

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 13, 2013

P. Hoffman
VPN Consortium
July 12, 2012

Proposal for Canonical and Other Formats for RFCs
draft-hoffman-rfcformat-canon-others-03

Abstract

This document proposes a new, XML-based canonical format for RFCs that explicitly allows external art as a normative part of the RFC. If the RFC Editor chooses this format, they will also publish non-canonical versions of RFCs in order to accommodate the largest target audience of readers. Having a simple, stable canonical format and a varying number of non-canonical formats that can change over time allows the RFC Editor to add useful formats, particularly in HTML, that can keep up with the needs of all RFC readers.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 13, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must

include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

1. Introduction

A clear result of the decades-long discussion about the format of published RFCs is that different RFC readers have different needs and desires. No single format will be sufficient, or even useful, to all people who read RFCs. Another clear result is that the format described in [RFC2223] and its follow-ons is no longer the best format for publishing protocols, process descriptions, research findings, and the many other types of documents that are part of the modern RFC series.

This document proposes to deal with these issues in a way that meets the needs and desires of the widespread RFC-reading community. For every RFC, the RFC Editor will publish both a canonical version of the RFC that is in XML format and multiple additional forms of the RFC, most notably at least in one or more HTML formats. The XML format used will likely be an updated version of that from [RFC2629], most notably to include in-line graphic art.

It is noted that XML files are not easily readable. However, it is also noted that the canonical version of an RFC doesn't need to be easily readable: only the non-canonical formats derived from the canonical version need to be readable.

Today, all RFCs are easily retrievable by all readers. In the future, all of the versions of an RFC and its art must be easily accessible as well. To make this easier, the RFC Editor will establish a permanent URL template for each RFC that leads to a page that lists all of the versions and art; a copy of that URL will be included near the beginning of the RFC in the boilerplate so that new RFC readers can find it. Further, it will be easy for advanced RFC users to mirror the entire collection of RFC material.

A major motivation behind the "one canonical, many non-canonical" proposal is to allow the RFC Editor to easily change the non-canonical formats in the future without having to change the canonical format. For example, the recent discussion of RFC formats has shown that many people strongly desire good HTML versions of RFCs, but there is not agreement of exactly what format the HTML should take. Further, it is completely clear that the HTML standard will evolve in the coming years and decades, and some of the new features that will be added will be quite useful in RFCs. Allowing the RFC Editor to add additional HTML formats to the RFC collection,

even for RFCs that have been published in the past, gives the greatest value to RFC readers without forcing any changes on the canonical RFC format.

Similarly, it is clear that HTML is not the only useful format for RFCs. Some people really like plain text; others want PDFs or other printer-ready paginated formats; still others want different formats that can be converted to different reading devices. Some people want detailed metadata for RFCs so that they can better mine them for relevant information; such metadata can be contained in either XML or HTML formats. All of these people can be accommodated by the RFC Editor publishing multiple non-canonical versions of RFCs. The canonical version of the RFC and all the non-canonical versions of the RFC should have predictable URLs so that tools can easily find (for example) an RFC in the reader's preferred HTML style just by knowing the RFC number.

The method that the RFC Editor uses to create the non-canonical formats for RFCs is left up to the RFC Editor. For example, they might generate it directly from the input files, through an intermediate format, or something else.

2. Canonical RFC Format and Content

Canonical RFCs are in XML format. The most salient rules for the format and content of those files are:

- o The format for the XML will be specified by the RFC Editor. It is likely that the XML format will be an improvement to that which is now referred to as "xml2rfc" ([RFC2629] and its informal successors).
- o The XML format will allow for art to be contained in the file. This art might be instead of text art in a document (such as for a diagram that is too complex to render well in text), or might be better variants for text art. The RFC Editor will determine which graphic formats are allowed, and it is likely that at least one vector format and one pixel-based format will be permitted.
- o The XML format will contain all of the metadata needed to produce any of the non-canonical formats for an RFC.
- o The text encoding for the document is UTF-8.
- o The RFC Editor can decide where it is and is not appropriate to use non-ASCII characters from the Unicode repertoire in the RFC. For example, the RFC Editor might make rules about using non-ASCII

characters in people's names, reference titles, examples in text, and so on.

