David Boundy, DavidBoundyEsq@gmail.com
Carl Oppedahl, carl@oppedahl.com
Neil R. Ormos, ormos-lists@ormos.org
Brian Siritzky, Brian@siritzkylaw.com
Richard A. Schafer, richard@schafer-ip.com

February 14, 2023

Dear Mr. Seidel:

This letter is a second follow-up to our meeting of February 1. In reviewing our notes and recollections, we realized that one idea was briefly mentioned but did not get sufficient attention because of the meeting time constraints. We think that idea is important and want to address it here.

Near the end of our meeting, you asked whether we thought the filing community would be comfortable if the USPTO were to require the filer to provide, in addition to an authoritative PDF file, the DOCX file that had been used to create that PDF file. You further asked whether we thought the filing community would be comfortable if the USPTO were to further require the filer to certify that the PDF file had been produced from the DOCX file.

While we briefly explained why these proposed requirements would be unacceptable, time constraints prevented our further elucidation.

In our short agenda for our February 2 meeting, we set out two principles that we think are fundamental: (1) The file that the applicant uploads must be the authoritative document for all purposes. (2) What You See Is What I Get. Anyone who opens or prints a filed document, anywhere, at any time, will see the same rendering in content and visual layout. We have no objection to the USPTO providing additional *options* (such as DOCX filing). But it is crucial that at least one option be available that strictly observes the two principles. We strenuously object to any *requirement* that would intrude upon either of the two principles.

The DOCX-as-auxiliary proposal (and as far as we can tell, ***any*** proposal that requires a submission of a DOCX file in ***any*** role) violates these two principles, as we will elaborate below:

- Reading between the lines, it appears that the USPTO has a basic misunderstanding of word processor technology and the ECMA and ISO standards. In this letter, we hope to put those issues to rest. DOCX is an inherently fragile, unreliable format. The ECMA-376 and ISO 29500 standards do not give the relevant and required guarantees. A DOCX document does not become reliable because it is given a name like “auxiliary” instead of “authoritative.”
- There are several practical problems: (a) for some applicants, no DOCX file exists, (b) where a DOCX file exists, it may not be available for filing, and (c) it may be impossible to make any certification

DOCX is Inherently Fragile, No More Reliable as an Auxiliary than as a Primary

We surmise you would not have asked about a “DOCX as auxiliary” approach unless USPTO staff harbored some lingering hope that it is feasible to work around the fundamental unreliability of DOCX. The technology mismatch is irreconcilable, if quality and correctness matter.

First, DOCX ***cannot*** be a reliable vehicle for the purposes the USPTO proposes, even in an auxiliary role. DOCX was designed for an incompatible purpose—from the outset, DOCX was *designed* to adapt the presentation of a document to exploit different resources that exist on different computers. Carl gave several examples in his letter of August 2019. We gave several examples in our Seventy-Three Practitioners letter at pages 13-19, and in the contrast between pages 1-34 vs. pages 48-80. Dozens of articles explain that various implementations of DOCX are not compatible with each other at the level of precision required for legal documents—some written by Microsoft employees (e.g., here, here, here, and here).

Patent attorneys around the world recognize the unreliability of DOCX as an interchange medium. Even where all parties are using Microsoft Word, differences among editions of Word, different hardware, fonts, and software installed on different computers, and the like, result in changes to the appearance of the document. Even if a document is written in English, when it is written in an edition for a non-Latin-alphabet market (*e.g.*, Israel, Japan, China, and Korea) and read in the U.S. edition, characters and paragraphs change, appear, or disappear. Recently, Dr. Siritzky received instructions from a foreign attorney to file a patent application. The foreign attorney sent both a DOCX (in case editing would be required) and a PDF. The cover letter included this sentence:

> **Note 1**. Please kindly take our PDF version into your consideration because of different version of software may result in unpredictable format errors (ex. formula, figures and so on).

