

Doris Mendiola - Additional questions and comments: Guidance Document

From: David Austin <david.a.austin@snet.net>
To: "Kenny D. Nguyen" <kdn@nrc.gov>, "Thomas E. Smith" <tes@nrc.gov>
Date: 07/17/2007 10:08 PM
Subject: Additional questions and comments: Guidance Document

Good day Tom and Kenny -
I have attached a PDF file with additional questions/observations.

6/28/07
72 FR 35521

A quick summary:

1. "data transfer rates"
2. ISO 9660 vs Joliet extension to ISO 9660
3. CD-ROM file systems
4. PDF Image Only / Searchable Image (Exact)
5. Section 508 and PDF Image Only / Searchable Image (Exact)
6. Section 508 but is it usable by the vision impaired?

4

RECEIVED

2007 JUL 18 AM 10:32

RULES AND DIRECTIVES
BRANCH
USNRC

Be well...
Dave Austin

Adobe Certified Expert - FrameMaker

*SONSI Review Complete
Template = ADM-013*

*E-RIDS = ADM-03
Add = T. Smith (TES)*

Mail Envelope Properties (469D75F8.681 : 2 : 34433)

Subject: Additional questions and comments: Guidance Document
Creation Date 07/17/2007 10:07:43 PM
From: David Austin <david.a.austin@snet.net>

Created By: david.a.austin@snet.net

Recipients

nrc.gov

TWGWPO02.HQGWDO01
 TES (Thomas Smith)

nrc.gov

TWGWPO01.HQGWDO01
 KDN (Kenny Nguyen)

Post Office

TWGWPO02.HQGWDO01
 TWGWPO01.HQGWDO01

Route

nrc.gov
 nrc.gov

Files

MESSAGE
 TEXT.htm
 Questions15_20.pdf
 Mime.822

Size

453
 598
 96207
 135848

Date & Time

07/17/2007 10:07:43 PM

Options

Expiration Date: None
Priority: Standard
ReplyRequested: No
Return Notification: None

Concealed Subject: No
Security: Standard

Questions

Q15.

Could the Guidance Document be updated such that the first sentence of the first paragraph of Section 4.2.1, OSM Content is less ambiguous?

The "best fit" for an OSM that is "closed" would be to call for a CD/DVD burn session to use Disc-At-Once (DAO), Closed. Such a disc may be used as a master for additional duplication with production level equipment. Use of such a process precludes the error precursor inherent in use of TAO.

The alternative, Track-At-Once (TAO), Closed would provide a "read-only" disc. However, TAO does not provide the added value of being a "master" for production duplication.

Q16.

Could the ambiguity of the first sentence in the first paragraph of Section 4.2.2, Packaging/Labeling be addressed?

"The OSM should be labeled with the Transfer Media Configuration (e.g., drive transfer rate)... "

The lack of specificity as to what the NRC actually means by "Transfer Media Configuration" is rather frustrating. What is being sought?

- file system used?
- CD/DVD Mode used?
- ??

Also, calling for the drive transfer rate is, well, egregious. Consider, "my" drive transfer rate is, in fact, utterly meaningless to "you". Data transfer rate during a "read" by a CD or DVD drive is solely dependent upon the drive being used. No "submitter" has any practical way to know what the drive transfer rate is on any given NRC desktop or laptop that has a CD/DVD reader drive. The NRC could receive a CD-ROM that could have data burned at some transfer rate ranging from 1X (150 KB/sec) to 52X. Any contemporary reader drive will dynamically adjust to a "drive/data" transfer rate to accommodate the media.

An example of why "drive transfer rate" is specific to the equipment used.

This is from the "burn" perspective.

I recently produced a "master" CD-ROM (DAO, Closed) which contained PDF files.

The CD-ROM used for the "master" was labeled as 1X-16X.

The CD burner transferred data to the disc at 32X.

Using the "master" disc I then used a TEAC CDW-D17 to duplicate the master to five CD discs.

These five discs were also 1X-16X discs. The TEAC "juke-box" can burn up to 52X.

The data transfer rate to the five discs was, again 32X.

In sum, "drive transfer rate" is valid only to the physical equipment used.

What the NRC has/uses will not be, in all cases, what the submitter has.

