[3510-16-P]

DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

37 CFR Part 1

[Docket No.: PTO-P-2012-2018]

Request for Comments on the Recommendation for the Disclosure of

Sequence Listings Using XML (Proposed ST.26)

AGENCY: United States Patent and Trademark Office, Commerce.

ACTION: Request for comments.

SUMMARY: The United States Patent and Trademark Office (Office) is seeking comments to obtain views of the public on the international effort to revise the standard for the presentation of nucleotide and/or amino acid sequences and the consequent changes to the United States rules of practice. The standard is being revised to require the use of extensible mark-up language (XML) format, to update the standard, and to more closely align requirements of the standard with those of public sequence database providers. Comments may be offered on any aspect of this effort.

DATES: Written comments must be received on or before [insert 60 days after

date of publication in the FEDERAL REGISTER] to ensure consideration. No

public hearing will be held.

ADDRESSES: Comments concerning this notice should be sent by electronic

mail message over the Internet addressed to seq_listing_xml@uspto.gov.

Comments may also be submitted by mail addressed to: Mail Stop Comments-

Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313–1450,

marked to the attention of Susan C. Wolski, Office of Patent Cooperation Treaty

Legal Administration, Office of the Associate Commissioner for Patent

Examination Policy. Although comments may be submitted by mail, the Office

prefers to receive comments via the Internet.

The comments will be available for public inspection at the Office of the

Commissioner for Patents, located in Madison East, Tenth Floor, 600 Dulany

Street, Alexandria, Virginia, and will be available via the Internet

(http://www.uspto.gov). Because comments will be made available for public

inspection, information that the submitter does not desire to make public, such as

an address or phone number, should not be included in the comments.

FOR FURTHER INFORMATION CONTACT: Susan C. Wolski, Office of

Patent Cooperation Treaty Legal Administration, Office of the Deputy

Commissioner for Patent Examination Policy, by telephone at (571) 272-3304, or by mail addressed to: Mail Stop Comments - Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, marked to the attention of Susan C. Wolski.

SUPPLEMENTARY INFORMATION:

1. <u>Background Information</u>

Patent applicants are currently required to submit biological sequence data in a standardized electronic format in accordance with World Intellectual Property Organization (WIPO) Standard ST.25, both within the framework of the Patent Cooperation Treaty (PCT) (Annex C of the Administrative Instructions) and under most national and regional provisions. The Rules of Patent Practice in the United States (37 CFR §§ 1.821-1.825) are substantively consistent with WIPO ST.25.

WIPO ST.25, which became effective in 1998 and has not been revised since that time, requires a flat file structure of numeric identifiers using a limited set of character codes. In October 2010, the Committee on WIPO Standards (CWS) established a Task Force, designating the European Patent Organization (EPO) as the lead, to propose a revised standard for the filing of nucleotide and/or amino acid sequence listings in XML format (hereinafter referred to as "the XML standard"). The work of the Task Force is accomplished through online

collaboration, restricted to Task Force members only, via the WIPO Web site. The XML standard (tentatively designated WIPO ST.26) is composed of three documents, namely, the main body of the standard, a first annex setting forth the controlled vocabularies for use with the sequence part of the standard, and a second annex setting forth the Document Type Definition (DTD) for the standard. Five rounds of comment/revision have taken place since March 2011, and discussion of the documents is ongoing.

It is expected that the XML standard will be adopted at a meeting of the CWS in early 2013. However, no decision has been made as to when it will enter into force for PCT purposes, and consequently, for national and regional offices. The work of the Task Force and issues pertaining to transitioning to the XML standard were discussed at the Nineteenth Session of the Meeting of International Authorities (MIA)(February 8-10, 2012). The Meeting agreed that the Task Force will look at the feasibility of developing a tool that would allow for the easy and complete conversion of sequence listings filed in one format (ST.25 or XML) into the other. Thereafter, the appropriate PCT bodies should commence a discussion on the most appropriate mechanism for transition from ST.25 to the XML standard. See the Meeting Summary available at

http://www.wipo.int/edocs/mdocs/pct/en/pct_mia_19/pct_mia_19_13.pdf.