In drafting *this very letter*, we had a problem when an entire line dropped out because one tool disagrees with the others—errors arise in different tools' handling of simple text documents.

Second, in the August 2020 Final Rule notice, the USPTO relied heavily on two standards, ECMA-376 and ISO 29500. In our view, the Final Rule's explanation raises grave concerns for USPTO staff's understanding of the role of standards. "Standard" in no way implies "interoperable" or "portable." We obtained all four of the relevant standards (the two DOCX standards and the two PDF standards)—did the USPTO? Even a brief reading of ECMA-376 and ISO 29500 will convince any reader that the two standards are the *problem*; they do not propose a solution. Each of the two DOCX standards leaves many dozens of attributes as "implementation-defined." Both allow an implementation (such as Microsoft Word) to add undocumented extensions. The ECMA-376 and ISO 29500 standards never purported to provide the interoperability or portability necessary for this use.

We block-quoted several problem sentences from these two standards in our Seventy-Three Practitioners letter at pages 14-17. We were deeply troubled that the USPTO chose not to respond to these comments in the Final Rule notice, despite a legal obligation to do so.

In contrast, the two relevant PDF standards, ISO 32000 and 14289, allow ***no*** "implementation-defined" differences in the meaning of a file. The only "implementation-defined" parameters are maximum sizes for certain storage compartments. "P" in "PDF" means "portable."

Third, as Carl explained at length in his *The Fool's errand that is DOCX* (see also an article by the Free Software Foundation), as of 2023, no commercial product implements an interoperable form of ECMA-376 and ISO 29500. First, as of the adoption of these standards in 2008, no company had a conforming implementation. Second, later in 2008 (only months after the ISO issued standard 29500), the standard forked into a "strict" fork (which would have had the

stability and interoperability the USPTO seeks) and a "transitional" fork in which every vendor is free to add new features at will, with no accountability for interoperability. No vendor has ever implemented the "strict" fork. Microsoft chose the "transitional" fork. Third, in almost every release of Word since, Microsoft has added new features that are not "standard," some of which are not back-compatible. Today's Word has diverged incompatibly from the 2008 standards. Fourth, the USPTO will be faced with regular "adds" of undocumented new features to Word that will break the USPTO's intake software, which will be an ongoing cost and headache for the USPTO. More importantly, if the USPTO's computers refuse a filing close to a deadline because of a new internal feature added to Word, the result may be catastrophic for that application and attorney.

The proposal for a DOCX as an auxiliary to the "authoritative" PDF is inherently untenable. The DOCX standard does not promise, and our experience confirms, that text cannot be reliably and identically extracted across any two implementations.

We believe the following statements are not just true; they are indisputable and beyond judgment call. If any member of your team has the *slightest* reservation about any of them, please contact us promptly, so that we can reach agreement on basic facts. We want to save you the effort of chasing illusions that cannot possibly lead to a reliable end result:

a. Opening the same DOCX file on two different computers sometimes yields different results. Those differences are of frequency and severity that cannot be tolerated in a system for legal documents.

b. Even if the DOCX file is offered in an auxiliary role, that still creates proofreading, error correction, and malpractice costs that most practitioners will not accept (and that, under the Paperwork Reduction Act, the USPTO has no legal authority to impose).

c. The two standards cited by the USPTO, ECMA-376 and ISO 29500, do ***not*** guarantee portability in the necessary respects. During their engineering careers, Dr. Siritzky and Mr. Boundy were compiler writers, implementing the standards for FORTRAN 77, FORTRAN 90, ANSI C (1989), and IEEE floating-point numbers. Mr. Boundy was Hewlett-Packard's alternate representative to an ANSI standards committee for multiprocessor extensions to FORTRAN and was one of the key people that coached HP's representative to the ANSI C committee. Those standards do not guarantee portability, and the DOCX standards are similar in this respect. If your team has ***any*** remaining questions that the DOCX standards do ***not*** guarantee the specific properties necessary for patent application filing, we have deep experience to assist.

