

**Notice of Complex Work Unit Pilot Program**  
**Request for Voluntary Submission of Electronic Complex Work Unit**  
**Files**

**Issue**

Applications containing complex work units such as chemical structures, mathematical formulae, and tables often add significant complexity and cost to the current processes for publishing patents and pre-grant patent application publications. This is largely due to the fact that the Office's publisher must take the complex work unit as filed in paper and convert it to an electronic format that is appropriate for publication. Further, especially in the realm of chemical structures, it is sometimes necessary for the Office to recreate the complex work unit in order to search relevant databases. Accordingly it would be desirable for the Office to obtain an electronic version of any complex work unit that is in a format compatible with Office electronic systems.

**Background Information**

*Tables*

In view of the processing problems caused by the submission of applications containing paper or image versions of lengthy tables, the Office continues to encourage applicants to submit lengthy tables on compact disc (CD) in ASCII text as provided in 37 CFR 1.52(e) or via EFS-Web as provided in the EFS-Web Legal Framework, as a part of the specification as provided in 37 CFR 1.77(b)(5). However, the Office continues to receive a significant number of patent applications that include lengthy table information submitted only on paper. The Office has an ongoing pilot program for patent document publishing purposes only wherein, in response to an Office request, applicants may submit lengthy tables in electronic formats other than ASCII text so the lengthy tables can be posted on a separate repository (PSIPS) instead of being published in the full text searchable databases. See Notice that Lengthy Tables will be Published Differently in Patents, Statutory Invention Registrations and Patent Application Publications And Notice of Pilot Program With Respect to Lengthy Tables, 1282 O.G. 108 (2004).

*Chemical and Three-Dimensional Biological Structure Data*

The Office published an Advance Notice of Proposed Rule Making (ANPRM) entitled "Acceptance, Processing, Use and Dissemination of Chemical and Three-Dimensional Biological Structural Data in Electronic Format," 70 Fed. Reg. 35571 (21 June 2005) seeking public comment on how best to provide the capability to manage, process, search, and disseminate 3-D biological structure data and chemical structure data as appropriate. In addition to general background information, the ANPRM provided background information specific to 3-D protein crystal data and chemical structure data. In particular, an effort was made to familiarize the public with standards and tools that are currently available to provide a frame of reference.

The USPTO anticipates that submission of chemical and 3-D biological structural data in electronic format in patent applications will improve the processing and examination of patent applications that include such data, as well as the publication of such data. Public comments in response to the ANPRM with regard to 3-D protein structure data were generally favorable, with most respondents favoring requiring electronic submission of such data, as it would benefit both the public and the Office without placing an undue burden on applicants, because most structures were already stored in electronic form. While public comments were mixed with regard to requiring the submission of chemical structures in electronic form, the respondents acknowledged that chemical structures appearing in patent applications are usually drawn using either free or commercially available software, and thus are created and stored in electronic form.

### **EFS-Web**

With the upgrade to EFS-Web 1.1, the Office is now able to accept text files as part of a patent application submission. Text-based formats are currently available for several of the types of complex work units frequently submitted by applicants. For example, chemical structure information can be created or saved as IUPAC International Chemical Identifier (InChI™) files; mathematical equations can be created or saved using the Mathematical Markup Language (MathML); and three-dimensional protein crystalline structures can be created or saved using the text-based, Protein Data bank (PDB)-compliant format. InChI™, MathML, and PDB formats are all comprised of ASCII text strings. As an added benefit, the worldwide PDB has indicated a willingness to start a pilot program to accept, store, and disseminate information from published patent applications which have been received by the USPTO in PDB format, a format commonly used by practitioners in this art.

### **Compact Disc (CD) Submissions**

With the advent of the Supplemental Complex Repository for Examiners (SCORE) system, the Office is able to receive files in ASCII text through EFS-Web, or in a wide variety of format types via compact disc (CD) submission, and make these files available to examiners through the Electronic Desktop Application Navigator (eDAN) and to practitioners and the public through a link in the Patent Application Information Retrieval (PAIR) system. However, the Office has not yet determined which file formats would improve the efficiency of the publication and/or examination processes.

### **Pilot to Accept Electronic Copies of Complex Work Units**

Beginning the date of the publication of this notice in the Official Gazette, the Office will begin a pilot program to accept voluntary submission of tables, chemical structures, three-dimensional protein crystalline structures and mathematical formulae recorded in electronic formats not compliant with 37 CFR 1.52(e) on CD as a supplement to (not as a replacement of) submission of these complex work units on paper or as a PDF file via EFS-Web.

Effective immediately, applicants may and are encouraged to electronically submit text documents in InChI<sup>TM</sup>, MathML, and PDB formats via EFS-Web as separate text files as a supplement to (not as a replacement of) submission of these complex work units as a PDF filed via EFS-Web.

