KILPATRICK TOWNSEND
ATTORNEYS AT LAW

KILPATRICK TOWNSEND & STOCKTON LLP
www.kilpatricktownsend.com

Two Embarcadero Center, Suite 1900
San Francisco, CA 94111
t 415 576 0200 f 415 576 0300

October 27, 2023

direct dial (415) 273-4301
ccretsinger@kilpatricktownsend.com

Comment on USPTO Submission to OMB Regarding DOCX Submission Requirements

This comment is submitted on behalf of Kilpatrick Townsend & Stockton LLP, a national law firm that files over 3,000 US patent applications per year on behalf of clients ranging from individuals and startups to Fortune 500 companies.

As patent attorneys and agents, our first duty is to protect our clients' interests. When we prepare patent applications for our clients, we strive to make sure that the text is complete and accurate because we know that even small errors can render a patent worthless. We count on the Patent and Trademark Office ("USPTO") to create a record that accurately reflects the text we intended to file.

Since approximately 2007, we have been electronically filing patent applications in PDF format using longstanding USPTO processes (the "PDF filing process"). The PDF filing process accurately captures the original application document as we uploaded it, and the PDF standard provides a predictable, reliable rendering of that document over the lifetime of the patent.

The USPTO is now seeking OMB approval for a proposed Information Collection[1] that would compel applicants to begin electronically filing text portions[2] of most patent applications in DOCX format (the "DOCX filing process"), under penalty of a surcharge of $400 per application for each application not filed in DOCX format (the "non-DOCX surcharge"). While the PDF filing process would remain available, applicants who are unable to file in DOCX format, or unwilling to take the associated risks, would incur considerable financial costs.

We urge OMB to reject the proposed Information Collection, for the following reasons:

1. DOCX is not capable of providing the same predictable, reliable rendering of patent documents that the PDF standard provides. To overcome the deficiencies of DOCX, the USPTO would need to publish a standard that establishes a definitive rendering of a DOCX file into a human-readable document and/or guarantee that applicants will continue to have the right to submit a PDF version of the application together with the DOCX file, with applicants having the

[1] USPTO "DOCX Submission Requirements," ICR Ref. No. 202309-0651-003, posted at https://www.reginfo.gov (September 27, 2023); "Comment Request; DOCX Submission Requirements," 88 F.R. 66414 (September 27, 2023).
[2] The "text portions" of a patent application are the specification, claims, and abstract.

right to rely on the PDF version as proof of the content that applicants submitted. The USPTO refuses to do either.

2. The USPTO significantly underestimates the burden imposed on patent applicants. Because applicants cannot know in advance how the USPTO will render their DOCX files and because rendering errors can be subtle (e.g., single characters), applicants must undertake detailed proofreading of the USPTO-rendered version of each application, a task that on average will take considerably more than the USPTO's estimated 30 minutes per application. The USPTO has failed to show that this burden is cost-justified.

The proposed Information Collection cannot be relied on to create an accurate record of the content of patent applications.

Patent applications are important legal documents that are required to clearly and accurately describe the claimed subject matter.[3] Errors in applications can result in loss of patent rights. After an application is filed, correction of errors is limited by what is in the USPTO's records.[4] If the USPTO alters the content of an application and fails to preserve the application's original content in the record, correction of errors may be impossible and applicants may forfeit important rights. This would be manifestly unfair to patent applicants.

