

WIPO



SCP/14/3.

ORIGINAL: English

DATE: December 18, 2009

E

WORLD INTELLECTUAL PROPERTY ORGANIZATION
GENEVA

STANDING COMMITTEE ON THE LAW OF PATENTS

Fourteenth Session
Geneva, January 25 to 29, 2010

**TECHNICAL SOLUTIONS TO IMPROVE ACCESS TO,
AND DISSEMINATION OF, PATENT INFORMATION***

Document prepared by the Secretariat

* Comments received from Members and Observers of the SCP on this document are available at:
http://www.wipo.int/meetings/en/doc_details.jsp?doc_id=154139

Table of Contents

I.	INTRODUCTION.....	2
II.	CONCEPTUAL FRAMEWORK.....	2
III.	AVAILABILITY OF PATENT INFORMATION IN DIGITAL FORM	3
IV.	CLASSIFICATION.....	6
V.	LANGUAGES.....	7
VI.	DISSEMINATION OF PATENT INFORMATION	7
VII.	ONLINE DATABASE SERVICES	8
VIII.	HIGH-VOLUME PATENT INFORMATION PRODUCTS	10
IX.	SEARCH AND EXAMINATION REPORTS	10
X.	TECHNICAL TOOLS FOR ACCESS TO PATENT INFORMATION.....	11
XI.	OVERVIEW OF WIPO'S PATENTSCOPE® SEARCH SERVICE.....	12
XII.	POSSIBLE AREAS OF IMPROVEMENT	12
ANNEX	Selected Databases	
Figure 1	Availability of patent collections in paper and electronic form	3
Figure 2	Availability of full-text patent collections in search systems used by patent offices	7
Figure 3	Use of free and commercial patent databases by national office	8

I. INTRODUCTION

1. Pursuant to the decision of the Standing Committee on the Law of Patents (SCP) at its twelfth session held from June 23 to 27, 2008, in Geneva, the Secretariat prepared a preliminary study on the dissemination of patent information (document SCP/13/5) which was discussed at the thirteenth session of the SCP, held from March 23 to 27, 2009, in Geneva. While the SCP agreed that the document constituted a good basis for discussions, it requested further elaborations and clarifications of various aspects contained in the document. Therefore, the SCP at its thirteenth session requested the Secretariat to prepare a concept paper on technical solutions to improve access to, and dissemination of, patent information.

2. Accordingly, this document elaborates on some of the elements described in document SCP/13/5, taking into account the discussions held on this topic at the thirteenth session of the SCP.

II. CONCEPTUAL FRAMEWORK

3. Patent information is a term that refers to the business, legal and technical information that is disclosed during the patent application, examination, grant and maintenance processes. Under most jurisdictions, patent information is disseminated in the form of a Gazette or Journal, published patent applications, published patents and legal status registers, all of which contain detailed information in a highly standardized form.

4. Users of patent information are typically those people directly involved in the patenting process in patent offices (patent examiners), or those interested in or involved with patent application procedures such as patent applicants and their agents. More recently, new uses of patent information are emerging in the areas of research and development, business/financial analysis, economics and other areas of policy analysis.

5. Patent information constitutes a unique source of technological, business and legal information, serving different users' needs. The technological data contained in patent documents allows users to learn from existing solutions to specific technical problems. Thus, patent information allows users to avoid duplicating research and development effort and build on the existing state of the art, ultimately promoting further innovation. Business information from patent documents can be used by enterprises to monitor the activity of their competitors and help them plan their business strategies. Accurate and up-to-date information about the legal status of a patent or patent application provides entrepreneurs with an indication as to their freedom to operate, preventing potential infringements. Legal information also contributes to an effective patent system by improving the visibility and transparency of rights owned by domestic and foreign inventors.

6. Patent searches can be carried out to retrieve various types of information relevant to users' need. They may be used to determine the state of the art in a certain field of technology (state-of-the-art searches). Patent offices, applicants and their agents may also seek to determine whether or not an invention is novel (novelty/patentability search). A higher quality of patent search results, crucial for patenting and examining procedures, can be achieved when access to the broadest possible coverage of patent information is ensured.

7. Increasingly, information contained in patent documents is also supplemented by scientific and technical information published in other forms, such as journals or research databases, but the record of previous patents and patent applications is still the most important resource for most patent searches. In some countries and under international or regional patent systems such as PCT and the European Patent Convention, a patent application is published with a search report that contains the result of the prior art search performed by a searching authority with citations which may be patent documents and/or articles in scientific and technology journals (non-patent literature).

8. Improving the accessibility of patent information involves several processes:

- Digitization of patent documents
- Dissemination of information
- Provision of online databases
- Provision of high-volume data products
- Training and awareness-raising
- Development of tools facilitating patent search

III. AVAILABILITY OF PATENT INFORMATION IN DIGITAL FORM

9. There are a vast number of patent documents available in digital form in databases around the world. There are up to 65 million patent records from around 100 countries available in public or commercial databases – although these numbers need to be interpreted with caution, because the information content and coverage is often very limited. In particular, many digitized patent records consist of basic information only, and less than half may have a searchable abstract. Complete published patent specifications and searchable text documents are only available for a small number of countries.

10. Many countries remain, for which no information is available in digital form or for which information is incomplete. This means that there is a lack of information about what rights may exist in those countries. According to a recent survey (Circular CN.3027¹) carried out by the International Bureau of WIPO, almost 30 percent of responding offices indicated that their patent collections are available in paper form only.

¹ The survey was conducted within the framework of a project being carried out under Recommendation 8 of the Development Agenda. The complete survey results and related study paper will be presented separately to the Committee on Development and Intellectual Property. Preliminary results are included in this paper where relevant.