- o Text art that internal to the document is limited to 95 columns. This is reasonable for printing on laser printers from the past 25 years, and allows much more expressive art than the current maximum of approximately 70 columns.

Text art encompasses many types of content. The unifying feature is that it is one or more lines of text that must be rendered with a monospace font in order to be fully understood in the context of the document. Thus, text art includes graphical representations such as packet diagrams, flow diagrams, and flow charts, but it also includes other text that needs to have column alignment such as multi-line ABNF.

This proposal does not deal with how mathematical equations might be included in the canonical RFC format. An author can do it as text or as art in an external file. The RFC Editor might allow an equation-specific format from external art files.

3. Additional Formats Provided by the RFC Editor

The RFC Editor will derive and publish non-canonical documents in multiple formats from the RFC. If the RFC-reading community agrees on a single HTML format, that will certainly be published. If the RFC-reading community cannot agree on just one HTML format, the RFC Editor might publish non-canonical versions of an RFC in multiple HTML formats. The RFC Editor will oversee the development of the tools needed to produce the non-canonical formats.

Depending on interest from the RFC-reading community, the RFC Editor will also publish non-canonical versions in other formats. For example, it is likely that the RFC Editor will publish in at least one format of PDF. Because some tools in widespread use rely on the current RFC format, the RFC Editor might also publish a non-canonical version in using the rules in RFC 2333 (line lengths, page headers, and so on).

4. Input to the RFC Series

The RFC Editor will allow submission of RFCs in the same XML format as the canonical version of an RFC. This allows an author to use differencing tools to track all changes that are made to the document that they submitted to the RFC Editor.

The RFC Editor will also possibly allow additional formats for submission based on agreement with the RFC streams. If other submission formats are allowed, the RFC Editor will convert the submission to the canonical format before performing any editing so that all editorial changes are easily tracked within the canonical format. This is similar to what they do currently with submissions that are not in the format the RFC Editor uses for its internal tooling.

This proposal in this document does not affect the allowed format for the publication of Internet-Drafts. The IETF Chair has indicated that such a change might happen after the choices are made for RFC format.

5. Metadata Needed to Create RFCs

The canonical format for RFCs must contain all of the body text for the RFC as well as all of the metadata that is used to mark up the RFC. RFC metadata is useful for many things such as finding RFCs with particular types of content and for making it clearer to a reader what the RFC author intended.

The following is a list of the metadata that needs to be part of the canonical RFC format. This list will probably be controversial, but the eventual list needs to contain all of the metadata that is intended for the final RFC format so that the format can be fully specified. Items marked with an asterisk are especially likely to need much more work.

Status

- Category
- RFC stream
- Obsoletes
- Updates
- Date published
- Draft derived from
- Title and short title

Per-author

- Name
- Initials
- Company
- Postal
- Email
- URL
- Role (editor, other)

Front content

- Abstract

- Copyright
- Other legal *
- Headings
 - Level
 - Number
- Paragraphs
 - Indented for quoting or emphasis
- Lists
 - Style (bulleted, numbered, unmarked)
 - Nesting
 - Elements
 - Definitions
- Validatable formats
 - ABNF
 - XML
 - JSON
 - ASN.1
- Inline art
 - Datagram layout
 - State diagram
 - Flow chart
 - Pseudocode
 - Table *
 - Non-validatable fragments of validatable formats
 - Other
- Tables *
- Special sections
 - Security Considerations
 - IANA Considerations
 - Normative References
 - Informative References
- Cross-references
 - To internal sections
 - To reference
 - To section in reference
- References *
 - RFCs
 - Specific versions of IDs
 - Non-specific versions of IDs
 - Common non-IETF documents (IEEE, ITU, ISO, ANSI, Unicode)
 - Corner cases

6. IANA Considerations

None

7. Security Considerations

None

8. Informative References

[RFC2223] Postel, J. and J. Reynolds, "Instructions to RFC Authors", RFC 2223, October 1997.

[RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, June 1999.

Author's Address

Paul Hoffman
VPN Consortium

Email: paul.hoffman@vpnc.org

Network Working Group
Internet-Draft
Updates: 2223 (if approved)
Expires: September 13, 2012

R. Braden
USC/ISI
J. Klensin
March 12, 2012

Images in RFCs
draft-rfc-image-files-03

Abstract

Documents in the RFC series normally use only plain-text ASCII characters and a fixed-width font. However, there is sometimes a need to supplement the ASCII text with images -- e.g., graphics, equations, or pictures. The historic solution to this requirement has allowed secondary PDF and/or Postscript files, but this approach has seldom been used because it is awkward for authors and publisher. This memo suggests a convenient scheme for logically including authoritative diagrams, illustrations, equations, or other graphics within RFCs.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 13, 2012.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect

to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. A New Scheme for Images: Composite RFCs	4
3. Construction of an Image File	5
4. Requirements for the Base File	6
4.1. Overview	6
4.2. Figures Section	6
4.3. Formatting Changes	7
5. Submission and Processing of the Image File	8
6. Bundled Composites	8
7. Implementation Considerations	9
8. RFC Repository File Formats	10
9. Internationalization Considerations	11
10. Approval and Authorization	11
11. Security Considerations	11
12. IANA Considerations	12
13. Acknowledgments	12
14. References	12
14.1. Normative References	12
14.2. Informative References	13
Appendix A. Change Log	13
A.1. Version -03	13
Authors' Addresses	13

1. Introduction

Published documents in the RFC series normally use only plain-text ASCII characters and a fixed-width font [RFC2223]. This simple convention has the advantage of a stable encoding for which a great variety of tools are readily available for viewing, searching, editing, etc.

Inclusion of diagrams, state machines, complex equations, and other graphics in RFC text has generally relied on the imaginative use of ASCII characters ("ASCII artwork"). However, in a few cases over the years, ASCII artwork has been inadequate for images needed or desired in RFCs. The solution to this dilemma has been to allow multiple versions of an RFC: a primary ASCII version as well as secondary versions that are encoded using PDF and Postscript. The PDF and Postscript versions are "complete", containing a copy of the text as well as the full images [RFC2223]. The textual content and layout of the PDF/PS version is required to match the base version as closely as possible. However, except in a few rare exception cases (see, e.g., [RFC1129]) the ASCII text version is considered the official expression of the RFC, and it is always normative for standards-track documents. We will refer to this scheme as "txt/ps/pdf" representation.

The three versions of an RFC using the txt/ps/pdf representation are in separate files in the primary RFC repository (<http://www.rfc-editor.org/rfc/>), with suffixes ".txt", ".pdf", and ".ps". The RFC Editor search engine returns links to all three versions when they are present in the repository.

Unfortunately, the txt/ps/pdf approach has been awkward for both editor and author, and it is error-prone. Therefore, it has seldom been used (roughly 50 out of 5000+ RFCs). The problem is that, in general, only the author has the tools to edit the PDF and Postscript versions. The RFC Editor can readily edit only the primary ASCII text, and then the author must incorporate the resulting changes into the PDF/PS version while maintaining the "look" of the RFC to the extent possible. There is no practical way for the RFC Editor to verify that this is done correctly, which may lead to editorial errors, reader confusion. It may also lengthen publication time.

This memo suggests a much better scheme for including figures, illustrations, and graphics to an RFC. We hope that the method proposed here will solve the image problem for RFC publication, while preserving the convenience, stability, and searchability of ASCII base documents. The txt/ps/pdf approach would still be possible (and in any case, RFCs using the historic scheme will continue to exist in the RFC repository forever).

Discussion of this document is being moved to the rfc-interest mailing list. For subscription information and archives, see <http://www.rfc-editor.org/mailman/listinfo/rfc-interest>.

