

**36th EPO-PatCom meeting
Friday, 12 October 2018, 11.00 hrs
Room 348, EPO Vienna**

Summary of the discussions

1. INTRODUCTION

The chairman said the EPO appreciated open and lively exchange at the regular meetings with PatCom. He noted that the new EPO President placed great value on feedback from users.

PatCom's President thanked the EPO for the ongoing dialogue with the EPO, saying it was a benchmark compared with other partners.

2. APPROVAL OF THE AGENDA

The agenda for the 36th meeting was approved with a new point "Espacenet" and the addition of two points under AOB:

- Support for PATLIB centres
- Non-patent literature

3. MINUTES OF THE LAST MEETING

The minutes of the meeting of 20 March 2018 were approved.

4. ESPACENET

The Office reported that it was in the final phase of internal testing for the new beta version of Espacenet and planned to present it at the EPO Patent Information Conference in Brussels. After that, a period of public beta testing, with volunteers ranging from students to private searchers, would follow. Initially "new" Espacenet would not replace "old" Espacenet, which would remain in operation for some months after the full release. One of the main features would be the replacement of the "advanced search" fields with a more flexible query builder. There would also be some filtering options, and the interface would be in responsive design.

PatCom asked if it would still be possible to link direct to the legal status information. The Office replied that it would be making sure that current links to documents continue to work, but would need to check about other data, such as legal status. The Office promised to investigate and report back.

The Office confirmed that there would continue to be a limitation on the number of search terms, but it would be different to those in operation today.

The Office also confirmed that the data coverage would remain the same with Espacenet today, noting, however, that the backend data would be managed in a different way.

Responding to a question on the EPO's policy with regard to commercial providers linking direct to Espacenet, the Office said that its goal was to block bulk data downloads (OPS was available for retrieving data in bulk). It was a question of keeping the load on the system to a level that avoided problems. The basic principle was to ensure that users could rely on the system for retrieving single documents. For the retrieval of single documents, it was also acceptable if commercial providers linked to Espacenet.

5. WORLDWIDE LEGAL STATUS DATA IN XML

The Office reported that the number of legal events in the INPADOC database had reached 250 the previous month. A lot of work was continuing in terms of adding SPC and paediatric extension data, notably for ES, but also other European countries (BE, IE, LU, NO, SE, SI).

The JPO has announced that it will change data delivery format in April 2019. The EPO said it was expecting sample data soon. The advantage of the new format will be that data becomes available two weeks faster than currently.

Responding to information from the Office that it had not uploaded data from the USPTO because the API had not been kept up to date, PatCom remarked that the API seemed to have improved of late. The Office agreed to check, but observed that updates at unknown, irregular intervals are not sufficient to guarantee the provision of reliable data.

With respect to OPS, the Office confirmed that OPS did not currently offer action to file inspection ("file wrapper") data, but that it hoped to be able to do this in the future. PatCom said they would very much appreciate such a development.

Turning to data on the entry of PCT applications into the national/regional phase, the Office reported a new backfile from WIPO meant that more countries were now available. The Office would continue to work on the remaining gaps, and on receiving more frontfile data.

Finally, the Office reported that it had introduced categorisation into INPADOC in June – all legal status codes used in the past 20 years had been categorised more or less in accordance with WIPO ST.27. The Office had also circulated a questionnaire on the more detailed

classification of legal status codes and was waiting to see what feedback it yielded before deciding how to proceed.

6. OPS (WITH POINT 8.: QUALITY AT SOURCE)

OPS had been largely stable since beginning of 2018, the Office said. The new flat-fee policy for using OPS seemed to be working well, with very few problems or complaints. A new release the previous week had not been widely announced because it contained no new features, just some additional authorities for full text (BE, CY, HR, IE, LU, MC, MD, NO, PT, PL, RO). With respect to the backfile for these countries, the Office said that it was expected soon for some of them, others would take longer. The corresponding information was available on the "statistics for full text" information in Espacenet. It expected the backfile to be in the form of OCR'ed scans; the frontfile should largely be high-quality XML, produced by the NPO in question.

7. DOCDB QUALITY

The Office reported that in DOCDB there was currently a tendency to fewer format changes; what changes there were, were announced in good time (at least four weeks in advance). The only change due in the near future was the introduction of "CPC International" expected for April 2019. CPC International related to expressing the classification symbols assigned by offices other than the EPO and USPTO; the main difference to CPC NO was that CPC NO was publication bound, whereas CPC International was family-based system. As soon as sample data became available, the Office would share it with PatCom.

Responding to a question from PatCom, the Office said it had no progress to report on CNIPA adding CPC information on Chinese utility models.

Returning to classification data, PatCom remarked that it was difficult for users to know which code to use for a code that had become invalid – a concordance table would be helpful, it said, noting that the information on the CPC website was quite complicated.

