EPO-PatCom meeting
Tuesday, 22 March 2011, 10.00 hrs
Room 348, EPO Vienna

Summary of the discussions

1. WELCOME

The EPO introduced Heiko Wongel, the new director for Patent Information Promotion and Training.

PatCom thanked the EPO for hosting the meeting, stressing that it and the XML day the following day were strongly supported and appreciated.

2. MINUTES OF THE LAST MEETING

Since the previous meeting, the President no longer requires a copy of the minutes to be distributed to the TOSC. As a result, the EPO expected to be able to finalise the minutes faster in future.

The minutes of the meeting of 21 September 2010 were approved.

3. FOLLOW-UP FROM PREVIOUS MEETINGS

3.1. AUTOMATION PLAN – LATEST STATUS

The previous practice of sending PatCom an excerpt of the EPO's Automation Plan has now been abandoned because the Office no longer produces an Automation Plan. This point will be removed from the agenda of future meetings.

A brief discussion took place on the hearing on the IT study commissioned by the Office, scheduled for 31 March in The Hague. The EPO said that one positive outcome of the study was that it praised the efforts currently made in the patent information area and confirmed that patent information was not a nice-to-have, but a must. PatCom hoped the study showed that EPO could not act alone and the all partners had to act together on the basis of standards, co-operation and communication.

3.2. EUROPEAN MACHINE-TRANSLATION INITIATIVE

The EPO described its on-going discussions with Google related to machine translation (MT):

- the Administrative Council and TOSC have said that there is a need to overcome the language barrier in patent information, and also to look for a technical solution that would ease the discussions on the unitary (EU) patent.
- the Office signed a memorandum of understanding with Google in October 2010. This is not an agreement as such, but it does lay down that any solution will be based on non-exclusivity, and that it will be free of charge, easily accessible and stable over a number of years.
- the memorandum expresses the hope that an agreement will be signed in the coming months. The EPO's "horizon" for this is the Administrative Council meeting of March 2011.

Responding to questions from PatCom, the EPO confirmed the following:
- the discussions with Google cover MT for the full scope of patent documentation, including backfile, frontfile and all languages.
- the task is to find an MT engine fit for the purpose of helping persons skilled in the art to decide whether or not they need a manual translation. The mandate is that the EPO should not build a MT engine of its own, but to make use of what is available on the market.
- the Office has not involved WIPO in its discussions with Google. The EPO wants to start its machine translation efforts with European languages, and later Asian languages. WIPO's priorities are likely to be different.
- there is no direct link to the KIPO-led IP5 foundation project on machine translation, but the EPO is certainly aware that it should avoid any duplication of efforts.
- the discussions include many complex issues, especially in the area of data protection

On the issue of exclusivity, the EPO stressed that had opened discussions with Google because it was looking for a place to start that was most likely to get a result for all languages. Other suppliers/providers interested in making a proposal based on the set of conditions laid down for the co-operation with Google were welcome to do so.

3.3. USER SURVEY RESULTS

The EPO reported that it had now received the detailed observations of the market survey company that had carried out the patent information survey on its behalf across member states in 2010. It was now in the process of finalising a report, to be published on the website as a PDF document. EPO will own the data, so deeper analysis will be possible.

3.4. VISUALISATION TOOL DEVELOPMENT

The EPO explained that its visualisation tool had been integrated into its "Patent information services for experts", divided into a public part (with stored data and graphics) and a password part where users can do more complex analysis.

The PRS (legal status) database has a new format so that it can be linked to PATSTAT. The introduction of a unique DOCDB-id will improve data portability in this regard.

A patent analytics workshop will take place in November, including a demo of some commercial products for more advanced patent statistics.

3.5. PATSTAT USAGE

PATSTAT currently had 150 users, the EPO said, the main users being statistical bodies like OECD, Eurostat and large universities.

3.6. BNS EXTRACTION

The EPO confirmed its announcement at last meeting that it wanted to offer data extraction from BNS. Some issues remained due to huge data volume. The pricing would be based on a minimum charge for one collection, plus a price per GB. PatCom repeated that this would be a very useful service.

3.7. STATUS OF XML ST 36 SCHEME

Work was continuing on an XML ST. 36 schema to replace the current DTD. The EPO said that no date for implementation could be announced yet, and promised to report well in advance so that users would have no problems with the transition.

The EPO has also been monitoring WIPO attempts to come to a way of describing ALL IP rights in one standard.

4. IP5 - PROJECT DEVELOPMENT

With regard to the Common Hybrid Classification project of the IP5 offices, the EPO said that there was a meeting of the IP5 Working Group 1 taking place in China in the same week as the present meeting. So far the CHC project is proceeding according to plan and five of the six pilot projects have already been transmitted to WIPO for introduction into the IPC in the next possible version. The WG1 meeting is expected to produce a clearer picture of the medium term roadmap for the Common Hybrid Classification project.

The EPO also drew PatCom's attention to the article in Patent Information News 1/2011 describing the "Cooperative Patent Classification" agreement with the USPTO. It gave a short presentation with further details.