Unless the NRC is willing to purchase and maintain common hardware/software platforms

for all submitters calling for "data transfer rates" is neither practicable nor meaningful.

As CD/DVD media ranges widely in price submitters will, of course, look for cost effective purchases.

With the ambiguity present in the current wording of the Guidance Document some might perceive the NRC as setting the stage to require specific types of CD/DVD media/burners be purchased.

Example - Blue ray CD/DVD burners provide more "precision" when burning/reading data.

BUT - The margin permitted for the laser and the read heads' voltage supplied and alignment is next to zero. The slightest deviation in either voltage or alignment will result in your computer being unable to "read" a perfectly good CD/DVD.

Conversely, a "yellow" dye in the CD/DVD (TDK approach) lacks the "precision" but is far more fault tolerant with regards to voltage and alignment. Thus, media produced on my computer will be readable on your computer.

Regarding "Modes"

We have -

For audio, no "mode" ... burn "raw" ...audio is "fault tolerant", so no error correction.

Mode 1 - for data, computer applications - has error correction

Mode 2 (CD-ROM XA)

Form 1 - for data, has error correction ...typically for compatibility for older drives.

Form 2 - for video (audio-video) ...vcd or svcd ...no error correction ...not for data

For submittals that are "data only" (i.e., UFSARs, the preponderance of content in COLAs, etc.)

Mode 1 is, specifically, the desired "mode" to use.

Mode 2 XA, Form 2 might be what the NRC should "cite" for vcd or svcd submittals.

Will the NRC develop and promulgate amplifying guidance in this regard?

Doing such would optimize keeping all involved on the "same page".

Q17.

When will the current Guidance Document be revised to remove the ambiguity and potentially misleading information regarding file format?

Specifically:

Appendix A to the Final "e" Rule called for the use of the Joliet extension to ISO 9660.

The current Guidance Document does not call for this.

Section 2.4, Naming convention states, for file naming -

"...should conform to ISO 9600 format."

However, the examples in the table "File Naming Example Table" are not ISO 9660 compliant.

About ISO 9660:

- ISO 9660 Specifications

- - Restrictions

- - - File and directory name restrictions. There are also different levels to this standard.

Level 1: File names are restricted to eight characters with a three-character extension, upper case letters, numbers and underscore; maximum depth of directories is eight.

[COMMENT: clearly, this is not what the examples depict]

Level 2: File names are not limited to 8.3 format, but may be up to the maximum allowed by the 1 byte counter in the dir entry and the filename length byte counter.

Typically this is close to 180 characters depending on how many extended attributes are present.

Level 3: Files allowed to be fragmented (mainly to allow packet writing, or incremental CD recording).

-- Other name restrictions:

--- All levels restrict names to upper case letters, digits, underscores ("_") and a dot.

[COMMENT: "upper case" letters only... The examples in the Guidance Document use mix case.]

--- Linux converts uppercase letters to lower case while mounting ISO file systems.

--- File name cannot start or end with dot.

--- File name cannot have more than one dot.

[COMMENT: examples use more than one dot]

--- Directory names cannot use dots at all.

[COMMENT: this may adversely impact what the NRC is calling for re: COLAs]

--- Directory depth limit

--- The restrictions on filename length and directory depth (to 8 levels including the main directory)

[COMMENT: The Guidance Document does not reflect this restriction.]

--- Many CD authoring applications attempt to work around this by truncating filenames automatically,

--- but at the risk of breaking applications that rely on a specific file structure.

[COMMENT: the NRC will not/cannot accomplish what it has promulgated for COLA submittals if the 8 level directory depth limit called out by use of ISO 9660. The truncating of filenames, by windows centric CD/DVD mastering applications is the norm, not the exception. Thus, directory names will be "smudged".]

Q18.

Will the NRC update the Guidance Document to reflect specific and correct information regarding the expected CD/DVD format?

Suggestion:

(a) use, in Windows, the Joliet extension to ISO 9660. (Linux, Apple, others would have to be given similar, explicit guidance.)

(b) For text and graphics (i.e., engineering drawings) use Mode 1.

(c) For all OSM, use Disc-At-Once (DAO), Closed sessions. This absolutely assures a "read-only" OSM.

Q19.

What specific guidance will the NRC provide submitters re: Section 508?

