

EPO-PatCom meeting
Tuesday, 27 March 2012, 10.00 hrs
Room 348, EPO Vienna

Summary of the discussions

1. MINUTES OF THE LAST MEETING

The minutes of the meeting of 20 September 2011 were approved.

2. FOLLOW-UP FROM PREVIOUS MEETINGS

2.1. WIPO - FRONTFILE DATA

PatCom reported that there seemed to be some progress with WIPO on the availability of their data. The EPO agreed, confirming there would be a paper concerning data on the entry/non-entry into the national phase at the forthcoming CWS meeting. Of the eleven countries for which the EPO does not receive entry/non-entry data due to agreements with WIPO, the EPO has been successful in obtaining permission directly from AP, MX, IL and UA to include their data in the legal status database.

PatCom agreed that this issue was primarily one for discussion with WIPO than with the EPO. For its part, the EPO said it would continue its efforts to obtain the data.

Currently, out of 61 patent authorities contributing to the EPO's legal status database, 43 deliver information about the entry into the national (regional) phase. Taking into account that several member states of the European and the Eurasian Patent Offices allow only regional designations, the number of countries covered is over 70.

For its legal status database the EPO takes data direct from national offices and WIPO. At the moment the country distribution according to the source of data is:

provided via WIPO and by the national office	18
provided via WIPO only	23
provided by national office only	2

This information is available in the 'Contents and coverage of the INPADOC legal status file' table on the EPO website, updated weekly (see www.epo.org/searching/essentials/data/tables.html)

2.2. EUROPEAN MACHINE-TRANSLATION INITIATIVE

The EPO described the status of its project with Google related to machine translation (MT):

- Patent Translate was launched in Espacenet and the Publication Server on 29.02.2012.
- The service provides translations between English and the following languages: ES, FR, DE, PT, IT, and SE. Further languages will follow. In some cases, where there is not a sufficient collection of patent texts in a particular language, the Office will work with other organisations, eg scientific institutes in order to collect more material for training the translation engine.
- Quality is better than before, based on tests using widely applied methodologies, but of course does not correspond to a legal translation.
- Non-European languages may also be added.

In relation to security issues, the EPO informed PatCom that when requesting a translation via Espacenet or the publication server, users' IP addresses would not be visible to Google, as the EPO gateway would act as the only visible "user" requesting the translation from the Google engine.

Responding to a question from PatCom, the EPO said it favoured the current solution of on-the-fly translation over bulk data translation. Since translation quality was not optimal yet, there seemed little point in trying to capture a snapshot a specific moment in time.

2.3. USER SURVEY RESULTS

The EPO said it was now ready to release the results of its recent survey into the use of patent information in Europe, including the details of the questionnaire and the raw data of the responses.

Some of the main findings were:

- awareness of patent information fell (albeit during financial crisis);
- the demand for the use of patent information for business decision making purposes was validated;
- a significant barrier was now the need to understand, make sense of and use patent information.

On the final point, the EPO explained that it was not its role to advise companies on strategy issues. It could, however, provide some "train the trainer" services and encourage the development of specialists in national patent offices and PATLIB centres to serve local needs around Europe. In a similar vein, it would be working on some simple visualisation tools to

stimulate interest in patent analysis, but there was no intention to become a software house. The EPO's task was to supply data, not to interpret it.

On the topic of improving understanding of the user landscape, PatCom agreed to assist the EPO in putting together a broad overview of users. Contact persons are G. Schultheiss (PatCom) and V. Gray (EPO).

2.4. VISUALISATION TOOL DEVELOPMENT

The EPO reported that the PATSTAT visualisation tool was now available as part of the "Patent information services for experts", in the form of a public part with graphics from pre-run queries and a password-protected part where users could do more complex analysis. The online hosted database provided for basic visualisation of outcomes. Users who wanted to do more could import subsets of the data into business intelligence tools for enhanced visualisation. The EPO said it was working together with KU Leuven in Belgium on how to deal with spelling variants of applicants' and assignees' names, and had achieved some improvements, allowing for better statistical outcomes.

