Internet Architecture Board (IAB)
Request for Comments: 7990
Category: Informational
ISSN: 2070-1721

H. Flanagan
RFC Editor

December 2016

RFC Format Framework

Abstract

In order to improve the readability of RFCs while supporting their archivability, the canonical format of the RFC Series will be transitioning from plain-text ASCII to XML using the xml2rfc version 3 vocabulary; different publication formats will be rendered from that base document. With these changes comes an increase in complexity for authors, consumers, and the publisher of RFCs. This document serves as the framework that provides the problem statement, lays out a road map of the documents that capture the specific requirements, and describes the transition plan.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Architecture Board (IAB) and represents information that the IAB has deemed valuable to provide for permanent record. It represents the consensus of the Internet Architecture Board (IAB). Documents approved for publication by the IAB are not a candidate for any level of Internet Standard; see Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7990.

Table of Contents

1. Introduction

"RFC Series Format Requirements and Future Development" [RFC6949] discusses the need to improve the display of items such as author names and artwork in RFCs as well as the need to improve the ability of RFCs to be displayed properly on various devices. Based on the discussions with communities of interest, such as the IETF, the RFC Series Editor decided to explore a change to the format of the Series [XML-ANNOUNCE]. This document serves as the framework that describes the problems being solved and summarizes the documents created to-date that capture the specific requirements for each aspect of the change in format.

Key changes to the publication of RFCs are highlighted, and a transition plan that will take the Series from a plain text, ASCII-only format to the new formats is described on the rfc-interest mailing list [RFC-INTEREST].

This document is concerned with the production of RFCs, focusing on the published formats. It does not address any changes to the processes each stream uses to develop and review their submissions (specifically, how Internet-Drafts will be developed). While I-Ds have a similar set of issues and concerns, directly addressing those issues for I-Ds will be discussed within each document stream.

The details described in this document are expected to change based on experience gained in implementing the new publication toolsets. Revised documents will be published capturing those changes as the toolsets are completed. Other implementers must not expect those changes to remain backwards compatible with the details described in this document.

2. Problem Statement

There are nearly three billion people connected to the Internet [ISTATS] and individuals from at least 45 countries have regularly attended IETF meetings over the last five years. The Internet is now global, and while the world has changed from when the first RFCs were published, the Series remains critical to defining protocols, standards, best practices, and more for this global network that continues to grow. In order to make RFCs easily viewable to the largest number of people possible, across a wide array of devices, and to respect the diversity of authors and reference materials while still recognizing the archival aspects of the Series, it is time to update the tightly prescribed format of the RFC Series.

All changes to the format of the RFC Series must be made with consideration to the requirements of a wide set of communities over an extended length of time. Examples of the preferences and specific needs are those of existing authors and implementers, lawyers that argue Intellectual Property Rights (IPR), educators, managers, and policymakers that need to know what to list in potential Request for Proposals (RFPs) for their organizations. The immediate needs of today's communities must be balanced with the needs for long-term archival storage.

3. Terminology

This document uses terminology from RFC 6949, repeated below for convenience.

ASCII: Coded Character Set - 7-bit American Standard Code for Information Interchange, ANSI X3.4-1986 [ASCII]

Canonical format: the authorized, recognized, accepted, and archived version of the document

Metadata: information associated with a document so as to provide, for example, definitions of its structure, or of elements within the document such as its topic or author

Publication format: display and distribution format as it may be read or printed after the publication process has completed

Reflowable text: text that automatically wraps to the next line in a document as the user moves the margins of the text, either by resizing the window or changing the font size

Revisable format: the format that will provide the information for conversion into a Publication format; it is used or created by the RFC Editor

Submission format: the format submitted to the RFC Editor for editorial revision and publication

4. Overview of the Decision-Making Process

Requirements, use cases, concerns, and suggestions were collected from the communities of interest at every stage of the project to update the RFC format. Input was received through the rfc-interest mailing list, as well as in several face-to-face sessions at IETF meetings. Regular conversations were held with the Chairs of the IETF, IRTF, IAB, and IAOC as well as the Independent Stream Editor to discuss high-level stream requirements. Updates regarding the status

of the project were provided to the IETF community during the IETF Technical Plenary as well as Format BoFs or IAB sessions at several IETF meetings [IETF84] [IETF85] [IETF88] [IETF89] [IETF90].

The output from the first year of discussion on the topic of RFC format was published as RFC 6949, which provided the first solid documentation on the requirements for the Series. RFC 6949 is a product of the IAB stream (following the process described in "Process for Publication of IAB RFCs" [RFC4845]). This is also the case with all of the RFCs that informed the format update work.