The issue is less "is the PDF 508 compliant" and more "if 508 compliant, is it *usable*". Once the NRC accepts a submittal that is "508" compliant it would seem that the NRC thereby incurs a legal obligation/liability vis-a-vis "usability" to those who must have 508 compliant files available. The crux of Section 508 is "usability". Without significantly enhanced guidance or "how-to" information, the NRC is likely to find itself in turbulent waters over the "usability" issue. Should this occur, what, if any, liability is placed on the organizations that are providing electronic submittals? If there is any, perhaps they would want to return to hard copy submittals, no?

Examples of why far more "usability" guidance is needed.

(1) Many submitters provide, as part of the submission, engineering drawings. Often such drawings are identified as Figure 1.2-1, Figure 1.2-2, etc. A text to speech function does not render/speak the "1.2-1" in a manner that provides vision impaired individuals a clear/valid understanding of what is being rendered.

No vision disability: For "1.2-1" we have "one dot two dash one". The visual layout/configuration dictates how the characters are perceived/recognized. The visual pattern defines what the data order and meaning is. These individuals do not use a text to speech function to achieve understanding and use.

With vision disability: For "1.2-1" the text to speech function renders the words "one point two one"; which, when heard connotes "1.21". This, of course, is incorrect. However, the PDF file is "Section 508 compliant". Just not "usable" by the vision impaired individual.

Solution:

When the AltText entry is written "1.2 dash 1" the text to speech rendering provides a clear, usable, and understandable piece of information ("one dot two dash one" which gives us "1.2-1", which is correct, understandable, usable). Also, entry of the colon and space characters(": ") will make the text to speech function pause. This makes a significant difference in usability.

(2) Elevations in a figure title - The AltText string for the image/figure tag in the PDF would, typically, reflect what is in the title block. Of course, the EL. 54'-6" represents the major premise that the users is not vision impaired. Entering the elevation information in the AltText field gives us a PDF that is 508 compliant. However, again, the file is not usable by the vision impaired. Why? the string is spoken "fifty four single hyphen six" (54'6 for the non-vision impaired). Incorrect, confusing, not usable or understandable.

Solution:

The AltText entry must be "54 feet 6 inches". This results in a clear, unambiguous text to speech rendering.

(3) Abbreviations. Use of abbreviations is pandemic. By both the submitters and the NRC.

Text to speech applications "don't do" abbreviations that are industry specific.

e.g., BLDG, GENL, ARRG, EL. TURB, etc.

Or "CM" for Configuration Management... ("cm" is read as "centimeter").

Use of abbreviations, even "standard" (to "us") abbreviations, cannot be used in an AltText entry.

As a best practice they should not be used in the text flow of digital textual content.

Not if the "508 PDF" is to meet usability standards.

Remember, we the non-vision impaired, do not identify or establish such standards.

Those who champion the needs of the vision impaired do so, sometimes by way of legislation, but most often via litigation in the courtrooms.

Q20.

The Guidance Document contains an incorrect statement that requires correction.

re: Section 2.3, File Formats

In the table "Preferred PDF Output File Formats"

Row Three, Column Four:

"Preferred format for graphic-, image-, and forms-oriented documents (PDF Image Only cannot be used for textual documents because it is not Section 508 compliant. See Section 2.21)"

FIRST:

The preferred (IRS, business engaged in Publishing Industry activity, etc.)

PDF output format for "forms-oriented documents" is **not** PDF Image Only.

SECOND:

Any and All PDF Image Only output files can be and are processed to be compliant to Section 508.

The scanned image that has been imported into PDF to produce the PDF Image Only output file is post-processed by Adobe Acrobat Professional (it is preferable to use the most current release with updates).

You designate the language, "tag" the image, "fence" the image, designate the fenced image as a "Figure" and add the required AltText... all done using Adobe Acrobat Professional.

In point of fact Adobe Acrobat Professional is the **only** application that can adequately post-process PDF files for compliance with Section 508.

Remember, the majority of "tools and plug-ins" are for web publishing and rely on the protocols established by W3C. PDFs are **not** in W3C space. They are the child of Adobe System Inc.'s and any Section 508 compliance for PDFs comes from Adobe and what they put into Acrobat Professional for the "developer" and into Acrobat Standard and Adobe Reader for the "user".