2. A New Scheme for Images: Composite RFCs

Under the suggested scheme, an RFC would be either a single ASCII file as generally used today, or a composite of two files: an ASCII-only "base file" containing the text of the RFC, and an "image file". When present, the image file would be a PDF file that contained only images, captions, and title information. Neither file of the composite would be complete without the other, and a reference to the RFC would be considered a reference to both files. An RFC would then be a logical entity whose complete representation could require two files, base and image.

The base file would be formatted exactly like current ASCII RFCs, with minor exceptions described below. An image file would contain one or more items that will be known collectively as "figures", whether they are actually diagrams, pictures, tables, equations, artwork, or other non-textual constructions.

If the approach of this document is adopted, terms like "the RFC" will refer to the combination of the base RFC file and an image file if the latter is supplied. Note that, just as with the txt/ps/pdf representations, an RFC is a logical entity whose complete representation requires multiple files. In particular, the IPR statement in the base file ("Rights Contributors Provide to the IETF Trust in Contributions BCP 78, RFC 5378 [RFC5378]) would apply to the composite, including the image file if one is present.

An ASCII RFC traditionally uses a file name in the form of "rfcN.txt", where N is integer RFC number without leading zeros. The image file that is associated with RFC number N could be named "rfcN.img.pdf". As noted earlier, the repository contains RFCs with file names "rfcN.ps" and "rfcN.pdf", using the historic txt/ps/pdf representations.

This "image file" scheme was inspired by a tradition of book publishing, in which pictures, figures, or "plates" may be grouped together following the text ("end figures"), or even bound separately from the main body of the text.

In principle, we could allow an image file to be encoded using both PDF and Postscript, since mechanical translation is possible in both directions. However, in the 20 years since the adoption of the txt/ps/pdf representations, the PDF format has become a de facto standard

for electronic documents, and readers for it are universally available. Furthermore, PDF has been standardized as a format for document archiving, as discussed further in the next section. Therefore, we propose to allow only PDF for image files, simplifying the new approach by not including a Postscript file option.

3. Construction of an Image File

An image file would be a single PDF file, consistent with the description in [RFC3778] and defined in [ISO32000-1]. The particular PDF form must be version-stable and must not contain any external references in scripts or otherwise. Those requirements are satisfied by the PDF/A [ISO19005-1] [ISO19005-2] profile. The RFC Editor may authorize other variants of PDF in the future.

There is an issue of whether particular PDF generators PDF that claim to satisfy PDF/A actually do so. Future experience may require published guidelines on PDF-generating software that claims to satisfy PDF/A but does not.

Except as otherwise specified in this document, an image file should contain only figures, supporting labels and captions, headers, and footers. It should not contain explanatory text or other materials that could reasonably be expressed in plain-text form in the base file. In particular, required sections of RFCs, such as IANA Considerations or references, must be completely understandable from the base text file. Any figures referenced from those sections may contain only supplemental material.

The image file would be paginated using the same 8.5 x 11" format as the base document, and the image file pages would be consecutively numbered. The first page number of the image file would follow the last page number of the base RFC. Each page of the image file would contain the same headers and footers as the base file, except for one change in the footer, suggested below. Since editing the base file may change its pagination, it may be simplest to ask the RFC Editor to overlay the headers and footers onto the image file near the completion of editing. Each page would need blank space at the top and bottom for this purpose. The amount of blank space is to be determined, but 0.5" might be a reasonable value.

Figures included in the image file would have to be labeled in a fashion that facilitated referencing from the base RFC. The labels should normally be numeric and monotonic. Simple consecutive integers will usually be the best choice, but in some cases it might be desirable to use a hierarchical scheme like: <section #>.<fig #>. An author who believes that another labeling scheme would increase

clarity should consult the RFC Editor.

4. Requirements for the Base File

4.1. Overview

A base file would be unchanged by the presence of an image file, except for the following.

- o The page number of the end-of-RFC boilerplate page would be changed to be logically one page after the last image file page.
- o A new unnumbered "Figures" section would be required. This is described below.
- o For a composite RFC, a minor modification to the first-page header of the base file and to the footers of both base and image files would tie the two files together. This is described below.

