E

Patent Cooperation Treaty (PCT) Text Processing Task Force

First Session
Geneva, January 29 to 31, 2024

PCT TEXT PROCESSING TASK FORCE: INTRODUCTION AND BACKGROUND

Document prepared by the International Bureau

INTRODUCTION

1.   The International Bureau would like to update the filing, processing and publication of PCT applications to use a full text format[1] throughout the international phase. This should:

(a)   grow the proportion of applications filed in full text format by making it easier for applicants to file international applications in full text format with legal certainty;

(b)   deliver more accurate patent information to designated Offices and patent information users;

(c)   allow patent documents to be presented in different accessible formats, convenient to the needs of different users;

(d)   reduce formalities errors;

(e)   help compare different versions of a patent application or a patent application with its priority document; and,

(f)   allow self-service correction and amendment of patent documents, with immediate visibility of how the application will appear in formal publication and other formats.

---

[1] Throughout this document, the term 'full text format' is used to represent a document format where the document text is provided in a processible format and only figures and other no-textual components are provided in an image format

2.   The associated updates to publication systems also offer an easy route to reliable handling of applications with color drawings and potential future formats, such as 3D models or videos.

3.   The resulting system should be easier and more beneficial for applicants, patent Offices and patent information users alike. However, the International Bureau is aware of concerns that need to be addressed to ensure that it can work effectively and provides the necessary legal certainty in an international system with roles played by different national Offices and the International Bureau, in many cases using different IT systems.

4.   The International Bureau has therefore set up the PCT Text Processing Task Force to seek expert input on technical and legal matters in the development of processing in full text format for all actions at the International Bureau. This should allow more informed decisions, considering the interests of Offices, users and other stakeholders worldwide.

BACKGROUND

5.   For many years, major patent Offices have published patent applications and granted patents not as direct reproductions of sheets filed, but as typeset documents. In the 20 years since the introduction of electronic filing and publication of patent applications, consumers have increasingly benefitted from patent publications being available in full text format. Full text based on flexible formats such as HTML, XHTML, and XML that are compatible with different browsers has allowed individuals to view documents using a variety of devices and viewing formats to consume information in its most convenient way possible.

6.   While many Offices publish patent applications as full text documents, the International Bureau has continued to publish international applications in image format. And, while the PCT provides for full text formats, the International Bureau continues to support image-based formats for the submission of filing and post-filing documents. In view of the volumes of documents exchanged in both text and image-based formats between receiving Offices, International Authorities, the International Bureau and designated Offices, the International Bureau has sought to publish an accurate copy of the documents received, both in text and image format. The common format across Offices has been image-based and this remains the official publication. A full text product, furnished alongside the official publication, is provided for search purposes with varying levels of accuracy dependent on the qualities of the original documents and processing.

7.   The International Bureau is seeking to move away from this arrangement to one where the official publication is an accurate full text product that can benefit from the use of modern viewer tools and be converted into an image format for consumers that continue to need this format. This requires a significant growth in the filing of international applications in full text formats, and would assist the following goals:

        (a)   enabling self-service submission of full text replacements of the text of international applications, thereby eliminating manual page image editing to prepare an application for international publication;

        (b)   allowing for the processing and publication of color drawings;

        (c)   reducing formalities errors as most layout issues are no longer relevant and image conversion will not be required for color drawings;

        (d)   reducing the need for OCR processing and the use of OCR-produced documents, eliminating the errors associated with these processes; and,

(e) improving the delivery of information into full text search and translation document consumer environments.

8. The vision for PCT filing at the outset, as set out in Section 2 and Appendix III (Basic Common Standard for Electronic Filing) of Annex F to the PCT Administrative Instructions, is for XML processing to be the normal case. PDF and paper were seen as transitional steps towards a common, full text system of processing. However, while some Offices have made XML processing essentially mandatory, some Offices still do not accept it and take-up in Offices where PDF is offered as an alternative is limited.

9. While paper filings of international applications have reduced dramatically in recent years, and were below 1 per cent in 2022, electronic filing in page image-based formats, predominantly PDF, remains popular with applicants. For example, XML filings of international applications have stayed at a similar level in percentage terms over the past 10 years at between 26 and 30 per cent of all filings. Even in Offices where XML filing is normal, amendments and corrections are normally done by submission of PDF documents containing replacement sheets.