2. Request for Comments.

The Office, leading the negotiations for the United States, is seeking public

comment on the current version of the main body of the standard and its two

annexes. The text of the current draft of the proposed main body of the sequence

listing standard, with its associated Annexes, is available via the Office's Web site

at http://www.uspto.gov/patents/law/comments/sequence_listings.jsp. The

documents are: Recommmendation for the disclosure of sequence listings using

XML (Proposed ST.26); Annex B.1. Controlled vocabularies; and

ST26SequenceListing-v1-0.

In light of the likely adoption of this standard in early 2013, the Office desires to

ensure that the XML standard is disseminated as widely as possible and the

opportunity to provide comments is correspondingly comprehensive. Written

comments may be offered on any aspect of the proposed standard or Annexes,

transition issues, or expected implementation in the United States. Comments are

specifically requested on the following issues:

(1) Comprehensiveness and Clarity. One goal of the development of a WIPO

Standard for sequence listings is to allow patent applicants to draw up a single

sequence listing in a patent application that would be acceptable for the purposes of

both international and national or regional prosecution. Any new standard should

represent the maximum requirements for any sequence listing submission, and each

national and regional office requiring compliance with the XML standard should have consistent interpretations of the standard.

The Office invites comments on whether the main body of the standard is sufficiently comprehensive and clear to achieve this goal, and in particular welcomes suggestions to add details or clarify the language as appropriate.

(2) <u>Absence of PCT Procedure</u>. Currently, ST.25 includes both substantive and procedural requirements for sequence listings because it fully incorporates PCT Adminstrative Instructions Annex C. The XML standard will be limited to substantive requirements, consisting of a general information part (to include information sufficient to identify the application of which it is a part) and a sequence data part (to include technical information pertaining to each sequence in the listing). It is expected that the PCT Administrative Instructions will be revised to separately specify procedural requirements pertaining to sequence listings filed under the XML standard in international applications. As an example, the XML standard itself will not require translation of free text that is in a non-English language to be included within the text of the application disclosure. However, international, national and regional offices would be free to make such a requirement.

The Office invites comments as to whether the XML standard includes any unnecessary procedural requirements or excludes any procedural requirements that should be retained.

(3) Feature Keys and Qualifiers. ST.25 uses a controlled vocabulary of feature keys to describe nucleic acid and amino acid sequences, with a very limited set of controlled vocabulary to further describe certain features. The XML standard includes, in addition to feature keys, a significant number of the qualifiers for the description of nucleotide sequences agreed upon by the International Nucleotide Sequence Database Collaboration (INSDC). Note that the XML standard does not include feature keys and qualifiers that are not relevant for patent data purposes. The XML standard also includes four qualifiers for amino acids. These feature keys and qualifiers form part of Annex B.1.

The INSDC revises feature keys and qualifiers on an occasional basis (i.e., there is no set schedule). While the goal of requiring INSDC feature keys and qualifiers is to improve compatibility with the public sequence database providers, it is not clear how often the international, national, and regional offices will be able to update submission software and procedures or rules to accommodate such changes.

Public comment is invited with regard to any feature keys or qualifiers that are not clear, or that are optional and should be mandatory (or vice versa). Comments are also welcome regarding how frequently WIPO should consider updating these

feature keys and qualifiers, recognizing the impact this will have on the Office

rules.

(4) Definition of a Sequence for which a Sequence Listing is Required. The XML

standard revises the definition of a sequence for which a sequence listing is

required. The following list sets forth the more significant differences from ST.25.

   (a) Prohibited sequences. ST.25 describes sequences for which a sequence

listing is not required, but does not specifically prohibit the presentation of such

sequences. In contrast, the XML standard (paragraph 4) prohibits the inclusion of

any branched nucleotide or amino acid sequences or any sequences with fewer than

ten specifically defined nucleotides or fewer than four specifically defined amino

acids.