4.2. Figures Section

An RFC that used this scheme (and had any figures) would need to include a Figures section in the ASCII base file. The Figures section should immediately follow the Table of Contents, if any, and precede the body of the document. The Figures section should list all figures in tabular form, indicating for each one the figure identification, title, and page number(s).

The style for the Figures section has not yet been fully specified. Here is a suggested example.

Table of Contents

1. Introduction	1
2. Philosophy	7
2.1 Elements of the Internetwork System	7
2.2 Model of Operation	8
2.3 The Host Environment	8
(etc)	

Figures

Figure 1: Protocol Layering	2
Figure 2: Protocol Relationships	9
Figure 3: TCP Header Format	15, *86
Figure 4: Send Sequence Space	20
Figure 5: Receive Sequence Space	20
Figure 6: TCP Connection State Diagram	23, *87
Figure 7: Basic 3-Way Handshake for Connection Synchronization	31, *88
(etc)	

*Page in Image file

(Page 1 follows)

An RFC that includes a base file may include ASCII artwork that is suggestive of a figure in the image file, but there is no requirement to do so. When such an approximate figure appears as ASCII artwork in the base file, its figure identification and caption must match those of the corresponding figure in the image file, and the entry in the Figures table should specify the page numbers in both the base and image file. In the example shown above, image file page numbers are marked with an asterisk. Note that very simple ASCII artwork need not have corresponding material in the image file.

Groups of mathematical equations form one particular case in which it may be desirable for a base file to include ASCII artwork approximations. This will ease searching for such documents using equation terminology. equations. There are well-established conventions for approximating fairly complex equations using ASCII artwork.

4.3. Formatting Changes

It would be necessary to tie the base and image files together, to make clear they are part of one RFC. Here is an initial suggestion for formatting.

The header line "Request for Comments: nnnn" in the base file could be changed to "Request for Comments: nnnn/Base". For consistency, the lefthand footer could become "RFC nnnn/Base". The lefthand footer in the image file could then be: "RFC nnnn/Image".

The following sentence could be placed in the "Status of this Memo" section: "This RFC is a composite of this base file and PDF image file <image file URL>."

5. Submission and Processing of the Image File

If an image file is needed, it should be submitted as an .img.pdf file along with the ASCII text file.

The image file could be submitted without headers or footers. The RFC Editor could then overlay the image file with the appropriate headers and footers, with correct pagination. The RFC Editor would do no editing of the image file beyond adding headers and paginated footers. If editing the base file revealed problems with figures in the image file, the authors would be asked to create a new image file.

6. Bundled Composites

The base/image file split should be very convenient for the process of editing and publishing RFCs, and all tools that return RFC meta-data should alert the reader to the composite structure. However, users may sometimes prefer to obtain an existing composite RFC as a single file in a bundled format.

Our suggestion for such bundling is to again use PDF encoding. Thus, corresponding to the composite file pair (rfcN.txt, rfcN.img.pdf), there would be a new file with a name like "rfcN.bun.pdf". The .bun.pdf file would be a single PDF file containing a facsimile of the .txt file (like the current .txt.pdf files) followed by the image file.

It is important to understand what is being suggested here. The .bun.pdf files would never be submitted for publication by authors; instead, the RFC Editor would mechanically generate the .bun.pdf files upon publication of the .txt and .img.pdf files (just as .txt.pdf files are now generated automatically). Some users might choose to consider the bundled .bun.pdf file as "the RFC", but the RFC Editor would consider "the RFC" to be the {.txt, .img.pdf} file pair. We note that:

The composite is logically a single RFC that can be normative for a standards-track document.

The .bun.pdf bundle, which would be mechanically derived from the composite pair, might reasonably be declared in the future to be equally normative.

The continued existence of a base file that is readily editable by the RFC Editor and readily searchable by users would maintain the advantages of the present ASCII-based scheme. We are not suggesting that, at least in the near term, the composite structure with its ASCII text component be abandoned.