This manifest unfairness is the risk created by the proposed Information Collection. The DOCX filing process does not place the application's original content (the DOCX file applicant uploads) into the record. Instead, the USPTO executes an automated "validation" process that checks for unacceptable features in the DOCX file and ***alters*** the DOCX file in an attempt to remove or replace unacceptable features.[5] It is the output of this validation process, which applicants do not control and have only partial information about, that becomes the authoritative source document.[6] In some instances, the validation process substantively changes the content of the document. For instance, documents or portions thereof may be reformatted in a manner that obscures intended meaning, or characters may be replaced with other characters that signify something different from what was intended.[7]

---

[3] *See* 35 U.S.C. §112(a) and (b).
[4] *See* 35 U.S.C. §132(a).
[5] *See* https://www.uspto.gov/sites/default/files/documents/DOCX_Feedback_Errors_and_Warnings.pdf (partial list of unacceptable items and features, with notes that certain content items are removed or converted to acceptable items).
[6] USPTO "Supporting Statement," in "DOCX Submission Requirements," ICR Ref. No. 202309-0651-003, posted at https://www.reginfo.gov (September 27, 2023) (hereinafter "USPTO Statement"). Other commenters have suggested that the USPTO should provide the source code for the validation and rendering processes. The USPTO has declined to do so. *See* USPTO Statement, Comment 6 and Response to Comment 6, at pp. 9-10 (all page numbers in the USPTO Statement refer to the DOCX file as rendered on the author's computer system, which may or may not match what OMB sees).
[7] For specific examples, see Comment by American Intellectual Property Law Association, August 7, 2023, posted at https://www.regulations.gov/docket/PTO-P-2023-0031/comments (comment ID PTO-P-2023-0031-0006) (hereinafter "AIPLA Comment"), pp. 30-37; Comment by 152 Patent Practitioners,

While applicants may be able to detect changes made by the USPTO's validation process prior to filing (e.g., by comparing the validated DOCX file to the uploaded DOCX file), it is not always apparent how to modify the file to prevent the validation process from changing it. Time constraints prohibit a back-and-forth with the USPTO to troubleshoot the process and await a resolution.[8] Consequently, where the validation process makes substantive changes, applicants cannot file in DOCX. Under the proposed Information Collection, these applicants would have to pay a $400 surcharge because of a defect in the USPTO's software. Forcing a member of the public to pay a fee for being unable to comply with information collection requirements due to defects in agency software is not acceptable.

In addition, a DOCX file is computer code, essentially meaningless until it is rendered into a human-readable form. Accordingly, after validation, the USPTO converts the (possibly-altered) DOCX file to PDF using a rendering process that, again, applicants do not control. This PDF file, which may not accurately reflect the substance of what the applicant uploaded, becomes the "record copy" of the application.[9]

The fundamental problem here is that — unlike with PDF[10] — there is no universally agreed-upon standard that defines "correct" rendering of DOCX into human-readable form. As other commenters have explained to the USPTO,[11] DOCX is a proprietary format developed by Microsoft Corporation ("Microsoft"). DOCX is based on OOXML but includes a large number of extensions, some of which are proprietary to Microsoft and any of which can be changed at any time at the sole discretion of Microsoft.[12] Other software developers that support interoperation with DOCX files make educated guesses as to how to interpret the proprietary elements or else ignore them, with the result that a DOCX document looks different depending on which software renders it.[13] Some of the differences are cosmetic; others are substantive.

PTAARMIGAN, August 7, 2023, posted at https://www.regulations.gov/docket/PTO-P-2023-0031/comments (comment ID PTO-P-2023-0031-0008) (hereinafter "PTAARMIGAN Comment"), pp. 77-82.

[8] Patent rights depend on the filing date of the application. *See* 35 U.S.C. §102. Any delay in filing due to technical issues may jeopardize an applicant's rights. Further, as discussed below, even if a particular filing date is not critical, the time spent on troubleshooting would usually cost applicants more than $400.

[9] USPTO Statement, Response to Comment 10, p. 11; "Filing Patent Applications in DOCX Format," 87 F.R. 25226 (April 28, 2022); "Submitting Patent Applications in Structured Text Format and Reliance on the Text Version as the Source or Evidentiary Copy," 86 F.R. 29571 (June 2, 2021).

[10] PDF is an open standard published by the International Organization for Standardization. The official definition is provided in "ISO 32000-2:2020 — Document Management — Portable document format — Part 2: PDF 2.0," available at https://www.iso.org/standard/75839.html.