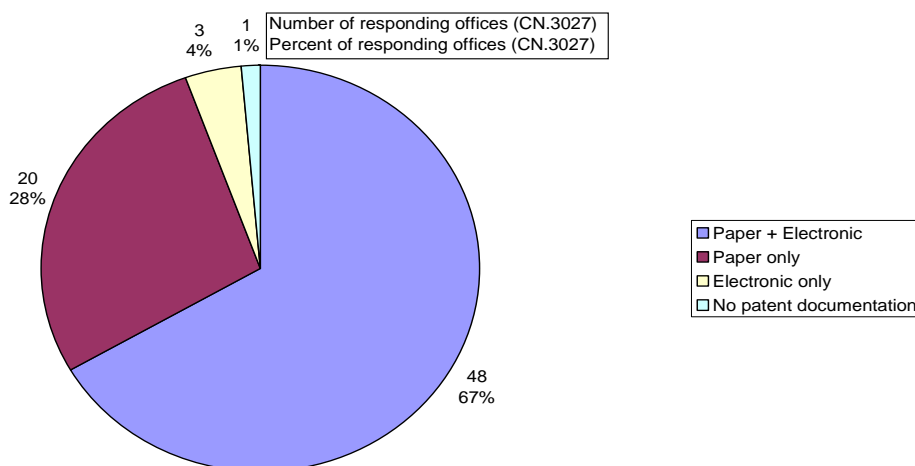


Fig. 1: Availability of patent collections in paper and electronic form

11. Legal status information (reliable information about the ownership of patent rights, about maintenance of patents and about litigation or other claims against a patent right) is difficult to obtain for many jurisdictions. This is an area where standardization of data and of dissemination mechanisms has the potential to greatly improve clarity on the status of patent rights. The issue of accessibility of legal status information is being examined more closely within the broader framework of the project “Intellectual Property and the Public Domain” being carried out under the WIPO Development Agenda.²

12. Digitization of patent information is therefore an activity where much work remains to be done in offices of all sizes and in all parts of the world.

13. There are several different levels of digitization that can be achieved. The first step is digitization of bibliographic information relating to published patent documents. Bibliographic information includes patent numbers, dates, applicant and/or inventor names, classification symbols, priority references, title of the invention and, in some cases, an abstract. Most offices achieve this level of information in computerized form, if only for the purposes of publishing a Gazette or Journal. For offices with automated administration systems, bibliographic data is generated as a result of the administrative processes.

14. There exist several WIPO standards for the coding of bibliographic information, which simplify understanding and exchange of information. Some relevant WIPO standards include:

- ST. 9: Recommendation concerning bibliographic data on and relating patents and SPCs
- ST.10/C: Presentation of bibliographic data components
- ST.16: Identification of different kinds of patent documents
- ST.8: IPC symbols on machine-readable records
- ST.32: Markup of patent documents Using SGML
- ST.36: Processing of patent information using XML

² See CDIP/3/4 Annex I.

15. A full list of relevant WIPO standards is available at http://www.wipo.int/standards/en/part_03_standards.html.

16. An important consideration for maintaining digitized patent records is quality. Although many offices maintain computerized administration systems, the data contained in those systems is not always accurate and correct. This may be because the information has been erroneously entered in the systems without the process of data validation (perhaps mistyped), because the information has not been updated as changes are made, because some information is missing (perhaps because it is not required for the administrative procedures), or because information has been lost (perhaps during migration from one computer system to another). In order to prepare for dissemination of patent information, many offices must first undertake a data quality improvement exercise which may include correcting and re-keying information, sometimes from the original paper records.

17. Once bibliographic information is captured, the next stage of digitization is the scanning of the patent documents themselves. This involves the scanning of published patent specifications or patent applications.

18. Scanning may be undertaken only for the purposes of patent information dissemination, in which case the final patent specification is scanned at the time of publication, or it may be an integral part of a paperless administration system. In the first case, the scanning operation is essentially an archiving process based on pre-prepared documents on paper. In the case of a paperless administration system, scanning is usually done on receipt of the documents and the documents are managed in an electronic document management system (EDMS) which will then be used to produce the final documents for publication and dissemination.

19. Two WIPO standards are particularly relevant to the creation and processing of documents in electronic form.

- ST.33: Data exchange of facsimile information of patent documents
- ST.36: Processing of patent information using XML

20. Unless the patent office has a system for electronic filing and processing of patent applications in XML full-text format (a subject beyond the scope of the present paper), patent documents will be scanned and processed in image format. Documents in image format may be viewed and printed, but they cannot be searched. To create searchable text from a scanned document, the document must be processed using optical character recognition (OCR) technology.

21. OCR is a process which is imperfect and the quality of the output varies greatly depending on the quality of the scanned images, language of the original document, and many other factors. The process can also be expensive as it requires high-performance computers and specialized software. Non-text elements, such as chemical and mathematical formulae, diagrams and tables, are usually not converted successfully to machine-readable text by an OCR process. Some offices employ manual correction and data entry to improve the quality of the OCR output, but this can quickly become an expensive exercise.

22. A final step in digitization is the making available of the electronic file of a patent application, including copies of correspondence, search and examination outputs, administrative documents and legal status records such as maintenance fee payments. This

information is extremely useful for patent applicants and practitioners to understand the quality of a patent.

23. The first efforts to create a global patent database led to the establishment of the International Patent Documentation Center (INPADOC) in 1972 by the World Intellectual Property Organization and the Republic of Austria. INPADOC collected and disseminated basic bibliographic data and patent documents and disseminated the information to patent offices and to private entities for a fee. Coverage was later extended to include legal status information. Given the technology available at the time, this service greatly improved the availability of patent information worldwide. INPADOC formally ceased to exist when its assets were integrated into the European Patent Office in 1991. Since then, the EPO has continued to collect and disseminate bibliographic data from its DOCDB database and legal status information from its PRS database. Since INPADOC, there have been no other coordinated worldwide projects to improve accessibility and availability of patent information. However, most patent information services today still rely on the assets that were originally created by INPADOC.