5. QUALITY OF DATA SUPPLIED BY EPO

The EPO reported on the following developments:

DOCDB/XML - New and upcoming features

- **Master Classification Database - IPC/2006**
"Rolled-up" IPC core-level classifications have been removed from the Master Classification Database. All related changes have been exchanged in the DOCDB/XML backfile of January 2011.

- **Unique and stable identifier for application-identifier - Q1/2011**
The new "doc-id" attribute on the application-reference provides a unique and stable identifier that will reduce the impact of corrections to application-identifiers (so-called "re-keys"). In the medium to long term this unique and stable identifier will be introduced to other EPO raw data products, including the legal status database. This will make it easier to connect raw data products in a reliable way.

- **US re-classifications, FI/Fterm and PAJ - Q3/2011**
A number of interesting new features such as US re-classifications, Japanese FI and Fterm and Patent Abstracts of Japan, will be added to DOC-DB. Target date: mid 2011.

- **In the pipeline ...**
Considerable enhancements to cited references that will bring improved transparency and enriched content for both patent citations and NPL citations. The EPO is also investigating the inclusion of applicants' and inventors' address data for a number of countries.

INPADOC worldwide legal status

- **Achievements since September 2010**
  - Brazil - scope extended with PCT entries into the national phase
  - Estonia - additional legal events, data extended back to 2003
  - Slovakia - legal status data introduced
  - Columbia - new LATIPAT country
  - US - extension of legal status with additional events related to renewal fees.

- **Codesheets**
The codesheet collection (explanatory fact sheets for each legal status event) now includes all legal status codes for the US complete with an exhaustive description of the US patent procedure. The codesheets for France are about to be delivered, and are in preparation for Russia and Taiwan.

- **Medium and long term**
  - Change of format for Spain, France and the Netherlands. All provide richer information in their XML data feeds.
  - Extension of the legal status for Poland with additional events
  - Extension of the legal status for Slovakia with additional events

With regard to **LATIPAT** data, the EPO confirmed it will be available as part of the raw data offered by the EPO. The exception may be full-text data (insofar as the EPO has it), where it will probably be necessary to contact the national patent office that owns the data.

The EPO said it would give more details during the XML/raw data event the following day on the attributes of the "**unique document identifier**", which would connect all data in the various databases, similar to the family ID, and would identify the patent application.

6. ESPACENET DEVELOPMENTS

The new Espacenet GUI was launched on 28 February 2011, and is strongly influenced by the design of the EPO website. Apart from the cosmetic changes, it includes:
- RSS feeds from searches
- adaptation of ECLA navigation and visualisation to GUI
- better citation coverage
- export function to MS Excel
- query history function
- improved navigation and document download.

Future developments for late 2011 are:
- improved ECLA navigation
- screen customisation
- implementation of the common citation document

The Office was also looking into improving full text searching and offering better help for the user (such as thesauri, pop help or Google-like prompts along the lines of "Do you mean to search for ... ?")

Usage of Espacenet was stable at 25 000 to 30 000 users, 190GB and 1.1 million PDF views, all per day.

PatCom said it was interest in knowing where the majority of Espacenet users were. The EPO replied that it had established there was a large proportion (approaching 50%) in Asia, and that it intended publishing news items on this.

7.   EUROPEAN PATENT REGISTER

The problem for users, explained the EPO, was that getting legal status information on EP patents, once granted, was complicated. A first step would be to make a live link from European Patent Register to national registers. This was still in the conceptual phase, but the EPO certainly hoped to make progress on this in the future.

8.   OPS DEVELOPMENTS

Since its launch in 2003, OPS has used the SOAP web service protocol. However, from March 2011, the EPO will offer the current version of OPS (version 2.6.2) as a fully REST-based service. Existing services will continue to support the old SOAP protocol, but this will not be the case for any new services that come on line.

Two new services available immediately in OPS RESTful only are:
* ECLA: European classification scheme and descriptions
* European Patent Register data in XML

In the next months the efforts will be on consolidating the RESTful interface and improving the access control to the service. Additionally, resources will be dedicated to offering Web service access to services available in Espacenet but currently not yet in OPS, such as classification search and a more structured source of non-patent literature.

Usage data shows a stable increase in number of users of OPS v.2 (an average of 2 000 to 2 500 individual IPs served weekly). The number of total requests has increased in 2011 to an average of over 4 million per day, corresponding to over 150 GB per day of data transferred.

9.   PUBLICATION SERVER DEVELOPMENTS

The EPO reported that the publication server was ready to process sequence listings. It also planned a corresponding raw data product for sequence listings.

The publication server, it said, was also a model for other patent offices. Patent Information News 1/2011 listed the countries that had adopted the EPO's technology for their own publication server.

10.   ANY OTHER BUSINESS

10.1.   WIPO FRONTFILE DATA

PatCom reported that WIPO no longer provided its data electronically, but only on DVD, the explanation being that WIPO had bandwidth issues, so commercial providers had to wait until their DVD arrived. This gave advantages to providers, depending on postal delivery times. The EPO agreed to pick up this point with WIPO.

10.2.   FREQUENCY OF MEETINGS

The EPO invited PatCom to reflect in time for next meeting on whether the current practice of six-monthly meetings was still the best frequency for their needs.

10.3.   DATE OF THE NEXT MEETING

It was provisionally agreed to hold the next meeting on 20 September 2011.