[11] *See, e.g.,* Comment by Carl Oppedahl, August 6, 2023, posted at https://www.regulations.gov/docket/PTO-P-2023-0031/comments (comment ID PTO-P-2023-0031-0004) ("Oppedahl Comment"), pp. 3-23.

[12] Microsoft publishes a list of its extensions to the OOXML standard at https://learn.microsoft.com/en-us/openspecs/office_standards/ms-docx/b839fe1f-e1ca-4fa6-8c26-5954d0abbccd. *See also* Oppedahl Comment at pp. 11-16.

[13] See n.7 *supra*.

The USPTO has not published its own definition of "DOCX format" or made its DOCX-to-PDF rendering software available to the public.

Instead of providing a definition of "DOCX format" for the purpose of patent application filing, the USPTO continues to assert that "DOCX is a word-processing file format that is part of OOXML, an XML-based open standard approved by the Ecma International® consortium and subsequently by the ISO/IEC joint technical committee" and that "[l]ike the PDF standard, DOCX presents documents, including text and formatting, in a manner that is independent of software, hardware, or operating system."[14]

This is simply false. DOCX is based on OOXML but is not "part of OOXML" or of any other standard.[15] It has also been demonstrated empirically that rendering of DOCX into human-readable documents is dependent on the software and hardware used.[16] In the absence of a published standard definition of "DOCX format," it is not clear how any dispute between an applicant and the USPTO over how a DOCX file should be rendered into a human-readable document would be resolved. Apparently, the USPTO does believe that there is a "correct" rendering, since it concedes that is rendering process has made errors in the past.[17] However, the USPTO has failed to set forth any definition of a "correct" rendering or any standards for proving that a rendering error exists.

To summarize, the DOCX filing process results in a USPTO record that contains two versions of the application, neither of which is the file that the applicant actually uploaded and either or both of which may be substantively different from what applicant actually uploaded. This is clearly unacceptable.

The USPTO has failed to address commenters' concerns on these points. It dismissively summarizes multiple comments outlining the technical issues as: "Several commenters asserted that the USPTO conversion tool and validation system are 'unreliable and error laden.'"[18] Its only response is that "[t]he system immediately detects and supplies the applicant with useful error and warning messages, allowing for adjustments to patent applications earlier in the process."[19] In fact, the system does more than supply messages. The "validation" process also modifies the DOCX file, which can and sometimes does result in substantive alterations.

In response to comments "express[ing] concern that errors would be introduced into applications submitted in the DOCX file format because USPTO systems are unable to recognize various technical symbols and characters, and thus may render them incorrectly,"[20] the USPTO

---

[14] USPTO Statement, Response to Comment 2, p. 7.
[15] For a thorough debunking of the USPTO's claims that DOCX is a standard, see Oppedahl Comment at pp. 3-27.
[16] See n.7 *supra*.
[17] USPTO Statement, Response to Comment 3, p. 8.
[18] USPTO Statement, Comment 4, p. 9.
[19] USPTO Statement. Response to Comment 4, p. 9.
[20] USPTO Statement, Comment 3, p. 8.

asserts that "Recent results have shown ***very few*** issues"[21] and that "[t]he USPTO has addressed many past issues."[22] The USPTO proceeds to cite examples of rendering errors that have been corrected, including one that was corrected ***last month***.[23]

Far from reassuring OMB or the public that the DOCX filing process can be relied upon, the USPTO's responses reveal that its software for processing uploaded DOCX files is not stable or reliable. "Very few issues" is not "no issues," and the USPTO is apparently still fixing bugs. Under these circumstances, it is at best premature to require applicants to either rely on the USPTO's software or pay a $400 surcharge.

