

Digitization Guidelines Survey

Response ID:181 Data

2. Identification page

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

PH - Philippines

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".

Intellectual Property Office of the Philippines

3. General questions

5. 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

Other (please specify): WIPO Publish, ASEAN IP Register

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

For industrial design collections:

PDF

Image-based

For patent collections:

PDF

Paper

For trademark collections:

PDF

Image-based

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

EN

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

For industrial design collections:

PDF

Image-based

Other (please specify): Plain text

For patent collections:

WIPO Standard ST.96

PDF

Plain text

For trademark collections:

Image-based
PDF

9. 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:

PDF
Plain text

For patent collections:

PDF
Plain text
XML

For trademark collections:

PDF
Plain text

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

NA, we do validation of produced XML

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

Yes

4. Digitization activities

12. 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

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

1. paper documents were scanned, OCR'd and uploaded to the EDMS of WIPO-IPAS;
2. OCR'd patent documents being extracted, indexed and uploaded in WIPO-IPAS and once passed the validation, it will be readily available at the WIPO Publish for publication
3. readily available indexed patent documents from WIPO Publish were converted to DOCX and uploaded to WIPO-IPAS for conversion to XML for validation and shared as ST. 96 to PCT Min Doc

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

NA, we are performing our digitization in-house

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

IPOPHL uses Mathpix to convert OCR'd patent documents per section (AB, DE, CL, DR, SL, whichever is applicable) to DOCX

which is being uploaded in WIPO-IPAS and a module is being ran to transform DOCX to XML package with corresponding HTML for validation. Once validated, the XML package being rendered into XML ST. 96 compliant, otherwise, we re-do the conversion, uploading and validation process. This process was made operational through the assistance from WIPO-SG

12.4. 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.

1. extracting drawings, tables, images or chemical formulae from the OCRed to DOCX and XML.
2. some of the patent documents were submitted as an image being paste to a word file and being saved as pdf, thus, there is a need to OCR them first before conversion
3. existence of numbers or bullets in the claim, which is being recognized as claim number, thus claim number is inconsistent; i.e. Claim 1.....comprising.....and selected from.....
 1. xxxxx.....
 2. yyyyy.....

5. Digitization activities

13. 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.

Yes, the Office is considering future digitization, especially the backfiles. As of now, we consider using the digitization process as cited at item 12.3, since we are using WIPO-IPAS and support from WIPO-SG. We may also consider accepting DOCX or XML format of patent documents upon filing. Further, we are considering additional personnel that would speed up the digitization of backfile records.

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

More training for our in-house digitization team, especially on document validation and quality checking, would be beneficial. Additional personnel would also help speed up the digitization of backfile records and reduce backlogs. WIPO may also consider supporting a backfile digitization project, with the IP Office focusing on validating and checking the quality of the digitized documents. This would help make the digitization process faster and more efficient. Continued investment in training, technology, and knowledge sharing will help support the long-term success of IPO Philippines' digitization efforts.