On the other hand, the .bun.pdf bundle could be a transitional step towards a future world where RFCs are published in pure .pdf.

7. Implementation Considerations

Implementation of the image file scheme has a number of implications.

1. The Internet Draft repository must allow submission and retrieval of both base and (when present) image files.
2. Internet Draft file names could be draft-...-vv.txt and (optionally) draft-...-vv.img.pdf, where "vv" is the normal version number. Updating either file of the composite RFC should increase the version numbers "vv" in both files. We DO NOT want two separate version numbers for one I-D.
3. The RFC Editor would need to be able to overlay headers, footers, and page numbers on a given image file. It is claimed that at least Adobe Acrobat Professional includes this capability, and that it also has limited editing capability.
4. The RFC Editor would also need a tool to verify that a given image file satisfies the constraints of PDF/A and that the original image can be overlaid with headers and footers.
5. Some RFC Editor scripts and tools would need extensions.
6. Small extensions to xml2rfc [RFC2629] would be useful to create base/image file cross-references in header and footers, and to build a Figure section.

8. RFC Repository File Formats

A frequent reaction to the suggestion given in this memo is some confusion over the different file formats that appear in the RFC repository. Here is a brief summary.

If a PDF image file exists along with a base ASCII RFC, then RFCs in any format other than that combination (e.g., complete PDF files, HTML, or Postscript) remain supplemental, with the reader taking responsibility for assuring that they are equivalent to the base RFC and image file. That arrangement is identical to the relationship between traditional all-ASCII RFCs and supplemental forms: the RFC Editor has never taken responsibility for guaranteeing that the two are identical in content.

The existing .txt.pdf files are not affected by this proposal, nor are any of the traditional non-ASCII formats. The .txt.pdf files are facsimiles of .txt (base files) in PDF, introduced to help Windows users read RFCs online. However, Microsoft has more recently provided an elementary ASCII editor, which probably makes the .txt.pdf files unnecessary in any case.

In summary:

- o rfcN.txt: ASCII-only file. In the traditional system, complete normative file. In the new system, text (base) part of normative composite RFC, or stand-alone normative text file when no image file is necessary.
- o rfcN.ps: Traditional system -- a Postscript file that includes figures and whose text is intended to be the same as the normative .txt file, but is generally non-normative itself. No new rfcN.ps files would be created after adoption of the image file proposal.
- o rfcN.pdf: Traditional system -- a PDF file that includes figures and whose text is intended to be the same as the normative .txt file, but is generally non-nformative itself.
- o rfcN.txt.pdf: Traditional system: Facsimile of corresponding .txt file.
- o rfcN.img.pdf: New system: image file component of a composite RFC.
- o rfcN.bun.pdf: New system: bundled composite file.

9. Internationalization Considerations

Our scheme of image files does not, and is not intended to, support character set internationalization for RFCs. It does not allow an author to omit the ASCII text from the base file and instead include the entire RFC text as one (very large) image file.

However, we should note two examples that illustrate special cases.

1. RFC 3743 [RFC3743] on internationalized domain names for Chinese, Japanese, and Korean contains a number of examples that may be hard to follow because they can represent those characters only in "U+nnnn" form. An image file could be used that would show the alternative Chinese characters for the examples. This would not diminish either the ability to search the base text or index the document or its readability for those of us for whom reading Chinese characters is difficult, but it should help those who can read them.
2. Suppose that a proposed RFC contained a section derived from Japanese text. The author might put an English translation into that section of the base document, note that the original was really in Japanese, and attach the Japanese as an appendix in an image file. This should raise no difficulties for informative documents. For normative documents, however, the existence of the Japanese original would raise some issues about what was actually authoritative, which is very undesirable.

[Note in Draft: A separate proposal [Hoffman-RFC-UTF8] is under consideration to permit UTF-8 strings to appear directly in text RFCs under restricted circumstances.]