Our principle 2 is "What You See Is What I Get. The [file uploaded by the applicant] must guarantee that anyone, anywhere, any time, that opens or prints [it] will see the same rendering in content and visual layout." DOCX ***cannot*** meet that requirement. We have confidence, supported by some experimentation and the conclusion of the USPTO's own 2015 "yearlong study," that PDF files—if the USPTO specifies document creation switches—will do a *better* job of delivering ***reliable*** texts that can be extracted, analyzed for feedback and internationalized, than DOCX can. Several technically viable solutions are available to handle specially formatted objects, including those mentioned in the USPTO's study.

We assume that reliability, reproducibility, and unambiguity of the record are as essential to the USPTO as they are to us, and that the USPTO recognizes that any "auxiliary" document must be as reliable as the primary "authoritative" document. In the short agenda we sent you for our

February 1 meeting, the first entry was "Correctness first" before any optimization for efficiency or cost. If the USPTO's quality and reliability targets are different than ours, please explain.

For Many Applications, No DOCX File Ever Existed

According to the USPTO's own survey, for nearly 20% of applications, 20% of applicants use non-DOCX software to author patent specifications. For practitioners who do not use Microsoft Word, a requirement to supply "the DOCX file used to create the PDF" may be impossible because no such DOCX document ***ever existed***. Libre Office, for example, stores its document in a native data format of ODF. WordPerfect uses WPD. Google Docs has a native data storage format (called KIX) that is not accessible to the user. While each has an "export" function, those exported DOCXs were not used to generate the PDF. Further, companies do not have a commercial incentive to make *outbound* exports reliable, so the "export as DOCX" functions in non-Microsoft word processors tend to be even more unreliable than the Word-to-Word issues we discussed above (an example of mangling by "import" and "export" is in our Seventy-Three Practitioners letter in the contrast between pages 1-34 vs. pages 48-80.)

A wide variety of software is used to author patent applications. That diversity of software can be seen even in the five of us:

- One uses LibreOffice and produces files in ODT format.
- One uses a version of Microsoft Word that produces files in Microsoft's older ".DOC" format.
- One uses Word for Mac, which regularly differs from Word for Windows.
- One uses a combination of LibreOffice, plain text editors, and other tools. There may be no single word-processing file that contains the whole text.
- Only one uses Microsoft Word for Windows that produces files in Microsoft's current DOCX format.

The practitioner community is even more diverse. We know of people using Word Perfect, Google Docs, LaTeX, TeX, and troff. None of these natively create files in DOCX format. Some *cannot* create DOCX files.

Finally, if the USPTO wishes to require an "auxiliary" document, then the Paperwork Reduction Act is very clear on this: an agency must track the public's existing recordkeeping practices "to the maximum extent practicable." The USPTO may not discriminate against the 20% of filers that do not use a DOCX-native word processor. Likewise, if the USPTO wants an auxiliary, we assume the USPTO would want the most reliable form. For these two reasons, if the USPTO requires an auxiliary text document, the USPTO will have to implement intake extractors for at least Libre Office ODT, Word Perfect WPD, and Word 2003 DOC, and perhaps others.

Even Where a DOCX File Exists, It May Not Be Available for Filing

Clients and foreign associates often give us instructions and applications to file at the last minute. Often a client or foreign associate provides the filing practitioner with a PDF file with instructions to file it. As discussed above, there may be no DOCX file in existence because whoever produced the application (e.g., the client or foreign associate) did not produce the PDF from a DOCX file. Even if a DOCX file exists that was used to create the PDF file, that DOCX file may be unavailable to the filing practitioner.