24. In cooperation with numerous national and regional patent offices, the International Bureau of WIPO contributes to the greater accessibility of patent collections by supporting the digitization of these offices' patent collection and making them available to the public through the PATENTSCOPE[®] search service. Under the project "Intellectual Property, Information and Communication Technologies (ICTs), the Digital Divide and Access to Knowledge", being carried out within the framework of the Development Agenda, member states are assisted in digitizing paper-based documents of IPRs and in creating searching facilities in order to provide public access to the database in a user-friendly manner.

25. Various WIPO Committees also play an important role in increasing the accessibility of patent information, including the Committee on WIPO Standards (CWS), which is tasked with the revision and development of WIPO standards relating to intellectual property information.

IV. CLASSIFICATION

26. Searching patent documents can be rendered more difficult as a result of language issues as well as diversity in technical terminology. To address these issues, patent documents are assigned classification symbols by patent offices, often as part of the search and examination process. Classification systems enable accurate retrieval of documents using search strategies that are language and terminology independent. A number of different classification systems exist, whose highly complex hierarchical structures and underlying technology paradigms are beyond the scope of the present paper.

27. The main classifications used are the International Patent Classification (IPC) which is used by many patent offices, the European Patent Office Classification System (ECLA), the U.S. Patent Classification System (USPC), the Japanese Patent Office classifications (FI and F-terms), and German Patent and Trademark Office (DEKLA classification). Apart from the classifications developed by Patent Offices mention should be made of the Derwent World Patent Index (DWPI) classification, which is the only classification/indexing system developed by a commercial company.

28. In order to assist users familiar with one particular classification system to retrieve information from documents classified under a different system, concordance tables have been widely established, relating symbols from one system to those from the other.

29. There are many developments underway to improve and further harmonize the different classification systems that are used worldwide, but they are beyond the scope of the current document. The IPC Revision Working Group (IPC/WG) is central to these efforts, preparing the revision of the IPC and dealing with search systems based on the IPC.

V. LANGUAGES

30. Patent applications and utility models are typically published in the national language(s) of the country or region in which they are filed. As a result, the information contained within these documents often remains inaccessible to users not competent in the language of publication. As patent filings have become increasingly dispersed among a larger group of patent offices, representing a greater range of different filing languages, and interest has increased in the patent collections of these offices, the need for access to documents in different languages has become increasingly critical.

31. Database services have employed a number of different means to address this issue. Many providers use manual machine translations to provide searches in pre-translated text of titles and abstracts, usually in one language only. Some search systems have indexed text in multiple languages and allow for searches in many languages. Due to the limitations of machine translation systems and the volume of data involved, some providers are switching to a strategy of providing cross-language search tools in which query terms are translated into multiple languages and the original documents are searched in multiple languages simultaneously. Such systems are complicated by the fact that the terminology used in patent documents is often new and not available in generic dictionaries, and can often be used differently in patent documents than in general documents.

32. Ideally, users would be able to use their own language to search, retrieve and read documents written in multiple other languages. Progress in this area has been rapid in recent years with new services, such as Google Translate, demonstrating that machine translation technology has a lot of potential. However, much work will be required by the patent information community to adapt generic tools to the specificities of patent documents.

VI. DISSEMINATION OF PATENT INFORMATION

33. Patent information may be disseminated in several different ways

- Data exchange agreements between IP offices
- Sale of patent data to the private sector
- Online databases and high-volume data products (discussed below)

34. Patent offices that maintain their own databases for search and examination purposes exchange patent data between themselves, usually in the context of a bilateral agreement specifying the terms and conditions of the exchange. In almost all cases, the information from other offices is not made available to the public. Even where patent documents are available in electronic form, they are not consistently digitized in such a fashion as to allow searching

in all parts of the document (the so-called “full-text”). In many cases, only titles or patent serial numbers (application or publication numbers) can be searched.

35. In the absence of multilateral agreement on full data exchange, the coverage of those patent collections included in the PCT Minimum Documentation³ appears to be weak in terms of searchable full-text available in search systems used by patent offices. As revealed in the survey carried out by the International Bureau of WIPO, nearly 50 percent of responding offices using any kind of search systems did not have access to all the patent collections included in the PCT Minimum Documentation. For certain collections, as few as 3 percent of responding offices could access full-text of patent documents.

