

## **Committee on WIPO Standards (CWS)**

### **Eighth Session**

**Geneva, November 30 to December 4, 2020**

### **PROPOSAL FOR THE REVISION OF WIPO STANDARD ST.37 (TASK NO. 51)**

*Document prepared by the Authority File Task Force Leader*

#### **BACKGROUND**

1. At its reconvened fourth session in 2016, the Committee on WIPO Standards (CWS) created the Authority File Task Force to deal with Task No. 51, the description of which reads:

“Prepare a recommendation for an authority file of patent documents issued by a national or regional patent office to enable other patent offices and other interested parties to assess the completeness of their collections of published patent documents.”

The European Patent Office (EPO) was designated as the Task Force Leader (see paragraph 122 (e) of document CWS/4BIS/16).

2. At its fifth session in 2017 the CWS adopted WIPO Standard ST.37 - “Recommendation for an authority file of published patent documents” referring the pending Annexes III and IV of the Standard to the following sixth CWS session (see paragraph 61 of document CWS/5/22). Consequently, the CWS modified the description of Task No. 51 as follows:

“Prepare and present for consideration at the sixth session of the CWS, to be held in 2018, Annex III “XML schema (XSD)” and Annex IV “Data Type Definition (DTD)” to WIPO Standard ST.37 “Recommendation for an authority file of published patent documents.”

3. At its sixth session in 2018, the CWS approved a new version of WIPO ST.37, version 1.1 with the adopted the new Annexes III and IV of ST.37, which are based on WIPO Standards ST.96 and ST.36 respectively. Consequently, the CWS updated the description of Task No. 51 as follows:

“Ensure the necessary revisions and updates of WIPO Standard ST.37.”

The version 1.1 of Standard ST.37 was published in December 2018 as well as links to the new XSD and DTD files on the WIPO website to allow IP Offices (IPOs) to validate their authority file XML instances.

4. At its seventh session in 2019, the CWS approved a revision to WIPO ST.37, including the improvements to both Annex III and Annex IV of ST.37 (see paragraph 150 of document CWS/7/29). The CWS requested the XML4IP Task Force to update Annex III, originally based on WIPO ST.96 version 3.1, but then updated to version 3.2 before publication, to improve the descriptions of components and to rename particular components to ensure they were more consistent with the design guidelines set out for WIPO ST.96 (see paragraph 147 of document CWS/7/29). In collaboration with the XML4IP Task Force, version 2.0 of this Standard was published in December 2019.

#### PROGRESS REPORT

5. In December 2019, version 3.2 of WIPO ST.96 was published, which moved reusable components developed for the authority file schemas to the WIPO ST.96 library. This was to allow these elements to be reused by components in other WIPO ST.96 namespaces in the future.

6. In June 2020, version 4.0 of WIPO ST.96 was published. Included in this new version is an update to the list of former country codes, namely ‘RH’ (Southern Rhodesia), in response to a request from an IPO to ensure they could validate their legacy authority file data in XML format. The former country of Southern Rhodesia is listed as a priority country for a number of their published patent documents.

7. The Authority File portal ([https://www.wipo.int/standards/en/authority\\_file.html](https://www.wipo.int/standards/en/authority_file.html)) has been updated twice in April and October 2020 with updated data sets provided by IPOs, including the authority file of published PCT applications. The PCT authority file is provided each week as a zip file containing three elements: i) the published patent documents for the past week, ii) the published patent documents for the year to date, and iii) the full data set. This type of authority file data set, covering multiple time periods, had not been foreseen, so a naming convention for files such as these had not previously been proposed within the WIPO Standard ST.37.

8. In July 2020, it was announced that the PCT Minimum Documentation (MD) Task Force intended to adopt WIPO Standard ST.37 as a means to accomplish their Objective C, which states: “*Propose clearly defined bibliographic and text components of patent data that should be present in patent collections belonging to the PCT Minimum Documentation*”. However, for the PCT MD Task Force to proceed with this proposal, the current element defined in WIPO ST.37 would need to be complemented with some new elements, when patent documents are available in a machine-readable version. Specifically, these elements would be a series of Boolean flags indicating: whether patent abstracts are available, both in the original language and/or in English, and whether a full text version of the patent specification is available.

9. In view of this, the Authority File Task Force proposes that the CWS requests the Authority File Task Force to propose revisions to WIPO ST.37 to meet the requirements for PCT documentations proposed by the PCT MD Task Force, specifically, to include in the authority file itself the following mandatory information:

- (a) full-text (yes | no): the description and claims is provided in a searchable text format;

- (b) original abstract (yes| no): available in searchable text format; and
- (c) English language abstract (yes| no): available in searchable text format in English.

It is proposed that a revision in this regard be presented by the Authority File Task Force for consideration at the ninth session of the CWS.

#### SUMMARY OF PROPOSED REVISION

10. Under the framework of Task No. 51, the Authority File Task Force, in collaboration with the International Bureau, has prepared a proposal for revision to WIPO Standard ST.37. These revisions are required, after the publication of WIPO ST.96 version 4.0, to allow validation of legacy authority file data, as well as the inclusion of a new example which recommends a particular naming convention for an authority file data set, such as the PCT authority file.