The USPTO has also stated that out of 183,685 applications filed in DOCX to date, "only five petitions have been received related to a conversion of data type error."[24] This proves little. The five petitions represent instances where an error occurred but the applicant went forward with DOCX filing anyway, presumably because the applicant did not notice the error until after filing. It does not (and cannot) indicate the number of instances where an applicant attempted a DOCX filing, encountered an error, and reverted to PDF filing. It also does not reflect instances where errors occurred that the applicants have not yet noticed or have not cared about, or instances where the applicants were unaware that they could petition for correction.

Frequent patent filers have expressed — repeatedly and emphatically — that they regard the USPTO's DOCX filing process as unreliable and a risk to applicants' patent rights.[25] As the USPTO's statistics reflect, few applicants are using this unreliable and risky process. OMB should take the public's concerns seriously and reject the USPTO's attempt to force an unreliable and risky process on an unwilling public.

The USPTO's "Auxiliary PDF" option provides insufficient protection against USPTO error.

As a safeguard against the risks of the DOCX filing process, the USPTO currently offers applicants the option to upload an applicant-generated PDF version of the application text (referred to as an "auxiliary PDF") along with the DOCX version, without incurring additional

---

[21] USPTO Statement, Response to Comment 3, p. 8 (emphasis added). Note that "very few issues" is an admission that at least some issues still exist.
[22] *Id.*
[23] *Id.* ("As another example, in September 2023, the USPTO resolved an issue with images, including images of equations and formulas embedded as Scalable Vector Graphics (SVG) format, that were not supported in multi-section DOCX documents.")
[24] USPTO Statement, Response to Comment 3, p. 8.
[25] *See, e.g.,* all nine comments posted at https://www.regulations.gov/docket/PTO-P-2023-0031/comments, responsive to "Comment Request; DOCX Submission Requirements," 88 F.R. 37039 (June 6, 2023). Note that two of these comments include large numbers of signatories, and a third was submitted by AIPLA, the nation's leading IP law organization. We submitted a comment on behalf of the approximately 150 registered patent attorneys and agents who practice at Kilpatrick Townsend & Stockton LLP.

fees.[26] Under current policy, if an auxiliary PDF is uploaded, the USPTO will make the auxiliary PDF available throughout the lifetime of the application (and eventual patent) and will allow applicants to refer to the auxiliary PDF to support requests to correct errors in the official application text generated from the DOCX file.[27] The USPTO says it intends to treat the auxiliary PDF as "an ongoing safeguard should any unexpected conversion discrepancies occur during the filing process,"[28] which would alleviate at least some concerns regarding the risk of uncorrectable error being introduced by the USPTO.

However, as things stand, the auxiliary PDF is not a sufficient safeguard. The USPTO has made no commitment to keeping this option available for future application filings. Instead, the USPTO is only allowing applicants to upload an auxiliary PDF without incurring additional fees "until further notice."[29] "Further notice" could presumably be published at any time. For example, the USPTO could obtain OMB clearance for the non-DOCX surcharge, then terminate the auxiliary PDF option a week or a month later. If that happens, applicants will face a choice of paying extra fees or taking the risk that the USPTO will substantively alter their application text and leave applicants without a way to correct the record.[30] This cannot be allowed.

The USPTO has failed to address this concern in its response to comments, stating only that "the USPTO has extended indefinitely the option to submit an applicant-generated PDF of the application along with the validated DOCX file(s)."[31] "Indefinitely" is no more of a commitment than "until further notice."

In the absence of any commitment by the USPTO to providing ongoing and sufficient safeguards against USPTO-introduced errors in future application filings, OMB should reject this Information Collection Request.

The proposed Information Collection is unduly burdensome and not cost-justified.