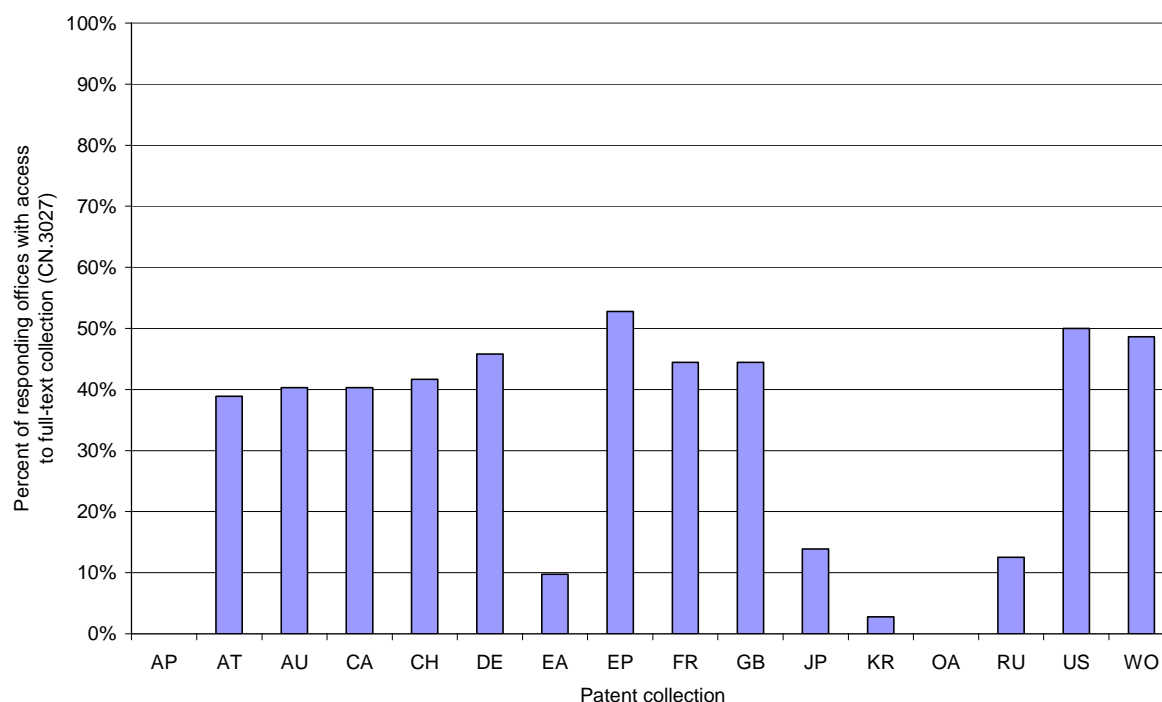


Fig. 2: Availability of full-text patent collections in search systems used by patent offices

36. Patent offices also sell data to commercial patent information service providers. Pricing policies are usually based on “marginal cost”, meaning that the office does not seek to make a profit from the sale of patent data, or to recover the full cost of patent publication.

VII. ONLINE DATABASE SERVICES

37. As mentioned above, various patent search types can be carried out, depending on the user’s profile and patent search needs. Irrespective of the type of patent search that a user carries out, the major tool for all patent searches (patentability, validity, infringement, clearance and state of the art searches) is the use of online searchable patent databases. The coverage of these databases varies: it may cover data from different countries, patent authorities, classifications, date range and content, from patent titles or abstracts to full text data and from patent applications to granted patents and utility models.

³ See Rule 34 of the PCT Rules and Regulations.

38. There is now a vast array of online databases for access to patent information. They fall into three main categories:

- (a) online database services maintained by patent offices and other organizations providing access to patent documents issued by the office or by also other offices.
- (b) commercial databases which are accessible to subscribers only.
- (c) independent free-of-charge databases.

39. According to the above-mentioned survey carried out by the International Bureau of WIPO, over 20 percent of offices – including those that do not carry out substantive examination but may provide other services to their stakeholders – use free-of-charge databases only, while nearly 15 percent do not use any kind of patent database at all.

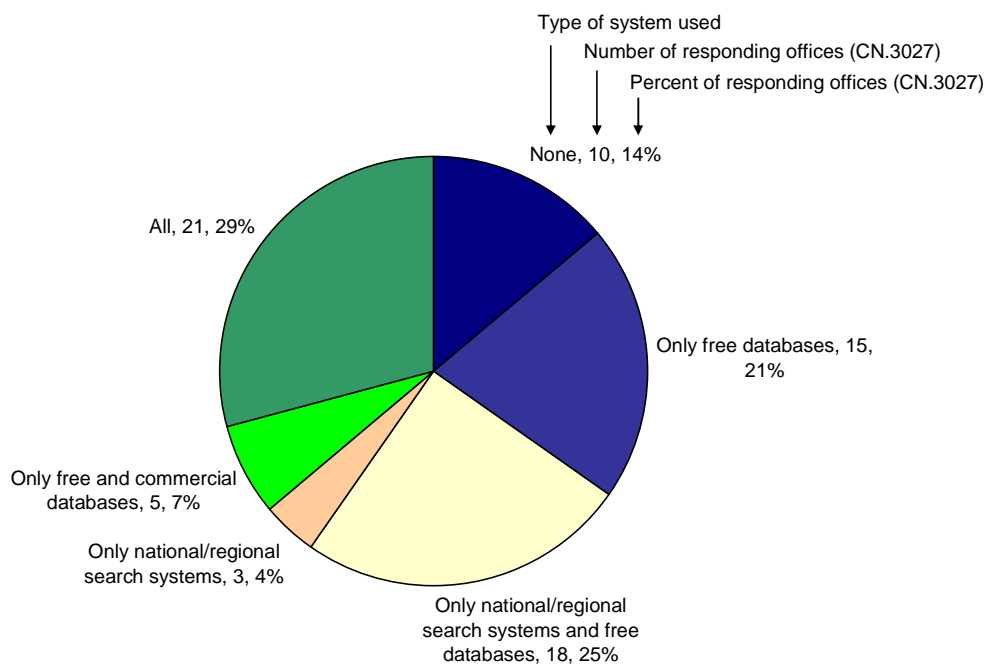


Fig. 3: Use of free and commercial patent databases by national offices

40. The differences among the free-of-charge and the commercial online databases are mainly their coverage as well as the search and analysis tools that they offer. The free-of-charge online databases maintained by national patent or IP offices usually provide access to their national records and usually do not provide searchable full-text or full historical coverage. Commercial services offer added-value features such as integration with non-patent literature data and analysis tools, as well as more sophisticated search features. A description of the various available search and analysis tools follows in the “Technical Tools for Access to Patent Information” section, while an indicative catalogue of currently available databases provided by national/regional offices and commercial providers is attached as an Annex.

