but it's a start.

What will happen with patents and patent information?

I am not afraid of a lack of business activity with patents. Obviously the recession will stop some research projects but it is well known that, in a recession, you shouldn't reduce your R&D because that is the business of the future. The R&D cycle is slightly longer than the manufacturing cycle. The large volumes of patents will continue to expand although there is a big question about how long the Chinese growth in patent filings, of 30 per cent per annum, will continue.

There will be a confirmation of today's trends with patent information and its provision. What has changed is that what is affordable now wasn't in the past. I remember working with machine translation in 1971 but if you wanted to apply it to text you needed to pay a lot for the computers. Now you can do it on your desktop. My laptop today is more powerful than Questel's first mainframe.

Interview by Siân Harris