#### Pilot Program Requirements:

- The patent application submitted on paper or as a PDF file via EFS-Web must include therein at least one complex work unit, i.e., at least one table, chemical structure, three-dimensional protein crystalline structure, or mathematical formula.
- Any CD or text file submitted in accordance with this pilot must be accompanied by a statement that the complex work unit file submitted in electronic form is the same as the file used to create the image of the complex work unit submitted as part of the paper patent application or PDF format patent application filed via EFS-Web. In the event of any discrepancy between the paper or PDF and the electronic version of a complex work unit, the paper or PDF version is the authoritative version.
- Each CD must be labeled "CWU Pilot" and only one copy of each CD is required. Each file must be labeled in accordance with the instructions given on the USPTO web site at: <http://www.uspto.gov/web/patents/cwupilot.html>. The requirements of 37 CFR 1.52(e) are hereby waived to the extent that they are inconsistent with the complex work unit pilot program requirements.

#### Duration and Results of Pilot Program

The pilot program is scheduled to run for six months. After the completion of the pilot the Office will evaluate the findings of the pilot and make a determination as to whether or not to (1) implement procedural changes based on the findings of the pilot program or (2) extend the pilot program for an additional length of time.

#### Exclusions from Pilot - Computer Program Listings, Sequence Listings, Tables in ASCII Format, International Applications

Computer Program Listings, Sequence Listings, or tables in ASCII format must be submitted on a separate set of CDs and comply with 37 CFR 1.52(e) or must be submitted in accordance with the EFS-Web Legal Framework.

In addition, international applications filed under the Patent Cooperation Treaty (PCT) are excluded from this pilot program. Thus, PCT applicants may not file complex work units on CD as a supplement to a patent application filed either on paper or via EFS-Web. Furthermore, applicants may not file text documents in InChI<sup>TM</sup>, MathML, or PDB formats via EFS-Web in international applications filed under the PCT.

## **Effect on Calculation of Application Size Fee**

The application size fee set out in 35 U.S.C. 41(a)(1)(G) applies to any application the specification and drawings of which exceed 100 sheets of paper.<sup>1</sup> However, electronic files submitted in accordance with this pilot are the source code for the complex work units that are included in the specification or drawings and thus are not considered to be a part of the specification or drawings. Accordingly, the following documents, when submitted in accordance with this pilot, will not be included for purposes of determining the application size fee:<sup>2</sup> tables, chemical structures, three-dimensional protein crystalline structures and mathematical formulae recorded on CD in electronic formats not compliant with 37 CFR 1.52(e); and chemical structures, mathematical formulae, or protein crystalline structures expressed as a text document in InChI<sup>TM</sup>, MathML, or PDB format and submitted via EFS-Web.

## **Further Legal Consideration**

Currently, 37 CFR 1.52(e) only allows an applicant to file Computer Program Listings, Sequence Listings, or tables on compact disc (with a duplicate compact disc) in ASCII format. Accordingly, for purposes of this pilot program, the Office hereby waives the requirements of 37 CFR 1.52(e) to the extent required to allow applicants to submit chemical structures, mathematical equations, three-dimensional protein crystalline structures, and tables not in compliance with 37 CFR 1.52(e) in electronic format on compact disc as a supplement to a paper copy of a patent application or a PDF copy of a patent application filed via EFS-Web.

Information obtained during the pilot will allow the Office to determine the feasibility of accepting and using chemical structures, tables, mathematical equations, and three-dimensional protein crystalline structures in specific electronic formats. Upon successful completion of the pilot, the Office will assess the desirability of proposing rules pertaining to the submission of chemical structures, tables, mathematical equations, and three-dimensional protein crystalline structures in electronic formats that are most useful to the Office (analogous to the current nucleic acid and amino acid sequence rules).

## **Detailed Information**

More detailed information pertaining to this pilot is posted on the USPTO web site at: <http://www.uspto.gov/web/patents/cwupilot.html>. This page includes links to the ANPRM and OG notices referenced above, a detailed description of InChI<sup>TM</sup>, MathML, and PDB formats, an explanation of the file formats currently used in the publication process, preferred naming conventions for the electronic files, and answers to frequently asked questions.

---


<sup>1</sup> Any sequence listing or computer program listing filed in electronic medium is excluded from this calculation.

<sup>2</sup> The application size fee is codified in the Code of Federal Regulations at 37 CFR 1.16(s).

**Contact Information**

Questions regarding this notice may be directed to Lisa Titcomb, CWU Pilot Program Manager, by phone at (571) 272-5723 or by e-mail at [lisa.titcomb@uspto.gov](mailto:lisa.titcomb@uspto.gov).

Date: 4/14/08

  
\_\_\_\_\_  
JON W. DUDAS  
Under Secretary of Commerce for Intellectual Property and  
Director of the United States Patent and Trademark Office