41. Options for improving access to online patent (and non-patent) databases are currently being explored under the Development Agenda project “Specialized Databases Access and Support”. This project aims to facilitate the access by national offices of developing countries, especially LDCs, as well as their regional and sub-regional IP organizations to

specialized databases for the purposes of patent searches, as called for under Recommendation 8 of the Development Agenda.

VIII. HIGH-VOLUME PATENT INFORMATION PRODUCTS

42. Many patent offices disseminate raw patent data through high-capacity mediums such as DVDs or through network transfer methods, in particular to other offices and commercial service providers. Most offices have now migrated their searchable database services to online services, although some searchable DVD products are still available, notably the ESPACE series produced by the EPO.

43. The PATENTSCOPE[®] data services provided by the International Bureau of WIPO include bibliographic data and abstracts (in French and English) and application images for the complete PCT collection as well as searchable full-text description and claims for all PCT applications filed in English, French, Spanish, German, Russian, and Japanese. Similarly, data services such as those offered by the United States Patent and Trademark Office, the German Patent and Trademark Office and the Japan Patent Information Organization (on behalf of the Japan Patent Office) include bibliographic data and abstracts, searchable full-text, and images for their local patent collections.

44. By contrast, some data services only include bibliographic data and, sometimes, abstracts or application images. The DOCDB data service offered by the EPO covers bibliographic data from around 80 patent collections, though only covers abstracts for some cases. Though data services generally provide access to descriptions and claims only in the original language of filing, services such as the Patent Abstracts of Japan and Korean Patent Abstracts include bibliographic data and abstracts in translation, into English in these particular cases.

IX. SEARCH AND EXAMINATION REPORTS

45. As indicated in document SCP/13/5, many patent offices make publicly available the search and examination reports they prepare through online file inspection systems or on demand by other means. Additionally, mechanisms exist for the exchange of search and examination reports among patent offices. The “Common Access to Search and Examination Results” Foundation Project of five intellectual property offices (IP5), namely, the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the State Intellectual Property Office of the People’s Republic of China (SIPO) and the United States Patent and Trademark Office (USPTO), aims to enable examiners to obtain access to dossier information of other patent offices, such as search and examination results, through a single source.⁴

46. Other similar projects include an initiative being carried out by the IP Offices of Australia, Canada and the UK (known as the Vancouver Group) with technical support from the International Bureau for the exchange of search and examination documents. Furthermore, eight Latin American patent offices (Argentina, Brazil, Chile, Colombia,

⁴ See SCP/13/5.

Ecuador, Paraguay, Peru, Suriname, and Uruguay) in partnership with the Inter-American Development Bank have launched a project to exchange search and examination data on patents and trademarks.

47. In June 2009, the ASEAN Patent Examination Co-operation (ASPEC) program was launched by Singapore, Cambodia, Indonesia, Malaysia, the Philippines, Lao PDR, Thailand, and Viet Nam. The purpose of the project is to reduce workloads and improve turnaround times and to improve the quality of the search and examination through the sharing of such information on request by the applicant. Participating patent offices may consider the information provided to them through this mechanism, but are not required to adopt the findings and conclusions reached by other offices.

X. TECHNICAL TOOLS FOR ACCESS TO PATENT INFORMATION

48. The access to and use of patent information through online databases can be supported by various search and analysis features. These include options to carry out searches in specific fields (e.g. abstracts, claims, specific classifications, collections or dates) and to more flexibly structure search queries, producing more relevant results. Since non-uniform classification and abstracting practices represent an obstacle to the effective searching of patent documents, certain search services offer additional classifications and standardized abstracts, notably through the Derwent World Patent Index. Diversity in the naming of applicants (e.g. IBM, International Business Machines, etc.) is addressed through applicant indexes incorporated into some commercial search services. Questel's QPat search service, for example, offers a look-up tool, which groups various names used for applicants, inventors, and assignees.

49. In some cases, users are interested in patent information in special fields of interest. According to the aforementioned survey (Circular CN.3027), the main fields of technology in which patent applications are filed at responding offices are, in order of appearance among offices' top three fields of technology: human necessities (71%), in particular medical or veterinary science and hygiene; chemistry and metallurgy (63%), in particular organic chemistry; performing operations (39%); mechanical engineering (31%); and electricity (24%), in particular communications. To address these users' needs, some services such as the STN search service operated jointly by FIZ Karlsruhe and the Chemical Abstracts Service, also allow searching using chemical structures and formulas as well as biological sequences in addition to text-based searching.

50. Some search services allow users to identify patent families (groups of patent applications related through one or more priority filings), largely using the European Patent Office's DOCDB and the Derwent World Patent Index patent family data sets. Relationships between specific patent documents and the patent documents that they cite (backward citations) or in which they are cited (forward citations) may also be highlighted. Thomson Innovation, for example, includes links from specific patent documents to citing and cited patent and non-patent documents and allows the creation of "citation maps", which graphically illustrate the citation relationships between different documents.

51. Commercial search services often provide tools for analyzing search results, e.g. allowing users to identify top applicants or highly relevant keywords associated with a particular search or to create custom charts, graphs, and reports based on results. LexisNexis'

TotalPatent search service, for instance, offers a feature that creates a customized report on a set of patents, where the user may define the date range, the patent section to be included in the report, report language and file format. Machine translation tools are also frequently employed to allow users to work effectively with documents presented in a variety of languages.