Finally, the Office reminded all PatCom members of their obligations under the EPO's terms and conditions on the distribution of bulk data.

8. QUALITY AT SOURCE (QAS) PROJECT UPDATE, IN PARTICULAR

Dealt with under agenda point 6 above.

9. LEGAL STATUS INFORMATION PRODUCTS

Reporting on the European Patent Register, the Office reminded PatCom– as already announced in March – that documents relating to the PCT procedure where the EPO is the ISA would be available.

For the Federated Register, 30 countries participating were now participating, including DE and GB. The remaining major country, FR, was making good progress on the necessary technical changes. Deep linking was available for 33 countries, including some extension and validation states. The Office also mentioned that it was beginning to discuss the concept of bi-directional registers with member state patent offices.

Switching to Global Dossier, the Office informed PatCom that it had shifted focus from extending the system to ever more countries to ensuring the reliability of the data already available. For the backfile from China, there was an issue associated with a change that took place in 2010, rendering it impossible to incorporate data from before that date. Latest infomationinformation from CNIPA seems to indicate that this will not change in the near future.

Finally, PatCom thanked the Office for the interesting article in PIN on European patents being valid beyond Europe.

10. PATENT TRANSLATE

Introducing this topic, the chairman remarked that the moved to neural machine translation had led to a major improvement in the quality of translations in Patent Translate, especially for Asian languages and less widely spoken European languages. User statistics showed a peak on Wednesdays, especially in the US, with the most frequent language pair being DE/EN; currently about 20 000 translations per day were being requested.

The Office confirmed that it-commercial providers were free to use the Patent Translate engine via Google, but that it was contractually prevented from sharing its language corpora.

The Office also mentioned that it had stored translated full-text for use by examiners and it expected Espacenet would also use this data set at some point in time. PatCom stated that if a body of English full text became available, PatCom members would be interested in acquiring it.

PatCom reiterated its standpoint that it gave Google an unfair advantage if the EPO was not prepared to share the same data with everyone that it shared with Google. The Office repeated that it was bound by its contractual conditions.

11. BIG DATA

The Office reported on the conclusions reached by the Fraunhofer Institute in a study on big data that it had recently done for the Office. The main conclusions were as follows:

- the way data is generally offered is for known experts, but not for data scientists outside the normal user sphere;
- it is important to offer a good resource map, showing at a glance what is available;
- the Office should provide a sandbox and documentation allowing people to play around with the data;
- the Office should give consideration to how to benefit from big data architecture, i.e. making EPO data available via third-party channels

The Office also reminded PatCom that its linked open data product was now in full operation (the service from the EPO being unique); up to now, there seemed to be a lot of interest in the CPC data.

PatCom was invited to give its view on linked data: What are the required formats? What could the EPO do to make the linked data service as useful as possible?

12. NAME STANDARDISATION PROJECT

In the area of name standardisation, PatCom remarked that INPI in France was managing to quote the French company number on some 90% of patent documents. Japan and Korea also publish company numbers on their patent documents. The EPO thanked PatCom for this information and said it would try to initiate a debate on the topic within the EPO.

13. AOB

13.1. SUPPORT FOR PATLIB CENTRES

The Office explained that it had invested efforts in recent years in reinforcing the services of Europe's PATLIB centres. Many PATLIB centres struggled with a low budget and could not afford to use commercial patent information tools and services. The Office therefore invited PatCom and its members to consider offering special price packages to PATLIB centres. PatCom thanked the Office for the idea, saying it sounded like a good way forward. It would respond formally in due course.

13.2. NON-PATENT LITERATURE

Non-patent literature, said PatCom, sometimes appeared in multiple versions, but unlike patents, these multiple versions were not linked by

any kind of family structure. There also seemed to be a proliferation of cases where innovations were described in patents and academic articles, again without any formal link between the two. This situation could cause confusion in the overall citation landscape. The Office agreed to discuss the point with its documentation specialists.

14. DATE OF NEXT MEETING

It was agreed to hold the next meeting in Vienna on 19 March 2019

36th EPO-PatCom meeting

16 October 2018, 11.00 hrs, EPO Vienna, room 348

Participants

PatCom:

President	Ann Chapman, Minesoft
Secretary	Jane List, Extract Information
	Felix Coxwell, Questel
	Jurjen Dijkstra, LexisNexis Univentio
	Christiane Emmerich, FIZ Karlsruhe
	Robert Fokkema, Lighthouse IP
	Margit Höhne, PatentGate
	Paul Peters, CAS
	Lee Smith, RWS

EPO:

Principal Director Patent Information and European Patent Academy	Richard Flammer
Director Publication Director Promotion	Pierre Avédikian Heiko Wongel Davide Lingua Roland Feinäugle Daniel Shalloe