11. The Authority File Task Force proposes the following revisions to the main body of WIPO ST.37 to:

- (a) Update the reference to WIPO ST.96 on page 3 of the Main Body so that the new title is provided:
  - i. “Recommendation for the Processing of ~~Industrial~~ Intellectual Property Information Using XML (eXtensible Markup Language)”;
- (b) Replaced all instances of ‘IP office’ in the Main body with the acronym defined: ‘IPO’ for consistency; and
- (c) Update paragraph 38 of the WIPO ST.37 Main Body to include a further example of file naming where the authority file data set covers more than one period of time. It should be noted that this example, proposed for inclusion within WIPO ST.37, does not capture instances where a zip file is provided, but it contains instead multiple files where each is organized by publication kind code, for instance. The proposed new item (c) under the paragraph 38 is as follows:

“(c) for the instance where multiple files are provided, and each covers a different time period, please refer to the table below for an example of recommended file naming:

CC_AF_gazetteNNXXXX_YYYYMMDD.ff	contains the authority file for publication NNXXXX where NN is the week number, XXXX is the year and ff is the file format (XML or TXT)
CC_AF_yearXXXX_YYYYMMDD.ff	contains the authority file for the year XXXX
CC_AF_YYYYMMDD.ff	contains the authority file with file format ff
CC_AF.zip	A zipped file which contains the three files above

“

12. The Authority File Task Force also proposed to make the following updates to Annex III of ST.37, XSD for Authority File:

- (a) Indicate a specific version number of WIPO ST.96 in the editorial note of Annex III. The new editorial note should read (proposed changes are shown):

*“Annex III of WIPO Standard ST.37 is the set of XML schema components to represent the minimum and extended data elements of an Authority File of patent documents issued by a Patent Office. Annex III is based on WIPO Standard ST.96 version 4.0, including the naming convention used to identify the names of the specific data components for Authority File. Annex III includes an Appendix which is a sample XML instance of an Authority File structured according to the XML schema, according to WIPO ST.96 V4 0.”;*

- (b) Update the import statement of XSD so that it references the new flattened schemas for V4\_0:

```
<xsd:import
namespace="http://www.wipo.int/standards/XMLSchema/ST96/Patent"
schemaLocation="ST96_Patent_V4_0.xsd"/>
<xsd:import
namespace="http://www.wipo.int/standards/XMLSchema/ST96/Common"
schemaLocation="ST96_Common_V4_0.xsd"/> ;
```

- (c) Update to version number to 2\_1 from 2\_0 and include the new update in the XSD declaration:

```
<?xml version="1.0" encoding="UTF-8"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:afp="http://www.wipo.int/standards/XMLSchema/AFPatent"
xmlns:com="http://www.wipo.int/standards/XMLSchema/ST96/Common"
xmlns:pat="http://www.wipo.int/standards/XMLSchema/ST96/Patent"
targetNamespace="http://www.wipo.int/standards/XMLSchema/AFPatent"
elementFormDefault="qualified" attributeFormDefault="qualified"
version="V2_1"> ;
```

- (d) Update the st37Version attribute to reflect the new version number V2\_1:

```
<xsd:attribute name="st37Version" type="xsd:token" fixed="V2_1">
  <xsd:annotation>
    <xsd:documentation>A specific release version of ST.37 XML
Schema</xsd:documentation>
  </xsd:annotation>
</xsd:attribute> ;
```

- (e) Update the schema modification date and reference to the release notes document within the annotation:

```
<xsd:annotation>
  <xsd:appinfo>
    <com:SchemaLastModifiedDate>2020-07-13</com:SchemaLastModifiedDate>

    <com:SchemaContactPoint>xml.standards@wipo.int</com:SchemaContactPoint>
    <com:SchemaReleaseNoteURL>http://www.wipo.int/standards/XMLSchema/AF
Patent/V2_1/ReleaseNotes.pdf</com:SchemaReleaseNoteURL>
  </xsd:appinfo>
</xsd:annotation> ; and
```

- (f) Update to the Appendix to Annex III, XML Instance, to revise the declaration to refer to the new version of the WIPO ST.37 XSD, V2\_1:

```
<afp:PatentAuthorityFile xmlns:xsi=http://www.w3.org/2001/XMLSchema-
instance xmlns:afp="http://www.wipo.int/standards/XMLSchema/AFPatent"
xmlns:com="http://www.wipo.int/standards/XMLSchema/ST96/Common"
xmlns:pat="http://www.wipo.int/standards/XMLSchema/ST96/Patent"
xsi:schemaLocation="http://www.wipo.int/standards/XMLSchema/AFPatent
PatentAuthorityFile_V2_1.xsd" com:officeCode="EP" com:creationDate="2020-
07-13" afp:st37Version="V2_1">
```

13. The Secretariat suggests that the revised Standard ST.37 be prepared in accordance with the decision by the CWS and be published on the WIPO website after this session.

#### UPDATE OF AUTHORITY FILES

14. The International Bureau has noted that many Offices are not providing updates to their authority files with the frequency that they have indicated. The International Bureau is therefore encouraging Offices to provide their regular updates on a specific date of the year. For Offices which update more frequently than annually, an update should still be provided at this time, in addition to any other required updates. The International Bureau is proposing March 1 as the specific date, or the next business day.

15. With the support of the CWS, the Secretariat intends to send a Circular inviting Offices, 1 month prior to the agreed specific date, to provide their authority file updates in advance of this new deadline.

16. *The CWS is invited to:*

*(a) note the contents of the present document;*

*(b) consider and approve the proposed revision of WIPO Standard ST.37 for version 2.1 as indicated in paragraphs 11 and 12 above;*

*(c) request Secretariat to prepare and publish the revised Standard ST.37 in accordance with the decision by the CWS as described in paragraph 13 above;*

*(d) request the Authority File Task Force to provide a proposal for necessary revisions to ST.37, in light of the proposal by the PCT MD Task Force, at the next session of the CWS, as indicated in paragraphs 8 and 9 above; and*

*(e) consider and approve the proposed date of March 1 for the provision of annual authority file updates and request the Secretariat to issue a Circular in February 2021, as indicated in paragraph 15.*

[End of document]