In addition to being unreliable and risky, the proposed Information Collection is unduly burdensome and not cost-justified. The USPTO estimates the time burden of DOCX filing at 0.5

[26] "Filing Patent Applications in DOCX Format," 87 F.R. 25226 (April 28, 2022) (initial announcement of auxiliary PDF option, available "on a temporary basis"). The fees waived by the USPTO include any applicable application size fees under 37 C.F.R. §1.16(s), as well as the non-DOCX surcharge under 37 C.F.R. §1.16(u), assuming that the surcharge takes effect before the fee waiver is rescinded.
[27] "Extension of the Option for Submission of a PDF With a Patent Application Filed in DOCX Format," 88 F.R. 37036 (June 6, 2023).
[28] USPTO Statement, Response to Comment 10, p. 12.
[29] 88 F.R. 37036 (June 6, 2023).
[30] If an auxiliary PDF is not filed, the only documents in the record are the USPTO's validated DOCX file and the PDF generated from the USPTO's validated DOCX file. As discussed above, there is no guarantee that either of these documents will accurately reflect what the applicant uploaded.
[31] USPTO Statement, Response to Comment 10, pp. 11-12.

hours per application.[32] In our experience, the time burden is roughly seven times higher. And the USPTO has not identified sufficient cost savings to balance this burden.

As discussed in our previous comment,[33] preparing a DOCX file for upload to the USPTO requires more work than simply printing to PDF.[34] In addition, some DOCX files fail to upload or are altered by the USPTO's validation process. When this happens, we face the choice between engaging in a potentially lengthy troubleshooting process or cutting our losses and filing in PDF format. In the real world, patent applications are often filed under tight deadlines; we and other applicants simply do not have the time for troubleshooting. Even in cases where a filing is not urgent, the time burden of troubleshooting quickly exceeds the $400 non-DOCX surcharge, and troubleshooting is rarely cost-effective. In cases where applicants cut their losses and file in PDF format, neither the public nor the USPTO gains any benefit from the time spent. Time wasted in failed attempts to file DOCX should be counted toward the burden.

Our staff members who have become fluent in the DOCX filing process estimate that the extra burden of preparing, uploading and verifying a DOCX file adds approximately 0.5 hours per application, as compared to preparing, uploading, and verifying a PDF file. Unlike the USPTO, we do not expect that this added burden will decrease significantly with additional experience.

The USPTO also ignores the added burden of proofreading the record copy. Due to the black-box nature of the USPTO's validation and rendering processes, applicants cannot assume that the USPTO's rendering of a DOCX file will accurately reflect either the source document or the DOCX file that they originally uploaded. A brief inspection of the USPTO-generated PDF file is insufficient for detecting rendering errors because substantive alterations can occur at the level of individual characters in a line of text.[35] Post-filing quality control therefore entails a close proofreading of the USPTO-generated PDF file, on top of the pre-upload proofreading and preparation, in order to detect such errors. At a typical rate of 10 pages per hour, close proofreading of a 30-page application adds a time burden of approximately **3 hours**.[36]

---

[32] 88 F.R. 37039, 37041 (June 6, 2023). Table 2 estimates a burden of 0.5 hours per application at an average rate of $435/hour.

[33] Comment by Kilpatrick Townsend & Stockton LLP, July 30, 2023, posted at https://www.regulations.gov/docket/PTO-P-2023-0031/comments (comment ID PTO-P-2023-0031-0003) ("KTS Comment").

[34] The USPTO boasts that the DOCX filing process "[e]liminates the need for patent applicants to convert structured text to PDF format." USPTO Statement, Response to Comment 1, p. 7. In our experience, converting "structured text" (which in our case means a Microsoft DOCX file) to PDF format is a trivial exercise, requiring no more effort than printing to paper. Converting our Microsoft DOCX files to DOCX files that will upload to Patent Center, the USPTO's electronic filing system, requires considerably more effort. KTS Comment, p.3.

[35] *See, e.g.,* AIPLA Comment at p. 33 (reporting an exponent that changed from 0.2u to 10.2u).

[36] Allowing applicants to file an auxiliary PDF that can be relied on to correct USPTO errors at any point during the life cycle of the patent may reduce the need for close proofreading of every application.

