User feedback on sequence listings – proposed ST. 26 standard – 2nd consultation round – summary of responses received

The EPO would like to thank the users who took part in the consultation.
The comments received have been grouped by topics as follows:

1. **Standard ST. 26 Main Body**

General comments: one user expressed the concern that the document is too complex to understand for the person skilled in the art (having a background in biology/chemistry but not a technical IT background).
EPO reply: The purpose of the ST.26 Standard is to allow full machine processing through the use of the XML language. The user is not required to fully understand the XML basis of the Standard but will be guided by the authoring tool which is being jointly developed by WIPO, EPO, USPTO, JPO and other offices.

a) **Exclusion of U (Uracil)**: one user expressed the concern that this will result in mislabelling errors of RNA as DNA.
EPO reply: A further aim of the ST.26 Standard is to allow smooth interaction with database providers which will electronically publish the contents of sequence listings. All major database providers (EBI, NCBI, DDBJ) follow the INSDC standard which does not allow "u" for uracil in nucleic acid sequences. Currently, already all "u"s in sequences are automatically converted into "t" before electronic publication by the database providers.

b) **Minimum requirements for compliance**: a section should be added in the Main Body to outline the minimum requirements for compliance in a clear and simple manner.
EPO reply: While acknowledging that a short section with minimum requirements would be desirable, it appears not possible to further limit the requirements set out in the ST. 26 Standard. That Standard has been drafted with the utmost effort to be as simple, as complete and as unambiguous as possible.

c) **Definitions**: the definition of the terms "must", "should" and "may" should be added to the Definitions section. ("must" means the indicated information is a requirement while "should or "may" means the indicated information is not a requirement).
EPO reply: The following has been added to the ST. 26 Standard and approved by the CWS in its fifth session on June 2, 2017:

4. For the purpose of this Standard, the word(s):
(a) "may" refers to an optional or permissible approach, but not a requirement.
(b) "must" refers to a requirement of the Standard; disregard of the requirement will result in noncompliance.
(c) "must not" refers to a prohibition of the Standard.
(d) "should" refers to a strongly encouraged approach, but not a requirement.
(e) "should not" refers to a strongly discouraged approach, but not a prohibition.

d) **Features keys and qualifiers**:

i. one user suggested that the qualifiers relating to geographical origins be removed as it would appear that they serve no particular purpose.
EPO reply: This qualifier is only optional and users are not required to use it.

ii. another user warned against the complexity of choosing the appropriate feature keys or qualifiers when preparing the sequence listing. A mistake could then be adversely interpreted against the applicant by an NPO. For this reason, the user recommended the use of feature keys and qualifiers with indicated uses that are not unduly burdensome and cannot be adversely interpreted and suggested that public databases implement the same feature keys/qualifiers. The following were cited as examples:

- "variation" as opposed to "misc_difference" means the indicated residue is a naturally occurring mutation or polymorphism
- VAR-SEQ vs VARIANT
- "MOD-RES" vs "SITE"

EPO reply: Exchange with database providers requires to follow the INSDC standard which prescribes certain feature keys and qualifiers.

e) **Use of lower case** to represent a nucleotide in a sequence: one user saw this requirement as too restrictive.
EPO reply: Use of lower case is prescribed by INSDC. The future authoring tool will automatically take care of this and other format requirements.

f) **Errors spotted**:

i. Two-letter language codes presented in upper case (section 46, examples 1 and 3, section 49, example) contrary to the requirement in ISO 639-1:2002 that lower case should be used.

ii. Annex VI is not listed in the Main Body as being annexed to it.
iii. The element “ApplicationFilingDate” cannot be mandatory if the application has not yet been filed.
iv. Use of < and > signs in the location example tables although paragraph 72 states that they must be replaced by the appropriate pre-defined entities.

EPO reply: These errors have been corrected in the current version of the Standard:
http://www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf

2. Guidance Document

a) **Complexity**: one user explained that the length of the Guidance Document by itself proves that the new standard is too complex for users.
EPO reply: The Guidance Document is meant as a reference document which also covers very rare and complex cases. In general, the consultation of the Guidance Document will not be needed. The authoring tool will also help greatly.

b) **Features keys and feature qualifiers**: users asked that the potential risk to use certain feature keys and feature qualifiers be identified and emphasized in the Guidance Document.
EPO reply: The Sequence Listings Task Force has taken account of this comment by drafting a dedicated “Added and deleted subject matter” document which highlights possible risks when transitioning from ST.25 to ST.26. This document is likely to be added to the ST. 26 Standard as a further Annex in one of the future CWS sessions.

c) **Errors spotted**: upper-case is used in the language codes in Annex VI-3.
EPO reply: These errors have been corrected in the current version of the Standard:
http://www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf

3. Authoring and validation tool

a) **Free text**: inclusion of a mechanism to identify and extract free text from sequence annotations in order to facilitate their inclusion in the description is seen as essential due to the need for translations in various languages upon entry into the national phase.
EPO reply: Such mechanism will be part of the authoring tool as it is a requirement under the PCT (Rule 5.2(b)).

b) **Stylesheets**:
   i. users warn against the risk of error in the conversion process to render the XML document in a human-readable format using stylesheets.
   ii. users suggested that an example stylesheet for the transformation of ST. 26 XML data into a human-readable form be included as an Annex to the Main Body of the standard.

   EPO reply: The XML document will always remain the legally binding text for sequence listings. Any rendering in a human readable format is only meant as a support but will have no legal effect. The Task Force is currently discussing whether to add a stylesheet as an Annex to the ST. 26 Standard.

c) **Availability of the tool**: users explained that they will need time to familiarise themselves with the tool and suggested that:
   i. it is made publicly available for open beta testing
   ii. the commencement date for use of ST. 26 be delayed until at least six months after public release of the final version.

   EPO reply: The first tests are planned for Autumn 2018. Test users are currently being selected. The tool will be available in good time before the ST. 26 Standard becomes mandatory under the PCT (namely as of January 2022 – tentatively).

4. **Further comments**:

a) Added matter:

i. Users expressed their concerns as to the risk of introducing "added matter" when converting an existing sequence listing from ST. 25 to ST. 26. The following applications were identified as particularly affected:

- New applications claiming priority from an earlier application including an ST. 25 sequence listing.
- New applications submitted under Art. 61(1)(b) EPC where the original European patent application includes an ST. 25 formatted sequence listing.
- Pending applications where an ST. 25 formatted sequence listing has already been submitted but which requires amendments or correction.
- Divisional applications where an ST.25 formatted sequence listing has been submitted for the earlier application.

ii. Users therefore request that the requirement to use ST. 26 when filing a sequence listing should start only for applications which date of priority (or date of filing if no priority is claimed) is after the start date.

EPO reply: See comments above under 2.b). Several transition scenarios are currently discussed between the Offices, see here: http://www.wipo.int/edocs/mdocs/pct/en/pct_mia_24/pct_mia_24_14.pdf

b) **Access to original sequence listing files**: users asked the EPO to make the original sequence listing files available online so that they can be readily obtained/downloaded by third parties without the need for a formal request to be submitted to the EPO.
EPO reply: The EPO will ensure that sequence listings are available to the public in the appropriate form.