   (b) Modified nucleotides. ST.25 specifies that sequences comprising

nucleotides other than those listed in Appendix 2, Tables 1 and 2, are specifically

excluded from the definition of sequences for which a sequence listing must be

provided. In contrast, the XML standard (paragraph 5) does not contain this

exclusion, and specifies inclusion of sequences containing any nucleotides that can

be represented using any of the symbols set forth in Annex B.1, paragraph 1, Table

1. This includes nucleotides that may contain, inter alia, a modified or synthetic

purine or pyrimidine base; a modified or synthetic ribose or deoxyribose, or a

8

modified or synthetic 3' to 5' internucleotide linkage, i.e., any chemical moiety that provides the same structural function as the phosphate moiety of DNA or RNA.

(c) <u>D-amino acids</u>. ST.25 specifies that sequences containing at least one D-amino acid are specifically excluded from the definition of sequences for which a sequence listing must be provided. In contrast, pursuant to the XML standard (paragraph 6), any unbranched sequences containing four or more specifically defined amino acids would be required to be included in a sequence listing, regardless of whether that sequence contains any D-amino acids.

(d) <u>Variants</u>. ST.25 does not specifically address how variants are to be represented in a sequence listing. The XML standard specifies how variants are to be represented. See paragraph 58 which reads as follow:

> 58. A variant sequence disclosed by enumeration of its residues and encompassed by paragraph 4 must be assigned its own sequence identification number and be presented in the sequence listing. A specific variant, i.e., deletion, addition, or substitution, disclosed only by reference to a disclosed primary sequence in the sequence listing, must be presented in the sequence listing either as a separate sequence assigned its own sequence identification number or by annotation of the primary sequence with appropriate feature keys and qualifiers. A specific variant containing multiple variations from the primary sequence at distinct locations, where

the variations at each location only occur together, must be presented in

the sequence listing as a separate sequence assigned its own sequence

identification number.

The Office requests comments on whether these changes as set forth in paragraphs

(a) through (d) above are desirable, and what difficulties, if any, are likely to be

faced in complying with the definition in the XML standard.

(5) Publications (references). ST.25 provided for the inclusion of publication

information (i.e., references to relevant prior publications) in numeric identifiers

<300> - <312>. The XML standard does not provide for such references.

The Office invites comments as to whether there is any perceived detriment due to

the non-inclusion of such publications or references in the sequence listing.

(6) Transition Issues. Transition to the XML standard will require Office analysis

of the time frame required to update systems to receive, process, and search

sequence listings filed under the standard. The date of entry into force may be as

long as two or three years after adoption of the new standard; however, as noted

above, discussions within the Task Force and WIPO are continuing. Current

thinking in the Office is that it would be preferable to have a clean transition from

current WIPO Standard ST.25 to the XML Standard. This would be accomplished

by having the XML Standard enter into force for PCT purposes (and

correspondingly, for US applications at the same time) at a particular date in the future (e.g., after such date, all sequence listings filed for the first time in an application (including a continuation, continuation-in-part, and a divisional) would have to be filed in compliance with that new standard).

(a) The Office invites comments regarding how much time is likely to be needed for applicants to transition to the XML standard (with the assumption that sequence listing authoring software will be publicly available).

(b) Given the divergent requirements of the proposed XML standard and ST.25 as described above, the Office invites comments on what difficulties an applicant should anticipate if national or regional offices required compliance with different standards (i.e., ST.25 and XML). Will the existence of separate authoring tools for each of the standards mitigate such difficulties?

Date:__May 9, 2012_____

_____
David J. Kappos
Under Secretary of Commerce for Intellectual Property and
 Director of the United States Patent and Trademark Office

[FR Doc. 2012-11755 Filed 05/14/2012 at 8:45 am; Publication Date: 05/15/2012]