

XML Filing and Exchange



IP5 Industry Meeting 18.01.2017

Overview of the Briefing

- Context
- Current situation
 - For Applicants
 - For the EPO
 - For exchanges with other Offices
 - The Challenges
- Proposal
 - Existing XML standard (Office Open XML)
 - Benefits for Applicants
 - Benefits for then EPO
- Next Steps
 - Annex F update
 - Implementation
 - Managing the Transition
 - User consultation

Context

- On-line filing has been operational since 2000
- Operationally it is working well
 - Approximately 95% of new filings at the EPO are now online
- There are new opportunities
 - Technology has moved on
 - PDF is no longer the only option
 - Publishing houses are now promoting online processing with word processor based XML
 - EPO's IT Roadmap has created new possibilities (e.g. eDossier)
- Questions
 - Can we be more efficient on both sides?
 - Are there better ways to exchange documents?

This presentation proposes a possible "state-of-the-art" answer

Current Situation

Current Situation - Applicants

Drafting

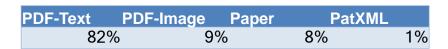
- 98% of filers at the EPO prepare application documents in MS Word
- Current word processors not able to generate WIPO ST 36 XML directly

Conversion

- 50% convert to PDF using AmyUni
- 40% convert using Adobe products
- 5% scan a printed copy
- 5% use other software to create PDF

Submission

- Current mixture of formats
- PDF-Image also includes 1mio pages per year by fax



Current Situation – EPO Receiving Office

- EP and PCT new applications 94,5% online v. 5% paper/fax
- Amendments 40% online v. 60% paper/fax
- PCT Ch I, II and Euro-PCT 100% image from WIPO or RO (except RO/EP)
- Estimated volumes in 2016

Pages	2016
Total EP Direct	2.616.000
Total EP SFD	724.000
Total PCT RO-EP	1.471.000
Total PCT Ch I	3.413.000
Total PCT Ch II	376.000
Total Euro-PCT	4.301.000
Total National	1.069.000
Grand total	13.970.000

- All online filings automatically loaded into PHOENIX** as Image and PDF
 - Never printed and scanned
- All paper filings scanned and loaded into PHOENIX as Images
- All incoming documents sent for Search Quality OCR (99,5%)
 - Used for EPOQUE and other EPO search and examination purposes

^{**} PHOENIX is EPO's current electronic dossier

Current Situation – The Challenges

- The volume of OCR
 - 10mio pages per year for incoming application documents
- The volume of re-OCR
 - Search Quality XML (99,5%) to Publication Quality XML (99.95%) for publication
- EPO Board of Appeal (BoA) Decision G1/10
 - No more correction of some obvious OCR errors after publication
 - Reinforced responsibility on the applicant to check the final specification
- Systematic exchange of image dossiers during the PCT phase
 - Image is the lowest common denominator in Annex F
- Applicants are asking for a simpler solution
 - But looking for something that will fit into their workflows

Proposal

Proposal – Word-processor XML standard

- Based on "Office Open XML" (ISO 29500:2008)
- Provide standard document templates and rules to ensure compliance
 - Single Unicode font
 - Complex tables, chemical and mathematical formulae foreseen as images
 - Paragraph styles for claims, description, drawings, abstract, headings
- Widely available
 - Native format for MS Word
 - Read and written by other word processors
 - e.g. LibreOffice, GoogleDocs
 - Available cross platform Windows/Mac/Linux
 - Already in use as the internal format by 98% of European Applicants
- Provide one, common validation engine for use:
 - By applicants before filing
 - By the EPO after filing
- Allows direct publication and use of the XML in the patent grant process

Benefits for Applicants

- "What you file is what we use"
 - An answer to BoA Decision G1/10 and correction of obvious errors
- Simpler process
 - Takes advantage of technology improvements in the last 15 years
 - Widely used on the publishing industry
 - No more need for conversion to PDF
 - Attach and send in-house document saved directly from word processor
- Secure process
 - Existing digital signature used to guarantee no changes after submission
 - Applicant has the digital "hash" to prove it
- Provides a solid basis for the future
 - Opens the door for colour support
 - An existing, well understood model for track-changes in amendments
 - Will also allow improved communications (e.g. hyperlinks for citations)

Benefits for the EPO

- No more OCR for valid XML filings
- Ready to publish document always available
- Content of XML filing directly loaded into EPO systems
 - Automatic extraction and Quality Control
- End-to-end electronic process in XML
 - Filing (eOLF/CMS*** Filing/WebForm)
 - Search (EPOQUE/ other search tools & eDossier)
 - Examination (eDossier & electronic publication editing)
 - A and B Publication
- Opens up similar opportunities for other offices that adopt the new standard

^{***} eOnLine Filing/Content Management System

Next Steps

Next Steps - Implementation

- Prepare updates to Annex F
- EPO implementation project
 - Includes building a prototype to demonstrate the new approach
 - MS Word Templates for applicants
 - eOLF/CMS*** Filing/Web-form filing to attach docx files
 - Validation engine for Applicants and EPO (same engine for both)
 - Loading XML text into existing EPO systems
- Design will include preparation for integration into new National Office hosted filing solution
- Pilot with selected users once Annex F approved
- Prepare any rule changes for approval by EPO's Administrative Council

^{***} eOnLine Filing/Content Management System

Next Steps – Managing the Transition

Support

- Communication on using the new templates
- User Days and Conferences to offer training
- Show how simple and secure it really is
- Explain advantages of re-filing in XML with any amendments
- Encourage
 - Look at adjusting fee policy to promote XML filing and staying in XML up to grant
- Balanced approach
 - Ensure a suitable transition period (e.g. two years)
 - Continue a parallel PDF-Text OCR chain for the diminishing number of remaining filings

User Consultation

Conclusions

Conclusions

- The current filing systems work well operationally
 - 94% online
 - Predominately PDF
- Challenges due to OCR conversion of the PDF
 - Costs
 - EPO Board of Appeal Decision G1/10
 - Multiple conversion steps
- Opportunity to adopt new technology and move forward
- Proposal to adopt word processing format already used by 98% of applicants
- This is expected to bring:
 - A simple and secure option for applicants
 - End-to-end XML processing inside the EPO
- Need to support and encourage users to transition
- Next step: Consultations with Users and other Offices