Certifying the DOCX File is Often Impossible

The seemingly-innocuous requirement that the filer hand in "the DOCX file that the filer used when the filer created the PDF file" is not technologically feasible, given the variability among the patent application filing community. For a significant fraction of U.S. filers, it is a simple fact that *no DOCX file exists* that was used when the filer created the PDF file. Even where the filer may have a DOCX file, the filer may not know or be able to discover the relationship between that DOCX file and the PDF file. Even if the applicant provides a document with a DOCX extension, that file may not be a DOCX file or satisfy the USPTO's (or Microsoft's) then-current requirements for DOCX.

Even where a DOCX file is provided alongside the PDF, the filing practitioner cannot be sure that the DOCX file was used to create the PDF file. Even the client or foreign associate who provides the files to the filing practitioner may not be sure of that relationship. Thus, practitioners often will not be able to satisfy a requirement to submit a DOCX file from which the also-submitted PDF file was produced because they will lack the knowledge required to sign a submission to the Office. Likewise, filing practitioners often would not be able to certify any particular relationship between a PDF file and the DOCX file from which it was produced if such certification were required.

A requirement that the USPTO knows could not be satisfied by all filers would be arbitrary and capricious.

Any Practical Requirement for a DOCX Auxiliary Document Will Almost Certainly Violate our Principles 1 and 2

We believe that any attempt to require a DOCX file as an "auxiliary" file will recreate the problems of today's DOCX filing. Because of the designed-in fragility of DOCX files, it seems inevitable that the USPTO will be backed into some variation on the approach used today—accept the uploaded DOCX files, alter them for the USPTO's purposes, and present them to the applicant for visual approval. But the visual approval necessarily occurs *on the user's computer*—which provides *no assurance whatsoever* that the *USPTO's computers* will see the same thing. We see no way for the USPTO to accept DOCX files for *any* purpose that does not lead back around to violation of at least one of Principles 1 or 2.

Recordkeeping Costs of Pairing Two Files

The DOCX-as-auxiliary proposal will require the practitioner to pair two files for the life of the patent, out to 26 years. In today's regime, this recordkeeping is easy and straightforward: the usual filing workflow creates one and only one PDF file (any earlier ones are all discarded), so it's easy to identify the one and only one file that should be permanently archived. The DOCX approach requires the practitioner to single out a DOCX file from among the multiple work-in-progress versions that may exist and keep it paired to the PDF for at least twenty-six years. There are no off-the-shelf tools to do this easily. This is impractical and an invitation to a malpractice suit. Under the Paperwork Reduction Act, this requirement to alter recordkeeping practices, the requirement for duplicative filing, and the costs entirely for the convenience of the agency with no benefit to the filer are all unlawful.

Legal hurdles

The Paperwork Reduction Act required the USPTO to "consult with the public" early on so that all the points we raise in this letter would have been well understood before the USPTO committed resources to this project. Despite inquiries, we have been unable to find any record or any consultation with the public before the PPAC presentation in 2018. Many of these issues were raised in notice-and-comment letters. The USPTO evaded answering by mischaracterizing the comment, not answering at all, or in some instances, simply lying ("To date, the Office has not received notifications of any issues resulting from the filing of applications in DOCX format."). Likewise, the USPTO was required to make the "survey" and "yearlong study" available for public vetting at the time of the Notice of Proposed Rulemaking. Neither was disclosed. The USPTO was required to estimate costs on the public, do several cost-benefit analyses, and make them available for public comment; the USPTO has never done so. The USPTO was required to make several filings at OMB; OMB's website shows no such filings ever occurred, and the USPTO's last filing at OMB admitted as much. The USPTO's representation of the findings of the yearlong survey in the August 2020 Federal Register final rule is 180º opposite the "Conclusion" of the study. Each of these incongruities was unlawful.

The Paperwork Reduction Act requires the USPTO to "implement [its information collection practices] in ways consistent and compatible, to the maximum extent practicable, with the existing reporting and recordkeeping practices of those who are to respond." 44 U.S.C. § 3506(c)(3)(E). Requiring 20% of applicants to change word processors would be unlawful.