XI. OVERVIEW OF WIPO'S PATENTSCOPE® SEARCH SERVICE

52. The World Intellectual Property Organization (WIPO) offers, through its PATENTSCOPE® search service, the possibility to search and retrieve full-text data from published international patent applications as well as for patent documents from the national/regional collections of ARIPO, Cuba, Israel, Korea, Mexico, Singapore, South Africa and Vietnam. For the PCT collection, complete file contents including International Search Reports, Written Opinions and Preliminary Reports on Patentability can be retrieved as well as national phase entry data for more than 40 offices.

53. Thanks to bilateral cooperation agreements between WIPO and participating offices, the coverage of the service will be extended to the national collections of several other offices in 2010 and beyond. Under these cooperation agreements, WIPO provides technical assistance to the participating offices for the digitization, formatting and dissemination of their patent data and documents. The level of technical assistance depends on the needs and capabilities of each individual office. These services are intended to better serve the needs of patent offices and other patent information users, with a particular emphasis on the needs of developing countries and new users of patent information.

54. The PATENTSCOPE® search service currently permits simple, structured, and command-line searching in 10 languages across 9 patent collections, of which three can be currently searched in full-text. The number of collections available with searchable full-text is to be expanded over time.

55. Users have the option to obtain graphical analyses of search results, but also analyses by clusters, i.e. patent office, IPC, applicant, inventor and filing date in table or graph format. They may also choose relevance or chronological ranking and translate records into all PCT languages of publication using an integrated Google translation tool.

56. Other PATENTSCOPE® services include a "Technology Focus" feature (<http://www.wipo.int/patentscope/en/dbsearch/analysis.html>), which offers an overview of trends in certain fields of technology, specifically traditional medicine and alternative energy.

XII. POSSIBLE AREAS OF IMPROVEMENT

57. As indicated above, access to patent information in digital format, in particular searchable full-text format, is still limited globally. Many patent collections are still only available in paper form or, if available in electronic form, are often not searchable in their entirety. Digitization projects are contributing to a greater availability of full-text patent data, however many collections still remain to be digitized. Similarly, the availability of legal status information, such as information on the validity of patents, needs to be improved.

58. Patent data can currently be obtained from a variety of sources, often through fee-based services. As a result, smaller patent offices, in particular, frequently lack access to the patent information resources they require to effectively carry out their functions in terms of search and examination. Individual inventors and SMEs also may lack the capacities to effectively use patent data dispersed among many different sources. Though certain search services such as EPO's esp@cenet search service already provide access to a number of different patent collections through a single source, these services are often not accessible in full-text and are limited in functionality. Improvements are need in the coverage and content of global search services, especially in the accessibility of full-text data, in particular among patent collections belonging to the PCT Minimum Documentation. This would contribute significantly to improving the quality of searches carried out by patent offices and applicants, as well as improving the scope of technical information available to the wider public.

59. Finally, search systems need to be more functional and user-friendly in order to be usable by non-specialist users. More investment needs to be made in search tools and language tools to enable better access to information by specialist and non-specialist users.

60. Technical improvements are being undertaken both at the national and international level. In particular, as described earlier, relevant WIPO Committees, various projects under the Development Agenda, and WIPO's Global IP Infrastructure Sector have been working towards technical solutions to improve access to, and dissemination of, patent information. Apart from the patent information relating to the PCT, patent information concerning national and regional applications is primarily held by each national/regional patent office. Therefore, further international cooperation is required in order for data to be provided in a more centralized and easily accessible fashion.

[Annex follows]

ANNEX

SELECTED PATENT DATABASES

1. Patent Offices and Organizations

- (a) World Intellectual Property Organization (WIPO) - PATENTSCOPE[®] search service
<http://www.wipo.int/pctdb/en>

The PATENTSCOPE[®] search service of the World Intellectual Property Organization includes PCT patent applications filed since 1978 as well as eight further national and regional patent collections.⁵ Bibliographic data, abstracts in French and English, and full-text descriptions and claims are searchable for all published PCT applications filed since 1978. For the Mexican and South African collections, bibliographic data, abstract, and full-text description and claims are searchable from 1992 and 1996 onwards, respectively. For other collections, bibliographic data and abstracts are searchable for varying time periods. The PATENTSCOPE[®] search service offers structured and command-line PCT search interfaces in English, French, Spanish, and Japanese. Based on search results, graphical analysis and clustering of top applicants, IPC subclasses, and other data can be carried out. For PCT applications filed from 2001, certain related documents (e.g. priority documents and declarations) can be accessed, while complete file contents are available for all PCT applications filed from 2009 onwards. PCT national phase entry data is also available for certain patent offices.

- (b) European Patent Office (EPO) - esp@cenet
<http://ep.espacenet.com>

The esp@cenet search service of the European Patent Office encompasses around 80 patent collections including European and PCT patent applications filed from 1978 onwards and the national collections of EPC member states.⁶ Bibliographic data and abstracts are searchable for European and PCT patent applications filed from 1978, with titles and abstracts being searchable in multiple languages (English and French as well as German for European patent applications). For other patent collections, data availability is mixed, both in terms of searchable document elements, language versions, and historical coverage. However, regardless of data availability, full-text description and claims are not searchable for any collection. The esp@cenet search service offers a number of different language interfaces including English, French, Spanish, Russian, and German-language structured search interfaces. File contents are partially available for European patent applications filed from 1978 onwards through the Register Plus system. For certain patent documents from all collections incorporated into the esp@cenet search service, legal status information and patent family data can be obtained through the INPADOC database.

