

From: carl oppedahl
To: [Fee.Setting](#)
Subject: USPTO must not impose the \$400 penalty on non-DOCX filings
Date: Monday, August 5, 2019 5:28:38 AM

USPTO proposes to charge a \$400 penalty for filing a patent application in non-DOCX format. This is a very bad idea, for reasons that I will discuss in detail. Only if USPTO were to make fundamental changes in its way of receiving DOCX files would it be acceptable for USPTO to impose a penalty for filing in a non-DOCX format.

USPTO needs to follow WIPO's example, permitting the practitioner to file a "pre-conversion format" version of a patent application along with the DOCX file. In the event of some later problem with USPTO's rendering of the DOCX file, the practitioner would be permitted to point to the pre-conversion format, which would control in the event of any discrepancy.

By way of background, the normal way to file US patent applications is in PDF format. With PDF format, the applicant has complete control over the appearance of characters and symbols.

Some years ago, the USPTO began beta-testing a system that would permit a practitioner to file a patent application in DOCX format instead of in PDF format. The undersigned was among the very first of the beta-testers of USPTO's system for DOCX filings. As implemented by the USPTO, the practitioner would upload a DOCX file, and USPTO would render the DOCX file in a human-readable PDF image format. As part of the e-filing process, the practitioner was expected to proofread the rendered image as provided by the USPTO's e-filing system. The notion was that the practitioner would be obliged to catch any instances of USPTO's system rendering the DOCX file differently from the way the practitioner's word processor had rendered that same DOCX file. If, for example, some math equation or chemical formula had gotten corrupted in USPTO's system, the practitioner would be expected to catch this *prior to* clicking "submit".

A first difficulty about this is that there is no single unambiguous thing called "DOCX" format. The history may be seen in the Wikipedia article here: https://en.wikipedia.org/wiki/Office_Open_XML DOCX exists in many variants, and in particular Microsoft has a history of making poorly

documented changes over time to the ways that Microsoft Word implements DOCX formatting of documents.

USPTO inaccurately characterizes DOCX as if one could be sure that any word processor will implement DOCX in the same way as any other word processor. For example, USPTO says:

There are several word processors that can create and save in DOCX format, including Google Docs, Microsoft Word 2007 or higher, Office Online, LibreOffice, and Pages for Mac.

That statement is disingenuous at best, and borders upon falsity given that there is no single unambiguous DOCX format. A more accurate statement would be:

There are several word processors that can create and save documents in variants of DOCX formats, including Google Docs, Microsoft Word 2007 or higher, Office Online, LibreOffice, and Pages for Mac.

USPTO also says:

DOCX is stable and governed by two international standards (ECMA-376 and ISO/IEC 29500).

This statement is simply false. There is no single DOCX standard to which Microsoft Word and the other word processors are all compliant.

To give a simple example, consider this math equation in a patent application that I recently filed as a PDF-based PCT application using Libre Office:

$$f(u) = \cos(u)^3 \exp(0.2u) \quad \text{(Equation 14)}$$

click to enlarge

As an experiment I uploaded the DOCX file of this PCT application to EFS-Web as if I were filing a domestic US patent application. The way the USPTO has designed EFS-Web, what happens next is that the practitioner sees this message in red letters:

The PDF(s) have been generated from the docx file(s). Please review the PDF(s) for accuracy. By clicking the continue button, you agree to accept any changes made by the conversion and that it will become the final submission.

It is easy to see that this filing procedure, as contemplated by USPTO, imposes an enormous professional liability risk on the practitioner. The practitioner is obligated to proofread the entire patent application, from top to bottom, for any corruption introduced by the USPTO's rendering system.

Here is how the USPTO rendered this math equation:

$$f(u) = \cos(u)^3 \exp(10.2u) \quad \text{(Equation 14)}$$

click to enlarge

The alert reader will notice that the USPTO inserted a spurious digit "1" into the math equation. Had I overlooked this corruption of the document by the USPTO, I might then have clicked "continue", at which point it would have been USPTO's position that I had agreed to accept USPTO's change of "0.2" to "10.2". TYFNIL the accused infringer would be able to seize upon this.

There are a dozen other places in this patent application where USPTO corrupted math equations; Equation 14 is merely the most striking so that is the one that I quoted here.

As a beta-tester of USPTO's DOCX systems, I have used a pretty simple way of choosing which of my patent applications I am willing to subject to the risks of filing in DOCX. Basically if there is any math equation or chemical formula, or anything other than very simple alphanumeric characters, I don't take the risk. Every now and then, on a whim, I will experiment with something like this "Equation 14" document, but I don't risk any actual substantive rights of a client by actually clicking "submit" in such a case.

But USPTO's proposed rulemaking would put me in the untenable position of having to pay a \$400 penalty for every case that I file that has a math equation or chemical formula in it.

If USPTO wants to pursue this, USPTO should follow the example of the World Intellectual Property Organization (WIPO). Like the USPTO, WIPO of course encourages practitioners to e-file using characters rather than images. Clearly all forward-thinking patent offices need to consider ways to try to collect characters, because that is more efficient in later workflow than collecting page images.

But what does WIPO do so that practitioners are protected from the kind of risks that we see above with Equation 14? WIPO permits the applicant, at the time of filing an international patent application, to provide not only the character-based version of the patent application (XML, in the case of PCT), but also the “pre-conversion format” of the document. You can see this in [Section 706](#) of the PCT Administrative Instructions. The idea is that if later it turns out that some flaw arose in the generation of the XML file, or some flaw in the way the XML got rendered into human-readable form, the applicant would be able to point to what the application looked like in its “pre-conversion format”.

It’s clear from this the simple thing that USPTO would need to do, as a precondition to imposing a \$400 penalty for non-DOCX filings, is to make a provision for the practitioner to be able to provide a PDF version of the patent application being filed, along with the DOCX file. This PDF version would serve as the controlling version in the event that (for example) the USPTO ended up inserting a spurious “1” into a math equation.

We can then circle around to the USPTO’s disingenuous statements about DOCX. If it were really true that there is some single unambiguous DOCX standard, then this spurious “1” would never have gotten inserted into the rendered patent specification in EFS-Web. The very fact that this happened proves that USPTO is wrong when it suggests that there is some single thing called DOCX that means the same thing in EFS-Web and in all word processors.

USPTO must scrap its planned \$400 penalty for non-DOCX filing, or must provide for the filing of "pre-conversion format" documents which will control in the event of any discrepancy in USPTO's rendering of the DOCX file.