

Consideration requests from the Citation Practice Task Force
By the ST.36 Task Force

(SCIT/SDWG/9/3)

12. To assist in the provision of uniform and reliable methods of referring to parts of patent documents, the Task Force recommends the following should be brought to the attention of the ST.36 Task Force for their consideration:

- 12.1 the ST.36 Task Force should note the importance of consistency of numbering (identifying) parts of a patent document filed in fully electronic form so that such numbers can be referred to easily during the creation of citation references at a later stage;

[ST.36 TF's observations (draft)]

The Task Force took a note, but can take no further action to support this request in relation to the current WIPO Standard ST.36.

The Task Force remarked that although the WIPO Standard ST.36 ICES contain the necessary elements to identify different parts of a patent document, e.g., paragraph numbering tags, there are no best practice guidelines or tools (such as style sheets) to apply the elements in a consistent way to patent documents. The said best practice guidelines could include, for example, recommendations for implementation of paragraph numbering, long paragraphs, and how documents should be rendered consistently. Such guidelines would be of benefit to optimize consistent rendering of documents across different formats of a single patent document (e.g. html, xml, or PDF) as well as across different IPO and commercial provider publication platforms.

- 12.2 the ST.36 Task Force should note the importance of minimizing long paragraphs when creating or amending patent documents to avoid problems (later) when material contained in long paragraphs is being cited;

[EP comments]

EPO has stated already that we cannot do anything about this except recommend this to our applicants but the whole description could be one long paragraph if the applicant wishes it.

[ST.36 TF's observations (draft)]

The Task Force took a note, but can take no further action to support this request in relation to the current WIPO Standard ST.36.

- 12.3 referring to Annex I, and noting the intention of some IPOs to count various parts of the document such as gene sequences and computer programs the ST.36 Task Force could review the following:

- 12.3(a) the elements “maths” and “chemistry” could be reviewed and if appropriate marked as current, instead of being reserved for future use, within the International Common Elements (ICEs) to WIPO Standard ST.36, and

[CH comments]

In relation to the citations practices, we sustain the introduction of the elements "maths" and "chemistry" as current elements. We also sustain the introduction of new elements for gene sequences and computer programs.

Within a search report we refer to the claims only of the enclosing document. We don't foresee any restriction in the paragraph length and don't apply any special paragraph numbering for corrected documents.

In the search reports we use the ST36 tags for patent citations and non patent citations. If an URL is contained in these elements it is represented as an hyperlink in the PDF. We don't generate a reference list for a document

[EP comments]

EPO has stated already that maths and tables are not reserved and are already part of ST.36. Chemistry is more difficult - there is no agreed international standard so far.

[ST.36 TF's observations (draft)]

With regard to XML tagging for Maths and Tables, the relevant elements already exist in WIPO Standard ST.36.

WIPO Standard ST.36 recommends to tag mathematic formulae by MathML2 which is maintained by W3C (download from: <http://www.w3.org/TR/MathML2/DTD-MathML-20010221.zip>) and tables by OASIS Open XML Exchange Table Model which is maintained by OASIS (download from: <http://oasis-open.org/specs/soextblx.dtd>). For further information, please see the SECTION I of Annex A to WIPO Standard ST.36.

No agreement on chemistry standard within W3C has been reached. Therefore, it was agreed that the addition of a chemistry standard to WIPO Standard ST.36 should be postponed until the W3C standard for chemistry would be ready.

For further information, the recommendation regarding Chemistry in Annex A is as follows:

```
<!--  
  Chemistry  
  Chemical formulae, ring structures, Markush structures, etc.  
  For now treat as image data.  
  Depending on industry support we may use, e.g., chemML, CML later  
  Recommended id = chem0001, chem0002, etc  
-->  
<!ELEMENT chemistry (img | (chem , img?))>  
  
<!ATTLIST chemistry id ID #IMPLIED  
  num CDATA #REQUIRED >
```

```
<!--
  This is a pointer to an external DTD for chemical markup - to be determined.
  For EPO probably CML; http://www.xml-cml.org/
-->
<!ELEMENT chem EMPTY>

<!ATTLIST chem  id          ID          #IMPLIED
                file       CDATA       #REQUIRED
                chem-type  CDATA       #IMPLIED >
```

- 12.3(b) consider if further clarification is required to identify gene sequence lists, particularly when image content is defined as “dna” or the data is tagged as the element “bio-deposit” or “sequence-list”;

[EP comments]

EPO marks up some DNA data but not all. Because of the layout (mono-spaced font in columns) of long DNA sequence listing and the difficulty of extremely accurate data entry we capture as image data (it is safer). However, applicants do send in this data as ST.25 data as a separate file for internal processing by examiners.