10. Approval and Authorization

Placeholder: In its capacity as the body that approves the general policy followed by the RFC Editor [RFC2850], [RFC5620]), the RSOC or full IAB would eventually need to review and sign off on this proposal after its tentative approval by the RFC Series Editor (RSE).

11. Security Considerations

This specification addresses documentation standards and adding additional flexibility to them. It does not, in general, raise any security issues. However, unless the specifications of this document are carefully followed, the image format recommended, PDF, may potentially contain external references or scripts that could

introduce security problems. That problem could be largely or completely alleviated, and long-term stability improved, by exclusive use of the PDF/A format as discussed in Section 3. The RFC Editor and other publishers should exercise due care to ensure that no such references or scripts appear in the archives.

12. IANA Considerations

This document requires no actions by the IANA. An intentional consequence of the model is that IANA will not need to inspect the image file in order to carry out its task of evaluating proposed RFCs for potential actions.

13. Acknowledgments

The impetus for this specification arose during a discussion during an RFC Editorial Board meeting after one of the IETF's many discussions about "modern" formats for RFCs. Aaron Falk made several specific suggestions that have been reflected in the document. The RFC Editor staff and other Editorial Board members contributed suggestions without which this version would not have been possible, as did Steve Bellovin, Alfred Hoenes, Olaf Kolkman, and Henrik Levkowetz.

14. References

14.1. Normative References

[ISO19005-1]

International Organization for Standardization (ISO),
"Document management -- Electronic document file format
for long-term preservation -- Part 1: Use of PDF 1.4
(PDF/A-1)", ISO 19005-1:2005, 2005.

[ISO19005-2]

International Organization for Standardization (ISO),
"Document management -- Electronic document file format
for long-term preservation -- Part 2: Use of ISO 32000-1
(PDF/A-2)", ISO 19005-2:2011, 2011.

[ISO32000-1]

International Organization for Standardization (ISO),
"Document management -- Portable document format -- Part
1: PDF 1.7", ISO 32000-1:2008, 2008.

- [RFC2223] Postel, J. and J. Reynolds, "Instructions to RFC Authors", RFC 2223, October 1997.
- [RFC3778] Taft, E., Pravetz, J., Zilles, S., and L. Masinter, "The application/pdf Media Type", RFC 3778, May 2004.
- [RFC5378] Bradner, S. and J. Contreras, "Rights Contributors Provide to the IETF Trust", BCP 78, RFC 5378, November 2008.
- [RFC5620] Kolkman, O. and IAB, "RFC Editor Model (Version 1)", RFC 5620, August 2009.

14.2. Informative References

- [Hoffman-RFC-UTF8]
Hoffman, P. and T. Bray, "Using non-ASCII Characters in RFCs", November 2008, <<https://datatracker.ietf.org/drafts/draft-hoffman-utf8-rfcs/>>.
- [RFC1129] Mills, D., "Internet Time Synchronization: The Network Time Protocol", RFC 1129, October 1989.
- [RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, June 1999.
- [RFC2850] Internet Architecture Board and B. Carpenter, "Charter of the Internet Architecture Board (IAB)", BCP 39, RFC 2850, May 2000.
- [RFC3743] Konishi, K., Huang, K., Qian, H., and Y. Ko, "Joint Engineering Team (JET) Guidelines for Internationalized Domain Names (IDN) Registration and Administration for Chinese, Japanese, and Korean", RFC 3743, April 2004.

Appendix A. Change Log

A.1. Version -03

Conversations about the formats of RFCs are restarting at IETF 83 in March 2012. This version of this draft is a minor update of the November 2008 version -02 to enter it into that discussion. It reflects changes in ISO standards and the RFC Editor model since 2008 as well as a few minor editorial corrections.

Authors' Addresses

Robert Braden
USC/ISI
4676 Admiralty Way
Marina del Rey, CA 90292
USA

Phone: +1 310 448 9173
Email: braden@isi.edu

John C Klensin
1770 Massachusetts Ave, #322
Cambridge, MA 02140
USA

Phone: +1 617 491 5735
Email: john-ietf@jck.com