PATCOM members who have created import filters to use PATSTAT data with their tools were invited to publish these for the user community on the PATSTAT forum: <http://forums.epo.org/epo-worldwide-patent-statistical-database/>

PatCom members interested in obtaining the EEE-PPAT table on harmonised applicant names, as well as the papers describing the underlying methodologies, should send an e-mail stating the nature of their request to TechnoInfo@ecoom.be

2.5. BNS EXTRACTION

Efforts to extract data from BNS had been successful for small extractions, but larger data quantities were problematic, explained the EPO. Thus, it was not possible to accept more orders at the current time. The EPO hoped, however, to be able to present progress at the Patent Information Conference in Hamburg. PatCom stressed that its members had considerable interest in a solution to this issue.

2.6. COOPERATIVE PATENT CLASSIFICATION (CPC)

There was significant discussion on the technical aspects of the implementation of the CPC, planned for 1 January 2013. The EPO explained that it had already invested a lot into cleaning up the ECLA scheme to get it ready for conversion to CPC, and the process was still going on. It was converting ICO symbols into ECLA, and taking over large proportions of internal keyword schemes - it had already created some 600 definitions for main technical fields and instructions on how to classify.

Current planning was as follows:

- 15 August 2012 - first version of CPC database, shared with USPTO
- autumn 2012 - ECLA+ICO-to-CPC correspondence table
- autumn 2012 - publication of the 600 definitions
- November 2012 - DOCDB available with CPC symbols
- December 2012 - CPC live on Espacenet

The EPO asked PatCom to give an opinion on the preferred format for the ECLA+ICO-to-CPC correspondence table it would prepare. PatCom said that probably XML would be best, but that they would take some time to consult internally and report back. They said that it was important to have the data as soon as possible.

The EPO emphasised that the correspondence table would be a one-off product, created at the moment of the switch from ECLA to CPC, and that it would not be maintained once CPC revisions started to occur.

In response to a question from PatCom, it was clarified that the EPO would almost certainly not display the CPC symbols on published patent documents. It was not clear, however, how the USPTO would proceed - they had a legal obligation to make the CPC symbols for a particular patent document available at the time of publication, but had not decided yet how to do this. For non-EP/US documents, the CPC can only be assigned after publication as the EPO has no access to the document before that.

PatCom asked if the ECLA symbols assigned today would be lost. There would be a record of old classes, said the EPO, but these would not be made available externally.

PatCom said they also urgently needed an updated DOCDB manual. The EPO said it was working on this. It could already confirm that there would be no schema change involved.

2.7. PUBLICATION SERVER

PatCom complimented the EPO on the Publication Server, saying it serves as an example to other offices.

2.8. EPOQUE

EPOQUE had been on the programme of the Patent Information Conference in Kilkenny. PatCom said they wondered where the discussions on the topic were going and said they were generating a certain sense of nervousness among its members.

The EPO said that the presentation in Kilkenny was in the interests of transparency and noted that EPOQUE had also been discussed at the IT hearing in March 2011, in The Hague, where PatCom was present. One

clear message coming out of these discussions was that EPOQUE might be a good tool for patentability searching but it was not the panacea to solve all the patent searcher's problems - this had been confirmed in a letter from the PDG to the Office.

It was important, said the EPO, to discuss what people understood under the word "EPOQUE". As a search tool, it was comparable to products on the market. The data was what makes EPOQUE special, but the EPO would most probably never be allowed to make, for example, commercial databases or copyrighted non-patent literature fully available externally..

PatCom summed up their view of the situation, saying that in releasing EPOQUE, the EPO would be perceived as acting like an online host. This would entail a major commitment, both technically (system maintenance and bandwidth), and in terms of user support.

3. STATUS REPORT ON THE UNITARY PATENT

The EPO reported that this project had been delayed until at least 1 January 2014. It confirmed that the EPO would be in charge of administrating European patents with a unitary effect and that it would be discussing how to present bibliographic information and other technical questions. It would make every effort to pass on information (DTDs, format changes, etc) as early as possible to PatCom members.