[ST.36 TF's observations (draft)]

The ST.36 Task Force thinks that no further clarification is required to identify gene sequence lists.

- 12.4 review and suggest further document structure elements or attributes that could be included as further tags associated with WIPO Standard ST.36 that would aid the creation of citation references, such as the addition an attribute *example* (found predominantly in chemical patent applications) to the *heading* element; and

[EP comments]

The example heading can occur anywhere in a patent description; we prefer the simple <heading>Example</heading> rule.

[JP comments]

Regarding citation references, JPO thinks that cited part in cited patent documents should be identified by the structured tag like <embodiments-example> etc. rather than heading tag in view of characteristics of structured language XML. Anyhow, although the discussion on new tags/attributes is useful, the JPO thinks that the discussion on use of tags in uniform manner among users/offices is more important.

[AU comments]

As stated in the citation practices questionnaire our use of citation data in xml format to-date is virtually non-existent. However, in our future XML bulkdata product we expect to provide some related citation information in XML format. This data is collated by examiners at the time of

acceptance and is a free text field. It is identify as prior-art as opposed to a specific type of citation data ie patcit or nplcit data.
See example

Prior art data will be provided as part of the exam details sections of the AusIP Discovery Series schema as highlighted below.

```
<xs:element name="au-exam-details" minOccurs="0">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="au-prior-arts" minOccurs="0">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="au-prior-art" minOccurs="0" maxOccurs="unbounded">
              <xs:complexType>
                <xs:simpleContent>
                  <xs:extension base="xs:string">
                    <xs:attribute name="sequence" type="xs:nonNegativeInteger" use="required"/>
                  </xs:extension>
                </xs:simpleContent>
              </xs:complexType>
            </xs:element>
          </xs:sequence>
        </xs:complexType>
      </xs:element>
      <xs:element name="au-direction-date" type="date-type" minOccurs="0"/>
      <xs:element name="au-exam-request-filing-date" type="date-type" minOccurs="0"/>
      <xs:element name="au-request-type" type="xs:string" minOccurs="0"/>
      <xs:element name="au-exam-request-status" type="au-exam-request-status-type" minOccurs="0"/>
      <xs:element name="au-deferment-date" type="date-type" minOccurs="0"/>
      <xs:element name="au-first-report-date" type="date-type" minOccurs="0"/>
      <xs:element name="au-further-report-date" type="date-type" minOccurs="0"/>
      <xs:element name="au-examination-section" type="xs:string" minOccurs="0"/>
      <xs:element name="au-acceptance-postponed" type="au-yes-no-type" minOccurs="0"/>
      <xs:element name="au-search-results-received-date" type="date-type" minOccurs="0"/>
      <xs:element name="au-third-party-examination-request-date" type="date-type" minOccurs="0"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
```

[RU comments]

Concerning the practice on citation references we could communicate that we don't have any specific tags for this purpose. If the reference is inside the paragraph of the document we use tags for text markup. If the reference appears in the bibliographic data section of the document (INID 56) we use the markup provided for INID 56 code.

[ST.36 TF's observations (draft)]

Regarding further XML tags to aid the creation of citation references, such as "example", the participants noted that the different XML tagging for citation references are available and has already been implemented by some IPOs, for example, ." <heading>Example</heading>" by the EPO and "<embodiments-example>" by the JPO.

- 12.5 note the use and examples of XML tagging of citation references (for example as provided by the EPO and available as Task Force background reference material at http://www.wipo.int/scit/en/taskfrce/citation_practices/background.htm).