

Digitization Guidelines Survey

Response ID:196 Data

1. Identification page

1. Please enter the ST.3 code of the member state or inter-governmental organization you represent.

KR - Republic of Korea

2. Please enter the name of the Office or Organization you represent. For instance, International Bureau of WIPO.

If your organization name is the same as your ST.3 code name, you may put "n/a".

MOIP(Ministry of Intellectual Property)

2. General questions

3. Is your national patent collection part of the PCT minimum documentation or does your Office contribute your IP documentation to one of WIPO's databases? Please select all that apply.

Patentscope
Global Brands Database
Global Design Database
PCT minimum documentation

4. What file or document format(s) does your Office accept at filing for IP applications? Please select all that apply.

For industrial design collections:

XML
Image-based
Paper

For patent collections:

XML
Image-based
DOCX
Paper

For trademark collections:

XML
Image-based
Media-based formats (e.g., MP3)
Other (please specify): Paper

5. In what language(s) do you publish your IP documents in? Please provide a full list of 2-letter language codes.

EN, KO

6. In what format(s) do you publish or disseminate your IP collection? Please select all that apply.

For industrial design collections:

WIPO Standard ST.96
Image-based
3D formats
PDF

For patent collections:

WIPO Standard ST.96
PDF

For trademark collections:

Image-based
WIPO Standard ST.96

Media formats (e.g., MP3)
PDF

7. In what format(s) does your Office publish the bibliographic data associated with the format(s) indicated in question 8? Please select all that apply.

For industrial design collections:

XML
PDF

For patent collections:

XML
PDF

For trademark collections:

XML
PDF

8. If your Office produces XML or JSON instances and does not validate these instances before publication, please provide your reasoning below.

N/A

9. Does your Office publish the color drawings they receive, if they are permitted during the filing of a patent application?

Yes

3. Digitization activities

10. Does your Office perform digitization of documents you receive before publication or data dissemination, for one or more IP rights, or has experience of digitization of documents associated with IP rights?

Yes

10.1. How are or were these documents or associated data transformed from the format you received? e.g., paper document manually scanned and OCRred or DOCX converted to XML.

Paper document → Scanning → OCR processing → E-filing software → XML conversion
DOCX file → E-filing software → XML conversion
HWP/HWPX file → E-filing software → XML conversion
International application XML format → Domestic application XML format conversion

10.2. If you rely or relied on a third party to perform the digitization of your documents, why did you reach this decision?

Entrusted to an institution equipped with the facilities and personnel prescribed by law to ensure efficient digitization operations.

10.3. Does your Office use or did you use any IT tools, in particular AI tools, in the digitization process that you find particularly useful?

None at present

10.4. What are or were the biggest technical challenges you have experienced in digitizing your documents? e.g., extracting drawings or chemical formulae from a patent application. Please provide further details in the comments box below, including how you overcame these challenges and the relevant IP right type.

Patent specifications contain various objects such as drawings, chemical formulas, mathematical expressions, and component diagrams. However, current OCR technology can extract only text, requiring operators to manually digitize non-text objects by capturing and inserting them as images. To improve efficiency, document analysis software is planned to be introduced to convert documents into XML format, followed by post-editing, thereby reducing the time required for the digitization input process.

4. Digitization activities

11. If your Office is considering future digitization activities, please provide details regarding the tools you are considering, in particular AI tools, and include your reasons for selecting them for potential use in digitization.

AI-powered OCR technology enables the extraction of key-value data from various document templates, facilitating data reuse. In addition, document analysis software can analyze document structures and convert them into XML format.

12. Please provide any other feedback regarding capacity building for the digitization of IP document collections.

Member Offices should share practical examples of adopting AI and other digital technologies in their digitalization efforts.