4. QUALITY OF DATA SUPPLIED BY EPO

PatCom protested at the twenty-fold increase in the data quantities in recent updates, saying they had received almost no warning and that the increase in data was causing disruption to their services.

The EPO explained that main reason for the increase was the ECLA clean-up process, getting EPO data ready for the CPC, and that this was likely to continue - and possibly get even more extreme - through the rest of 2012. The magnitude of the changes to the classification scheme had not been foreseen by the EPO, thus it had not been possible to issue a warning. The high data volumes were likely to continue in 2013 due to the CPC revisions after the launch. It was not yet decided how to handle these revisions, and what their frequency should be.

In some areas of the CPC, the USPTO did not feel that the ECLA structure adequately reflected what they needed, so there may be some big changes in the pipeline, for example in the area of business methods. Details will follow later.

PatCom said that it also needed advance information about smaller changes to the databases, citing the recent PRS code change for US A1 documents as a case in point. The change had caused problems for

providers that could have been avoided with more active communication from the EPO.

The EPO reported on the following developments on the databases:

DOCDB/XML

"Mega" re-key of JP publication numbers

Q4/2012 - the rekey will take place to synchronise the DOCDB formats for JP publication numbers with the number formats as supported by the JPO, and it will concern the complete JP collection before 2000. The change will make it possible to load pre-2000 JP cited references. It affects 11 million JP publications proper plus 12 million publications citing a JP document. In all, this means there will be 34 million exchange records in the raw data product. There will be a full DOCDB backfile release in February 2013 with these changes.

Citations

March 2012 - the launch of a procedure to extend the collection with "pre-grant" citations for US-A1, adding approx. 3000 new citations a day to the collection resulting in 10 000 to 20 000 cited references. The collection will also be extended to include citations filed by the opponent in the EP-B1 records.

INPADOC worldwide legal status

Japanese legal status data

2012 - Legal status data from Japan is now available. The backfile to April 2011 will be complete by summer 2012, and then the EPO will start loading further backfile data, extending back to 2003.

NL legal status data

After a hiatus of approximately two years, the capturing of NL legal status data will soon be resumed.

New countries

Argentina, Cuba Columbia and El Salvador data is now available. Legal status data for German utility models is also included.

Full text data in ST.36

Q2/2012 - The EPO expects to offer full-text data for GB and FR. AT and CH are likely to follow.

5. ESPACENET DEVELOPMENTS

The EPO reported that it was continuing to fine-tune the new GUI, introducing a more compact bibliographic screen and improved navigation throughout the application. Further changes are still planned, especially

with a view to making Espacenet compliant with generally accepted accessibility criteria.

Additionally, the "Patent translate" feature based on a partnership agreement with Google was recently released in February.

Developments planned for 2012 include:

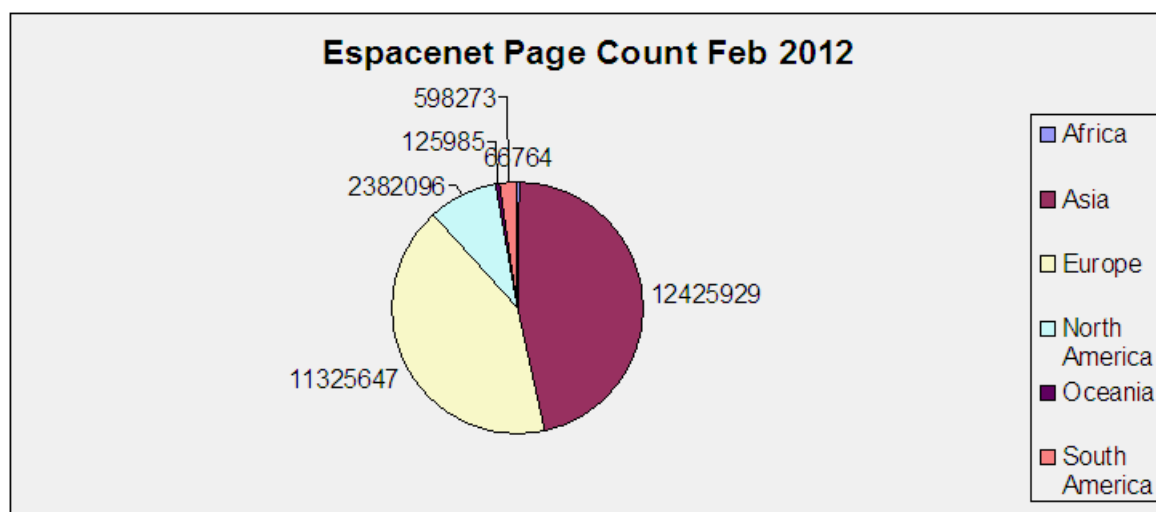
- migrating from ECLA to the CPC
- integrating features from the Common Citation Document application into Espacenet
- enhancing search functionality
- enhancing viewing

