

CA/T 8/08

Orig.: en

Munich, 18.03.2008

SUBJECT: Report of the meeting between the EPO and the PatCom Group of
11 March 2008

SUBMITTED BY: President of the European Patent Office

ADDRESSEES: Working Party on Technical Information (for information)

The PatCom Group was established towards the end of 1999 and is an association of commercial companies that provide patent information products and services that include EPO data. It has been agreed to hold regular meetings with PatCom on a six-monthly basis. The meeting on 11 March 2008 was the 16th such meeting.

This document has been issued in English only.

TABLE OF CONTENTS

Subject	Page
I. INTRODUCTION	1
II. REPORT OF THE LAST MEETING	1
III. REFORMULATION OF EPO PATENT INFORMATION POLICY	1
IV. AUTOMATION PLAN	2
V. RAW DATA DAY (10 MARCH 2008)	2
VI. QUALITY OF DATA SUPPLIED BY EPO	3
VII. DATA, FORMATS AND MEDIA	3
VIII. DEVELOPMENTS OF <i>esp@cenet</i>	3
IX. DEVELOPMENT OF OPS	4
X. DEVELOPMENT OF PUBLICATION SERVER, OPEN WEB SERVICE INTERFACE	5
XI. ANY OTHER BUSINESS	5
A. CHINA	5
B. GOOGLE	6
C. PATENT INFORMATION SURVEY	7
XII. DATE OF THE NEXT MEETING	7

I. INTRODUCTION

The PatCom Group was established towards the end of 1999 and is an association of commercial companies that provide patent information products and services using EPO data. The EPO has agreed to hold regular meetings with PatCom on a six-monthly basis.

The 16th meeting between the PatCom Group and the EPO took place in Vienna on 11 March 2008. The present document is a report of that meeting.

The agenda for the meeting, chaired by the Vice-President of DG 4, was based on a list of points submitted in advance by PatCom.

PatCom was represented by:

Lighthouse IP (PatCom Presidency), Incom, Lexis-Nexis Univentio, Minesoft, Questel, Thomson Scientific, FIZ Karlsruhe, Matrixware, Prous Science, and CAS.

The chairman commenced the meeting by informing those present that an announcement was expected in the coming few days concerning the successor to Wolfgang Pilch in the post of Principal Director Patent Information. For the time being, Mr Pierre Avédikian would be performing that role on an *ad interim* basis.

II. REPORT OF THE LAST MEETING

PatCom confirmed their full agreement with the report of the previous meeting (CA/T 39/07).

III. REFORMULATION OF EPO PATENT INFORMATION POLICY

With regard to the reformulated patent information policy (CA/11/07 rev. 1), PatCom said they would submit their official position in writing.

Replying to a question on how the reformulation was being implemented, the EPO said that no changes had been implemented yet, but that there would be changes, hopefully driven by the "barrier free" concept described in the policy. The EPO was investigating possible avenues forward for making patent information more available, more easily available, easier to understand. It would be one of the main tasks of the new Principal Director Patent Information to put life into the reformulated policy. PatCom could rest assured that thorough discussions would take place before embarking on this. It was expected that the first consequences may take shape around the beginning of 2009, although all parties should

understand that the high level discussions within the Organisation ("raising the bar", "European Patent Network", etc) might influence how the priorities were set.

PatCom asked if any progress could be expected on the Community Patent issue under the Slovenian EU presidency. The EPO confirmed that the Slovenian (and forthcoming French) Presidency of the EU were active on this and on the EPLA issue, but said that the Community Patent would still be a difficult subject to find consensus on.

IV. AUTOMATION PLAN

PatCom confirmed receipt of the excerpt of the EPO's Automation Plan sent to them in December 2007.

V. RAW DATA DAY (10 MARCH 2008)

A short discussion took place on the technical seminar held for the commercial vendors the previous day. The EPO said it was pleased to have noted the attendance of real technical experts. It was also glad to have managed to attract a speaker from the Office for Harmonisation in Alicante, because it was working on passing the message of data quality across all players in the field. PatCom agreed that broadening the invitee list had been positive and encouraged the EPO to keep the idea of inviting other major IP offices or institutions such as WIPO.

PatCom thanked the EPO for organising the day and took the example of converting data systems to compliance with EPC 2000 as a demonstration of how these meetings were very helpful for avoiding misunderstandings.

PatCom added that data delivery remained an issue, and there was clearly still a bottle-neck on Thursdays. Sending data by DVD should remain an exception (once or twice a year). The EPO confirmed that the topic had been discussed the previous day and that it was looking to implement a solution for splitting the delivery dates for certain types of data. New data had to be made available as early as possible, but there was some flexibility with correction/amendment data. PatCom said they would support delivery of new data on Thursdays and correction data on Fridays. The EPO offered to distribute the correction/amendment part of the data on Mondays, i.e. three days earlier than it is currently done. Both parties agreed to continue studying the matter.

A short discussion took place on how to transmit the importance of data quality to all users of the patent system, especially applicants and other patent offices. PatCom stressed that they would help by putting the issue on the table wherever

they could. There was some consensus that it would help if efforts were undertaken to explain to people what the consequences were of poor data were.

VI. QUALITY OF DATA SUPPLIED BY EPO

PatCom praised the EPO for the way it was managing the move of its DOCDB database to XML 2.0. The period of parallel delivery of the old and new formats had worked very well. They thanked the EPO for the way it endeavoured to do a very professional job, both in terms of timeliness and in its response to issues. They said the EPO was ahead compared to other patent offices.

