TPTF/1 January 29-31, 2024

Agenda

1. Opening of the session
2. Introduction of the agenda
   (document PCT/EF/TPTF/1/1 Prov.)
3. Text Processing Task Force - Introduction and Background
   (document PCT/EF/TPTF/1/2)
4. Text Processing Task Force Initiatives
   (document PCT/EF/TPTF/1/3)
   - Full text processing applicant and Office landscape
   - Full text processing requirements
   - Discussion on priorities and proposed outline of activities
5. Summary of discussions and next steps
   (document PCT/EF/TPTF/1/4)
6. Closing of the session

WIPO Document Converter demonstration site

The WIPO Document Converter is available fully integrated in ePCT filing for use when filing applications.

It is also available for demonstration and testing purposes through the following standalone webpage. This document converter standalone webpage must not be used in the preparation of a production filing when the document is still highly confidential.

Meeting Summary

The first session of a text processing task force was held from January 29 to 31, to help us understand and address some of the issues . This was attended by delegations from 23 national Offices, 2 intergovernmental organizations and 7 non-governmental organizations representing user groups.

The task force reviewed the current landscape of electronic filing. As a result of the China National Intellectual Property Administration converting all international applications filed with it as receiving Office to XML format, the International Bureau now receives over half of international applications in XML format and we have begun to publish international applications based on the XML not only where they are unchanged from filing, but also where corrections and rectifications have been made.

Delegations took a very positive approach, but recognized that many issues would need to be addressed.

Some of the issues on the side of applicants and attorneys included:

- the need to be able to submit a document as approved by the applicant;
- to have a clear and reliable long term view of the application as filed;
- to have consistency in formats accepted and processed across different Offices;
- to be certain before filing that a document submitted is in a form that will be accepted;
- to have clear processes and timelines for addressing any conversion issues without affecting applicants' rights;
- to reduce the number of formalities errors and conversion errors;
- have effective ways of dealing with residual paper and image-based PDF filings; and
- to see real benefits for the applicant emerging from the process.

From the Office side, there was a need to:

- be able to plan and justify the investment in new processing systems;
- deal with existing applications in any new processing environment;
- have a clear understanding of what format is considered to have been filed and the formats used for processing and document exchange;
- to be able to make lossless conversions between ST.36 and ST.96 application bodies;
- to have a common approach for markup of amendments and corrections;
- to be able to deal with drawings filed as a separate unit from the description and claims; and
- to be able to handle special features accurately, including chemical and mathematical equations.

Addressing these issues and requirements would involve a complicated set of legal issues, technical standards, new procedures and IT development that would need to be considered and coordinated.

As we need to break the work down into manageable sections that help us to move forwards, we propose the next session of the task force will deal primarily with the question of file formats, and what can be used for filing that is easily prepared by applicants to a well understood standard, how can a satisfactory long term archive of the application as filed be provided, and what formats can deliver optimal processing implementations.

Example Applications

Example types

Level 1 - PDF Filing, sheet processing, sheet master published

Level 2 - XML Filing, sheet processing, sheet master published

Level 3 - XML Filing, parallel XML, sheet processing, sheet master published

Level 4 - PDF Filing, converted to XML, XML processing, XML master generates sheet publication

Level 5 - XML Filing, XML processing, XML master generates sheet publication

Example type	IA Number	Publication date	Features
Level 1	BR2020050475	03.06.2021	Formulae in description (see description tab (OCR)) Conversion delivers poor quality drawing images R26 sheet markup
Level 1 (-)	RU2020000337 (paper original)	14.01.2021	Formulae in description (see description tab (OCR)) R26 sheet markup
Level 1 (-)	IB2020020012 (Contingency upload ?)	17.09.2020	Embedded image in description (see description tab (OCR))
Level 2			
Level 3	IB2022052618 (DOCX Filing)	29.09.2022	WIPO docConverter created XML from DOCX uploaded in ePCT Original had line numbered sheets Amended claims sheet processed in parallel - see full text tab contains amended claims (XML), Official publication sheets contain stamps
Level 3	IB2022055678 (DOCX Filing)	29.09.2022	WIPO docConverter created XML from DOCX uploaded in ePCT Original had paragraph numbered sheets Color drawing - see figure 6B
Level 3	IB2022055496	22.12.2022	WIPO docConverter created XML in French
Level 3	IB2022050019	14.07.2022	WIPO docConverter created XML Nested ordered lists (see paragraph 16 in full text tab) Tables in XML (see paragraph 121 in full text tab)
Level 3	JP2023014556	02.11.2023	XML application body from applicant (JPO-PAS i480) R26 drawing markup in XML (e.g. Figure 21) and sheet stamps (e.g. sheet 21) based on parallel processing Drawing texts listed in reference signs section (unfortunately another drawing was selected for the front page)
Level 4	CN2023088337	11.01.2024	Chinese PDF to XML conversion Rule26 markup in paragraphs 79, 82
Level 5	CN2023130669	25.01.2024	Chinese XML Filing (English language) application body from applicant Amended claims
Level 5	CN2023130677	25.01.2024	Chinese XML Filing (English language) application body from applicant Drawings rectified Rule91 (In Patentscope, see full text section in XML and see black-line markup on PDF rendered publication)