The enhancements relating to searching and viewing will take place as part of the IT roadmap over the next three years. The option of including some semantic search possibilities is also under investigation.

Usage statistics (February 2012)

- 1.6 million separate visits
- 550 000 unique IP addresses
- 48 million page views

The EPO said the statistics showed little change to the geographical user spread, with the largest proportion of users coming from Asia, then Europe.



6. OPS DEVELOPMENTS

OPS has progressively moved from using the SOAP web service protocol to RESTful. Legacy SOAP-based OPS services are still supported,

however the recently introduced OPS Register and ECLA services as well as any new services that come on line will only be available as REST.

The plans for the year 2012 include the following activities:

- enabling access to ES fulltext data
- migrating from ECLA to the CPC
- providing access to file inspection data as available under the "All documents" tab in the European Patent Register

Usage statistics

- stable at 3 to 4 million requests per day, amounting to just under 100 GB
- weekly average of 4000 IP addresses
- geographical spread: about 60% in Europe (first two months of 2012), then Asia, then a small amount US and Oceania.

The EPO reported that it had observed about 0.5% of users were taking 75% of the data. In the spirit of fair treatment of all subscribers, it was necessary to think about how to uphold equal and fair treatment. The EPO said it would prepare concrete proposals that would probably distinguish between high volume and low volume usage.

7. ANY OTHER BUSINESS

7.1. UPDATES TO STANDARD LICENCE AGREEMENTS

Due to the evolution over the years, some older licence agreements are out of date and no longer reflect the current situation. The EPO said it would therefore be contacting providers with these older licence agreements in order to switch them to the newest text. It stressed that this procedure would not involve any change to the substance or spirit of establish practice.

7.2. PATLIB2012

PatCom members were reminded about the PATLIB2012 conference and exhibition which will take place in Manchester from 30 May to 1 June 2012. PATLIB features an exhibition only every two years, on even years. www.epo.org/patlib

7.3. DATE OF THE NEXT MEETING

It was provisionally agreed to hold the next meeting between 5 and 8 November 2012, during the EPO Patent Information Conference in Hamburg.

EPO-PatCom meeting
Tuesday, 27 March 2012, 10.00 hrs
Room 348, EPO Vienna

List of participants

PatCom

President of PatCom
Secretary of PatCom

Willem-Geert Lagemaat, Lighthouse IP
Georg Schultheiss, formerly FIZ-Karlsruhe

Ann Chapman, Minesoft
Jurjen Dijkstra, LexisNexis Univentio
Cinda Harrold, CAS
Armin Förderer, FIZ Karlsruhe
Paul Peters, CAS
Rob Willows, Thomson Reuters

EPO

Director Publication
Director Promotion
Director Special Services
Director Data Resources

Pierre Avédikian
Heiko Wongel
Günther Vacek
Miguel Albrecht

Nigel Clarke
Valérie Gray
Davide Lingua
Johannes Schaaf
Daniel Shalloe