⁵ For more details on PATENTSCOPE[®] data coverage, see:
<http://www.wipo.int/patentscope/search/en/search.jsf>

⁶ For more details on esp@cenet data coverage, see: *<http://www.epo.org/gpdc>*

- (c) German Patent and Trademark Office (DPMA) – DEPATISNet
<http://depatisnet.dpma.de/DepatisNet/depatisnet?window=1&space=main&content=index&action=index>

The DepatisNet search service of the German Patent and Trademark Office incorporates around 80 patent collections including German patent applications filed from 1877 onwards and European and PCT patent applications filed from 1978 onwards.⁷ For the German patent collection, certain bibliographic data and full-text description and claims are largely searchable for applications filed from 1877 onwards, while titles and abstracts are generally searchable from 1946 onwards. For other patent collections, bibliographic data, abstract, and full-text description and claims are searchable to varying degrees. The DepatisNet search service offers German and English-language structured and command-line search interfaces. Patent family data is available for many patent documents through a separate family search interface.

- (d) Japan Patent Office (JPO) - Industrial Property Digital Library (IPDL)
http://www.ipdl.inpit.go.jp/homepg_e.ipdl

The Industrial Property Digital Library (IPDL) of the Japan Patent Office provides access to patents granted from 1996 onwards, patent applications filed from 1971 onwards, examined patent applications from 1922 to 1996, and patent specifications from 1885 onwards. Bibliographic data, abstracts, and full-text description and claims are searchable for granted patents from 1996 onwards and for patent applications from 1993 onwards. For patent applications filed prior to 1993, only selected bibliographic data (including FI/F-terms) and, in certain cases, abstracts are searchable. The full range of search options is only available through a Japanese-language structured search interface. English-language structured search interfaces are provided to directly access specific patent documents through their application, publication, or patent numbers, to search FI/F-terms, or to search the Patent Abstracts of Japan (PAJ). The PAJ include English-language translations of titles and abstracts, Latin-character transliterations of applicant and inventor names, and IPC codes for patent documents filed after 1976. Legal status information is provided through the IPDL, including English-language legal status information for documents belonging to the PAJ.

- (e) Korea Intellectual Property Office (KIPO)
<http://patent2.kipris.or.kr/pateng/searchLogina.do?next=GeneralSearch>

The Korea Intellectual Property Rights Information Service provides searchable full-text data of Korean patent and utility models applications (since 20.06.1948), granted patents and utility models (since 25.03.1983), as well as Korean patents published in English abstracts (KPA – patent applications since 05.02.1999 and granted patents since 31.01.1973). The full-text is searchable in Korean, but the users may enter a search keyword in English and the system will search the Korean translation of the entered keyword within the Korean full-text. Information on the legal status is included, along with a presentation of the drawings. KIPRIS offers a K2E-PAT (Korean to English-Patent Automatic Translation) service which is fee-based; the users may translate the results through other free-of-charge services, such as Google Translate.

⁷ For more details on DepatisNet data coverage, see:
<http://depatisnet.dpma.de/DepatisNet/depatisnet?window=1&space=main&content=statb&action=statb>

- (f) State Intellectual Property Office of the P.R.C. (SIPO)
http://218.240.13.210/sipo_EN/search/tabSearch.do?method=init

The State Intellectual Property Office of the People's Republic of China (SIPO) provides a patent and utility model searchable database of Chinese patent and utility model applications and granted patents and utility models dating from 1985. The users may carry out searches within the following search fields: application number/date, publication number/base, IPC classification, applicant/inventor/patent agent's name, priority number or country/province code, allowing a combination with some operators, while the results can be sorted by any of these fields. There is also a quick search option, allowing to search in a search field in title, abstract, application number, applicant, inventor or IPC.

An English abstract translation is being shown along with the search fields' information, with a link to the free-of-charge English-Chinese translation service of the China Patent Information Centre ("China Patent Machine Translation - C. English full-text search is not possible

- (g) United States Patent and Trademark Office (USPTO) - PatFT and AppFT
<http://patft.uspto.gov>

The PatFT and AppFT databases of the United States Patent and Trademark Office provide access to US patents granted from 1790 onwards and to US patent applications filed from 2001 onwards, respectively. Bibliographic data, abstract, and full-text description and claims are searchable for granted patents from 1976 onwards and patent applications from 2001 onwards. For patents granted prior to 1976, only selected bibliographic data (patent number, issue date, and current US classification) is searchable. The PatFT and AppFT databases offer only English-language structured and command-line search interfaces. File contents and legal status information are available through the Patent Application Information Retrieval (PAIR) system.

2. Selected commercial providers

- (a) Chemical Abstracts Service (CAS) and FIZ Karlsruhe - STN
<http://www.stn-international.de/index.php?id=123>

The Scientific and Technical Information Network (STN) is an online database service that provides –through various interfaces like STN Easy or STN Express- various coverage for global patent and non-patent information (published research, journals, patents, structures, sequences and properties) of various data collections⁸ through various interfaces. It supports the specialized chemical/biochemical/chemical engineering search with access to MARPAT, CAS Registry and DWPI data. STN is offered in English, French, German, Spanish and Japanese as interface languages.

The users may search by various criteria, allowing the retrieval of information on chemical compounds through CAS Registry number or chemical name search. STN allows access to the Chemical Abstracts Plus (CAplus) database, which includes bibliographic chemical information, as well as the CAS Registry with organic and inorganic chemical substances and sequences content. Both databases also contain analysis based on patent and non-patent literature. Furthermore, the user has access to the MARPAT database, entailing

⁸ For full-text data coverage, see <http://www.cas.org/support/stngen/clusters/pnttext.html>

Markush chemical structures. Moreover, SciFinder tool allows drawing of chemical structures and then exact, substructure or similarity search. The results usually include bibliographic data and an abstract in English, with a hyperlink to full-text, where this is available.

STN AnaVist tool analyzes and visualizes various data (e.g. CAplus and DWPI), shows the relationships among various fields, provides summary or detailed reports with charts by e.g. patenting countries, priority applications, classifications, Derwent class or number. It offers the option of research landscapes with indication of areas where a company is active and the patents and articles which are related to it.