VII. DATA, FORMATS AND MEDIA

The EPO reported that the harmonisation process for its data was continuing, and that a milestone had been achieved in the integration of some data that had been separate from DOCDB, especially citations and NPL citations, made possible by harmonising to WIPO's ST 36 format.

Progress had also been made on converting the EP document from SGML to XML, the work on which was already complete back to 1992. It was planned to complete the collections by the end of 2008, at least for EP-A documents.

In response to a question from PatCom on the acquisition of missing Italian data, the EPO said that there was no progress to report. It acknowledged that the matter was a major issue and said it was continuing discussions with the Italian office.

VIII. DEVELOPMENTS OF *esp@cenet*

The EPO announced that the following enhancements were planned for *esp@cenet* (for implementation by mid-2008 unless indicated otherwise):

- examiner citations (not only EP/WO, but other authorities)
- export of lists (e.g. bibliographic data for result lists)
- increase limits: page limits / my patents (100 records)
- all publication steps available (A1, A2, A3, etc)

- date range search
- one-line, one-box search form (like Google), taking the form of a simple search or with search modifiers (e.g. IN for inventor name). There will be a limit on the number of terms, which will not exceed the existing limits on the full search mask, although users will be able to use "OR" between fields. The use of brackets will be supported.
- Late in 2008: full-text search for EP documents
- 2009 and beyond: links to national registers, sorting of result lists, enhancements to ECLA.

The machine translation feature in *esp@cenet* was also discussed, but the EPO said it had no information available on the introduction of further language pairs, and that there were no plans to provide machine translation of Asian languages. It confirmed, however, that its intention was to make the dictionaries for any new language pairs available as a raw data product as soon as they were ready. PatCom stressed that some commercial operators were holding back on developments pending the availability of the technical dictionaries and urged the EPO to provide more concrete information on timing.

IX. **DEVELOPMENT OF OPS**

The EPO outlined some of the technical improvements planned as a result of the Open Patent Services rebuild, which was well under way. It added that a public beta test version would be ready in March 2008 and that the production release was expected for May/June 2008. Versions 1 and 2 of OPS would be maintained in parallel for a six-month period to allow users time to convert their systems.

The new features will include the following:

- bibliographic search (in line with the *esp@cenet* command line functionality; under the same restrictions as in *esp@cenet*)
- delivery of the "equivalents" list (simple family)
- full text inquiry, to be able to check whether we have claims and/or description for a specific publication

- long term: web services will be extended to cover data from the European Patent Register, and the descriptions of ECLA classes; the EPO is also committed to extending full text coverage (currently only EP/WO) of first publications (i.e. "A" documents) to other European countries: DE, AT, BE, CH, FR, GB, and possibly CA and AU, depending on the agreement of the office concerned.

The EPO reported that the statistics showed a relatively stable use of OPS over the past six months (one million requests per day). The number of individual IP addresses was continuing to increase gradually, and was currently at about 1200 per week. Well over 50% of users were from Europe.

X. DEVELOPMENT OF PUBLICATION SERVER, OPEN WEB SERVICE INTERFACE

The EPO informed PatCom that the call for tender for hosting the Publication Server would be awarded to Jouve.

In terms of usage, it had detected an increase due to better awareness and to the increased visibility through the use of site-map protocols (SMPs), which meant that people browsing on the internet "hit" documents on the European Publication Server more often. Usage was running at about 3000 IP addresses per day, peaking on Wednesdays with new publications, with European users in the majority. SMP had only been implemented as a pilot, and it was too early to predict its impact.

The EPO reminded PatCom that the technology used for the Publication Server was available to member states. AT, FI, GB, IE and SI were either already using it or working on implementing it. PL had also shown interest.

XI. ANY OTHER BUSINESS

A. CHINA

Following up on information given at the EPO Patent Information Conference in Riga the previous October, the EPO said it had initiated a pilot operation to obtain manual translations of claims, titles and abstracts for Chinese patent documents. An analysis of the cost benefit aspects of such an operation resulted in opting for a new approach, aimed at populating the database with automatically translated abstracts and claims, with the option of ordering a manual translation when needed by the examiner.

The impetus for this strategy had been some progress on the co-operation with the Chinese Patent Office (SIPO), with the EPO acquiring full text data on CN patents. Similar data existed for utility models, so it should be possible to obtain a machine translation version of the complete text, albeit not up to the standard achieved for Korean and Japanese. Furthermore SIPO was working on a translation machine adapted to patent documents.

Whether or not it would be possible to sell full text data from China as an EPO raw data products was not clear yet. The EPO would discuss the matter with SIPO.

The EPO added that some commercial providers offered full-text searches in Chinese, which it would be testing. Further announcements were planned for the East meets West event in April 2008.

For the Chinese image collection, the EPO expected to have a full backlog for patents and utility models by end of 2008. As for all patent image collections, this data is available via *esp@cenet* and OPS.

B. GOOGLE

PatCom asked if there was information on two matters concerning Google:

- the EPO was trying to persuade Google to start loading EP data.
- Google activities in the area of machine translations

The EPO replied that there had been little progress since the last meeting. However, the introduction of SMP on the Publication Server (see point X above) meant that any search engine could index the data.

With regard to translations, the EPO remarked that it was interesting to observe what Google was doing, but there was nothing happening specific to patent translations as far as the EPO was aware.

C. PATENT INFORMATION SURVEY

The EPO informed PatCom that it planned to follow up its 2002-2003 survey in 2008. It had consulted with PatCom during the first survey and would propose to do the same again. EPO would therefore be in contact with PatCom when putting together the questionnaire. PatCom thanked the EPO.

XII. DATE OF THE NEXT MEETING

It was provisionally agreed to hold the next meeting on 26 September 2008 in Vienna.