The USPTO may not require "unnecessarily duplicative" filing. § 3506(c)(3)(B). If the PDF has the information in it (as the USPTO's own "yearlong study" concludes), the USPTO may not require a duplicative DOCX.

An agency should reduce its own costs, but "shall not do so by means of shifting disproportionate costs or burdens onto the public." 5 C.F.R. § 1320.5(d). The USPTO was required to do a cost estimate, and show that the costs on the public are not "disproportionate" to its $3.15 savings. The USPTO never did the cost estimation, let alone the cost-benefit analysis. (The public did—two estimates came in at about $200 million per year.) Costs of duplicate filing and error checking will necessarily be many times the USPTO's cost savings.

It must be deeply frustrating to the USPTO that we're asking you to do additional work in order to accommodate the requirements of a broader cross-section of filers than the USPTO had originally considered. We suggest that these costs are the inevitable consequence of the USPTO's neglect of the law. The Administrative Procedure Act, Paperwork Reduction Act, Regulatory Flexibility Act, Information Quality Act, and Executive Order 12866 required the USPTO to research certain topics, confer with the public, develop certain cost-benefit analyses, publish those cost-benefit analyses in the Notice of Proposed Rulemaking for public vetting, fairly answer public comments (as opposed to unfairly misparaphrasing some comments and entirely skipping others), make filings with OMB, and certify that it has taken certain steps to reduce burden on the public. The USPTO skipped all these steps. These are not just legal requirements; they are sound engineering. Any engineering costs that fall on the USPTO are, we suggest, a consequence of shortcutting sound engineering and legal processes.

Conclusion

As we clarified in our meeting, we accept that some filers have adopted DOCX filing. Nobody speaks for all patent practitioners, and we have no desire for the USPTO to tell those people they cannot file in that format, even though we think doing so is misguided. Nor do we have any objection to the USPTO being willing to accept filings that provide both a DOCX file and PDF file *at the filer's option*, even though we would never want to do so.

As to the two questions you asked at the end of our meeting, for at least the reasons set forth above, we and many others strongly object to any *requirement* that the filer must provide a DOCX file as an auxiliary (regardless of whether the USPTO imposes any financial penalty), particularly if the requirement includes any expectation or requirement that the applicant specifies or certifies any specific relationship among the files.

While this letter repeats many of our issues with DOCX, albeit now within the context of the USPTO's proposed "auxiliary" document with a certification requirement, this letter should also be read in the context of other proposals or endorsement of DOCX proposals the USPTO may be receiving from others.

As we have laid out above, it is evident to us that support for ***any*** proposal that requires a DOCX requires willful blindness to the technological facts, the experience and opinions of the patent procurement community, the content of the ECMA and ISO standard documents, and the governing law. While we don't purport to speak for all filers, we do have the benefit of hundreds of emails exchanged on Carl's email lists—the membership includes many hundreds of attorneys, agents, and paralegals, from small firms, large firms, and in-house, and a few foreign participants. We think our two principles, and the implementation we suggested on February 1, represent a broad consensus of many parts of the filing community.

While the USPTO may have received support for some variant of a DOCX proposal from professional associations such as the AIPLA and the IPO, we note that those associations do not represent even a majority of stakeholders, let alone the full spectrum necessary to meet the USPTO's obligations of fairness and uniform treatment under several statutes. The USPTO should not rely on "approval" from any of these associations to represent approval or agreement from the legally-necessary spectrum of stakeholders.

Time constraints in our meeting prohibited a complete discussion of those ideas. We hope this letter helps you understand why the DOCX-as-auxiliary proposal would be a serious mistake with severe adverse consequences.

We look forward to the follow-up call you offered in our meeting and will be happy to discuss these issues in more detail in that call. If issues of technology or standards arise in the interim, we can discuss them immediately. If you would like suggestions from us about better serving all applicants' needs while still getting the USPTO what it needs, we would be glad to work with you in a two-way discussion.

You can reach me, David Boundy, at 646 472 9737.