(b) Lexis Nexis - TotalPatent

<https://www.lexisnexis.com/totalpatent/signonForm.do>

TotalPatent provides bibliographic and full-text search in patent applications and granted patents stemming from various patent authorities⁹ and allows the users to search both in the language of publication and in English language Machine Translation, images, legal status, citations, patent family data as well as compressed and searchable PDFs. The interface language is English.

The users may search in various search fields: full Text, title, abstract, and Claims with selection among the various full text or bibliographic data authorities and assignee/inventor's name or publication date. TotalPatent allows the users to generate "word clouds" based on concepts, terms or phrases and to adjust the weighting of the terms. The suggested concepts can be reviewed and there can be assigned relative importance. In addition, the users can retrieve comments registered by TotalPatent user groups on certain patent documents. The user can display/sort the results by a range of fields including: publication date, assignee, class, etc. and select the "family" option and view either "main family" data, thus documents having exactly the same priority data or "extended family" data with family members sharing at least one priority data, and also family reports by using the relevant link.

The Generate Reports feature which creates a customized report on a set of patents, where the users may define the date range, the patent section to be included in the report (front page, bibliographic data, abstracts, claims, description, legal status, family info, images) document language, format (RTF, HTML, plain text, PDF or Word). Furthermore, the Conduct Analysis tool delivers analytics in various chart formats on selected data of the saved work folders (authority, kind code, inventor/assignee/attorney/agent/US examiner, filing/first published/priority/granted date, class, legal code and status) and citation mapping is also offered.

(c) Questel – QPAT

<http://www.qpat.com/index.htm>

QPAT covers full-text search in the following collections: Argentina, Austria, Belgium, Brazil, Canada, Chile, China, Denmark, EP, Finland, France, Germany, India, Japan, Mexico, PCT, Russia, Spain, Sweden, Switzerland, United Kingdom and the USA¹⁰, and IPC, US,

⁹ For more details on TotalPatent data coverage, see:

http://corporate.lexisnexis.com/Cms_managed_files/documents/TPlistofauthorities.pdf

¹⁰ For full coverage information see

http://www.questel.com/customersupport/Coverage_and_Updates_FullText.htm

ECLA and JP (FI and F-terms) classification. The Patent Number Wizard allows the users to format application and priority numbers into a Questel standardized format and the Browse Index constitutes an assignee/applicant lookup tool. In case of US documents lacking ECLA class, the user may retrieve through a statistical concordance tool, the “PCL/ECLA Correspondence” option, the most frequently used ECLA/PCT class that co-appears with documents with the same US classification. Moreover, a Family and Extended family search (FamPat and PlusPat file) can be carried out. Legal status, abstracts and citations can be included in the search results which can also be presented in a graphical display with priority data a timeline display. QPAT also offers Citations search (“cited/citing patents”), covering forward/backward citation, search by application/priority/publication and the possibility to display only a results list or also synthesis and graphs. The interface languages are English, French and Japanese and the search can be carried out in English, French or German. QPAT features multilingual search with automatically generated translation of the search keywords in different languages. Furthermore, graphical representation of the various family/citation relationships (by date, inventorship or ECLA/IPC classification) (Family Citation and PatCitation) is offered. Through the ANALYZE feature the most frequently appearing Assignee names or classification codes are identified and the top 15 Assignees, classifications and citations are displayed. For chemical patent searches in particular, Questel offers access to the Merged Markush Service (MMS) database which contains generic but also specific chemical structures.

- (d) Thomson Reuters – Thomson Innovation
<http://www.thomsoninnovation.com>

Thomson Innovation covers searchable full-text data of DWPI from 1963 onwards, PCT applications from 1978 onwards, US patent applications from 2001 and US granted patents from 1836 onwards, EP patent applications from 1978 and EP granted patents from 1980, British applications from 1916, French applications from 1971, German applications and granted patents from 1968, German utility models from 1983 onwards. With regards to Asia-pacific partially full text coverage in English is offered for Japan (limited coverage for granted patents from 1998-2004 and full text for granted patents from 2005 onwards, bibliographic data for applications from 1976-2003 and full-text coverage from 2003 onwards, bibliographic data for utility models from 1971 with gaps and full-text utility models from 2008 until present), China (limited coverage for patent applications and utility models from 2007 onwards) and Korea (bibliographic patent applications and granted patents from 1978-present, full text applications and granted patents from 2008 onwards, bibliographic data for utility models from 1979 and full text utility models from 2008). Patent Innovation also provides access to the DocDB file with data collection from over 60 countries and the Derwent World Patents Index (DWPI) from more than 41 patent-issuing authorities.

The language interfaces are English and Japanese and the users may search in English, German and French. Thomson Innovation allows simultaneous patent, Non Patent Literature and Business data search. The users can look for patent families and through the Analyze menu can create charts, but also view keyword clusters of the delivered patent and NPL data selecting text from specific fields, but also generate with ThemeScape a topographical content map based on the search results. Patent and NPL Citation maps can also be created, by generation, time or backward/forward citation.

- (e) Google – Google Patents
<http://www.google.com/patents>

Google Patents provides the capability to search with the full-text description and claims of US patents, generated using an optical character recognition process for patents issued prior to 1976. Users may search by patent number, title, inventor and assignee name, US Classification and IPC, issue and filing date, document status (issued patent or patent application), and type of patent (utility, design, plant, etc.), though assignment of data to specific fields is imperfect for patents issued prior to 1976, due to the fact that the assignment is carried out by an automated process. Google Patents provides English-language structured and command-line search interfaces.

[End of Annex and of document]