The USPTO has failed to address our estimate of burden in any meaningful way. Its response to "[c]ommenters estimated that they needed between 3 to 6 additional hours to review a single application filed in DOCX"[37] is only to state that: "The USPTO considered the most recent feedback received and maintains that the estimated time burden of 30 minutes is an adequate amount of time and does not warrant a further increase."[38] The USPTO fails to explain why close proofreading of the record copy is unnecessary or to offer any other justification for ignoring the informed estimates of burden provided by people who actually file patent applications.

This burden is not cost-justified. At the average hourly rate of $435 applied by the USPTO,[39] the 3.5-hour burden on applicants for filing and proofreading would exceed $1,000 per application. Even at a reduced proofreading rate of $150/hour, the burden on applicants would typically exceed the proposed $400 non-DOCX surcharge. When these costs are combined with the risk of USPTO-introduced alterations going undetected until it is too late to correct, many applicants may opt to upload a PDF and pay the non-DOCX surcharge, thwarting the USPTO's stated goal of receiving more applications in structured-text format.

The USPTO has not shown any offsetting savings that would justify this substantial added burden on applicants. The USPTO previously published an estimate that its cost for extracting electronic text from an application submitted in PDF format is approximately $3.15.[40] The USPTO now asserts that "this amount does not include other costs incurred in processing these documents" and that "[t]he use of image-based PDFs incurs many costs over the lifetime of an application."[41] The USPTO fails to provide any estimate or explanation of these alleged "other costs." Unless these "other costs" are a hundred times the cost of initial extraction of electronic text, the proposed $400 non-DOCX surcharge cannot be justified by any cost savings to the USPTO.

Practitioners' opinion of the DOCX filing process is revealed in the data. The USPTO has stated that "183,685 total applications [have been] submitted in DOCX since the USPTO began accepting new applications in the DOCX format."[42] The USPTO does not specify when it began counting these 183,685 applications. We estimate that over 2,000,000 utility applications have been filed since the non-DOCX surcharge was proposed in 2019.[43] Therefore, over this

---

However, as discussed above, the USPTO refuses to permanently establish the option of filing an auxiliary PDF.

[37] USPTO Statement, Comment 13, p. 13.

[38] USPTO Statement, Response to Comment 13, p. 14.

[39] "Setting and Adjusting Patent Fees During Fiscal Year 2020," 85 F.R. 46932, 46947 (August 3, 2020).

[40] *Id.*

[41] USPTO Statement, Response to Comment 12, p. 13.

[42] *Id*.

[43] The non-DOCX surcharge was proposed in "Setting and Adjusting Patent Fees During Fiscal Year 2020," 84 F.R. 37398, 37412-14 (July 31, 2019); although the USPTO began accepting applications in DOCX format two years earlier. *Id.* at 37143. Our estimate of "over 2,000,000" utility applications filed since July 2019 is based on USPTO data indicating that over 500,000 utility applications have been filed annually since 2012. "Table of Annual U.S. Patent Activity Since 1790," available at

four-year period, the DOCX filing rate is, at most, 10%. Few filers are choosing to file in DOCX. Whether that is because of the added risk or the added burden or both is unknown.

OMB should reject the proposed Information Collection for imposing a disproportionate burden on patent applicants.

Conclusion

The proposed Information Collection imposes both unjustified burdens and unnecessary risks on patent applicants. We urge OMB to reject the proposed Information Collection and advise the USPTO to develop an alternative that would better balance its desire to receive structured text with applicants' need for a record that accurately reflects what applicants actually filed.

For questions regarding this comment, please contact
Cathy Cretsinger
ccretsinger@kilpatricktownsend.com
415-576-0200

https://www.uspto.gov/web/offices/ac/ido/oeip/taf/h_counts.htm. This data set ends in 2020; however, other reports indicate that annual patent filings have not decreased in recent years.