10. In moving forward, it is important to recognize why image-based formats remain popular and find solutions to address concerns in a way that can achieve the buy-in from users. Filing of international applications and related documents needs to be straightforward for users and ensure legal certainty, noting that several Offices are involved in the processing of an application during the international phase. While common word processors can prepare a document in DOCX format and this is the most common format for sharing draft applications between applicants and attorneys in most parts of the world, filing a patent application in this format needs to address the following issues:

(a) the legal status both of the files submitted and any files generated by a validator or convertor of the DOCX submission;

(b) the record of the files submitted in case of a need to audit the submission;

(c) the reliability and consistency of DOCX convertors;

(d) whether compatibility can be maintained between Offices using WIPO Standard ST.36 and those using ST.96;

(e) the importance of accepting replacement content at the level of headings, paragraphs, claims and drawings, rather than pages, and the consequential effects on paragraph numbering and cross-referencing, which would also affect the specifications of application processing software at Offices;

(f) rules for file formats such as tables, drawings and formulas (mathematical and chemical), as well as references to citations;

(g) the procedure and timeframe available for the applicant to resolve errors that might occur in the XML conversion process; and,

(h) consistency in any images referenced from the XML file.

11. Until recently, the International Bureau has processed international applications that it receives in full text format through the maintenance of a sheet image master version and the editing of a parallel full text version of the international application. From October 2023, the International Bureau is transitioning to process these applications in full text format with publication, in page image format, generated from the XML original. This page image publication of these applications includes generated change markup at the paragraph, figure

and claim level in place of sheet level stamped indications of changes. For the time being this limited to international applications filed at the China National Intellectual Property Administration as receiving Office where changes to the application have been recorded, corresponding to around 2 per cent of all international applications. The proportion may be grown by extending to all international applications filed in full text format and thereafter, a key challenge is to increase full text filings of international applications.

12.   A system of DOCX filing should be beneficial for all parties, eliminating most formalities issues, facilitating the introduction of publication of color drawings, helping ensure the identification at an early stage of issues that may cause problems with the preparation of typeset publications, and delivering reliable full text to patent search systems. However, such a system needs to be transparent, intuitive, consistent, reliable for long term records, properly understood by applicants and attorneys and equipped with safeguards giving confidence that any errors that do occur are few, likely to be spotted early and can be corrected at any stage.

ESTABLISHMENT OF PCT TEXT PROCESSING TASK FORCE

13.   The fifteenth session of the PCT Working Group "invited the International Bureau to continue to investigate the relevant issues with processing international applications in full text format to identify practical ways forward that would give clear and simple arrangements for applicants and would deliver the necessary technical results and legal certainty" (see document PCT/WG/15/19, paragraph 27).

OBJECTIVE, COMPOSITION AND MODALITIES

14.   The International Bureau has therefore established this PCT Text Processing Task Force to inform the International Bureau in developing legal and technical proposals for moving away from international applications based on page images to applications that are filed and processed in full text at all stages in the international phase, with the aim of creating a system that is easy to use and has the confidence of applicants and Offices alike.

15.   The Task Force will comprise representatives of IP Offices and user stakeholder groups in the PCT System, who should be experts on relevant aspects of this subject. The International Bureau intends the Task Force to meet around once per quarter until the issues are adequately understood and then as required to help monitor implementation and address any new issues arising. Most meetings will be with remote participation only, but there may be opportunities for the Task Force to meet at least partially in person, for example at the same time as the PCT Working Group. Work should continue between meetings using an electronic forum in the form of a wiki space, like those used by many task forces of the Committee on WIPO Standards.

OUTLINE PROPOSALS

16.   The International Bureau has identified an initial set of issues and requirements, further outlined in document PCT/EF/TPTF/1/3, that need to be addressed in moving forward towards a complete full text-processing infrastructure. They are grouped under three main initiatives listed below:

    (a)   Growing full text format filing volumes through customer experience improvements;

    (b)   Common functionality across Offices for the processing of international applications in full text format; and,

    (c)   Delivering confidence in filing certainty through coordinated legal provisions.

CONSIDERATIONS OF THE TASK FORCE

17.   In establishing the PCT Text Processing Task Force, the International Bureau anticipates that the first step in the work of the Task Force, is to review the status across Offices regarding full text filing and processing of patent applications. This review would enable information sharing across the Task Force members that would enable the review and prioritization of the proposals supporting the main initiatives, along with any additional proposals, to enable the transmission of the processing of all international applications in full text format.

[End of document]