THIRD:

One cause for "rejection" is non-compliance with Section 508.

One of the three acceptable PDF output types is PDF Image Only.

Yet, the table entry states PDF Image Only "cannot" be 508 compliant.

Ergo, PDF Image Only must not be "acceptable" for submission.

Fact: PDF Image Only, like the other two acceptable PDF output types

can be made 508 compliant.

SOF: *No* PDF file, of any output type is, currently, created such that it is Section 508 compliant "out of the gate". Yes, the PDF files can be "tagged" in the authoring environment (currently only Adobe FrameMaker, Adobe InDesign and MS Word provide this capability - yes, other MS Office applications provide some, very limited tagging... but, they do not merit consideration as an "authoring environment").

However, "tagged" does not equate to Section 508 compliance.

"Tagging" is step one. Then there is "the rest of the story" where full compliance is established via use of Adobe Acrobat Professional.

Observation:

It is my understanding that NRC staff reviewers prefer that PDF Image Only files (e.g., engineering drawings) be OCR'd to create Searchable Image (Exact) PDFs. This is to support PDF Find/Search.

The OCR'd characters, when rendered to speech by a text to speech reader are meaningless. That engineering drawings scanned into PDF format are "graphic-, image- ..." files is unequivocally clear. To produce Section 508 compliant (meaning "usable") files will require a submitter to *not* OCR the PDF Image Only file. Producing the Section 508 compliant (and usable) PDF file is, while time consuming, very straight forward.

Note: "Tagging" the OCR'd hidden/invisible characters/fonts of a Searchable Image (Exact) to establish Section 508 compliance (read order of usable/understandable text flow) is *not* do-able in any practicable manner.

The core "data" of the Searchable Image (Exact) PDF is the image - not the OCR'd characters. To produce a Section 508 compliant PDF the Searchable Image (Exact) PDF is processed in the manner I described above. No different than a PDF Image Only file. That the text has been OCR'd means two things:

(1) The OCR characters are a rendering of characters, that if the source hard copy was "clean", provides a very close *approximation* of the text captured in the image.

The less "clean" the source hardcopy, the greater the deviation between OCR and actual.

Note that any emphasis on the OCR text provided by Searchable Image (Exact) can easily be construed as emphasis on that which is, in no venue, the "legal" data (see the NRC disclaimer associated with ADAMS files).

The "exact" in Searchable Image (Exact) is the image.

It is *not* the OCR content.

It occurs to me that any requirement that all PDF files containing a scanned image of text be OCR'd or risk rejection might engender a problematic legal situation. Something to have the legal eagles look at... it'd be interesting.

(2) The OCR'd text is processed by the text to speech software. The spoken OCR character content is often not understandable or usable. After all, the OCR characters are *not* the real textual content. Often, such "Section 508 compliant" PDF(s) do not meet any practicable usability criteria. As these PDF files are now the NRC's, the NRC becomes exposed to a "liability" issue.

Certainly PDF Formatted Text and Graphics (formally known as PDF Normal) can be post processed to be Section 508 compliant.

However, many submitters have legacy/historical textual content that was never transcribed into a word processor application of a publishing authoring environment (FrameMaker, InDesign, Quark, etc.).

To provide PDF files for electronic submittals the hardcopy was scanned. A PDF Image Only file is the result. These can be and are used as legal documents. PDF Searchable Image (Exact) provides a *convenience* - that of the use of find/search. Again, a convenience, not an essential. The image is the legal copy. The OCR characters, in the USA, are not. The PDF Searchable Image (Exact) is, in fact, at its core, just another PDF Image Only file.

What is made Section 508 compliant is the image. Not the OCR'd hidden/invisible text.

SubQuestion:

Does the NRC expect the legacy/historical textual information to be provided as "Section 508" compliant?

If yes, the NRC may have to provide a legal basis that would compel submitters to incur the significant expense of re-mastering such content into a digital authoring environment.

The task requires re-typing and independent checks. The information, being part of a submittal, is "legal stuff". Thus, care and diligence is required. All this requires significant expenditures of monies and people resources.

Suggestion - re: establishing meaningful guidance for Section 508 compliance "how-to" - Get active involvement of the actual "508" stack holders. American Federation of the Blind, etc.