After the high-level requirements were published, the RFC Series Editor (RSE) brought together an RFC Format Design Team to start working out the necessary details to develop the code needed to create new and changed formats. The Design Team discussed moving away from the existing xml2rfc vocabulary, but with such a strong existing support base within the community and no clear value with other XML vocabularies or schemas, the decision was made to work with the xml2rfc version 2 (xml2rfc v2) [RFC7749] model and use it as the base for the new format environment. Part of this discussion included a decision to stop using an XML document type definition (DTD) in favor of a Regular Language for XML Next General (RELAX NG) model using a defined vocabulary. While the biweekly calls for this team were limited to Design Team members, review of the decisions as documented in the documents produced by this team was done publicly through requests for feedback on the rfc-interest mailing list. Several of the documents produced by the Design Team, including those on xml2rfc v2 [RFC7749] and v3 [RFC7991] and the SVG profile [RFC7996], were sent through an early GenART review [GEN-ART] before starting the process to be accepted by the IAB stream.

While the IETF community provided the majority of input on the process, additional outreach opportunities were sought to gain input from an even broader audience. Informal discussions were held with participants at several International Association of Scientific, Technical, and Medical Publisher events [STM], and presentations made at technical conferences such as the TERENA Networking Conference 2014 [TNC2014] and NORDUnet 2014 [NDN2014].

In order to respond to concerns regarding responses to subpoenas and to understand the legal requirements, advice was requested from the IETF Trust legal team regarding what format or formats would be considered reasonable when responding to a subpoena request for an RFC.

Given that several other standards development organizations (SDOs) do not offer plain-text documents, and in fact may offer more than one format for their standards, informal input was sought from them

regarding their experience with supporting one or more non-plain-text formats for their standards.

Finally, the entire process was reviewed regularly with the RFC Series Oversight Committee [RSOC] and regular updates provided to the IAB and IESG. They have offered support and input throughout the process.

Where consensus was not reached during the process, the RSE made any necessary final decisions, as per the guidance in "RFC Editor Model (Version 2)" [RFC6635].

5. Key Changes

At the highest level, the changes being made to the RFC format involve breaking away from solely ASCII plain text and moving to a canonical format that includes all the information required for rendering a document into a wide variety of publication formats. The RFC Editor will become responsible for more than just the plain-text file and the PDF-from-text format created at time of publication; the RFC Editor will be creating several different formats in order to meet the diverse requirements of the community.

The final XML file produced by the RFC Editor will be considered the canonical format for RFCs; it is the lowest common denominator that holds all the information intended for an RFC. PDF/A-3 will be the publication format offered in response to subpoenas for RFCs published through this new process and will be developed with an eye towards long-term archival storage. HTML will be the focus of providing the most flexible set of features for an RFC, including JavaScript to provide pointers to errata and other metadata. Plain text will continue to be offered in order to support existing tool chains, where practicable, and the individuals who prefer to read RFCs in this format.

6. Canonical Format Documents

6.1. XML for RFCs

Key points regarding the XML format:

- The canonical format for RFCs is XML using the xml2rfc version 3 (xml2rfc v3) vocabulary. The XML file must contain all information necessary to render a variety of formats; any question about what was intended in the publication will be answered from this format.

- Authors may submit documents using the xml2rfc v2 vocabulary, but the final publication will be converted to use the xml2rfc v3 vocabulary.
- SVG is supported and will be embedded in the final XML file.
- There will be automatically generated identifiers for sections, paragraphs, figures, and tables in the final XML file.
- The XML file will not contain any xml2rfc v3 vocabulary elements or attributes that have been marked deprecated.
- A DTD will no longer be used. The grammar will be defined using RELAX NG [RNC].
- The final XML file will contain, verbatim, the appropriate boilerplate as applicable at time of publication specified by RFC 7841 [RFC7841] or its successors.
- The final XML will be self-contained with all the information known at publication time. For instance, all features that reference externally defined input will be expanded. This includes all uses of xinclude, src attributes (such as in <artwork> or <sourcecode> elements), include-like processing instructions, and externally defined entities.
- The final XML will not contain comments or processing instructions.
- The final XML will not contain src attributes for <artwork> or <sourcecode> elements.

[RFC7749] describes the xml2rfc v2 vocabulary. While in wide use at the time of writing, this vocabulary had not been formally documented prior to the publication of RFC 7749. In order to understand what needed to change in the vocabulary to allow for a more simple experience and additional features for authors, the current vocabulary needed to be fully described. RFC 7749 will be obsoleted by [RFC7991].

[RFC7991] describes the xml2rfc v3 vocabulary. The design goals were to make the vocabulary more intuitive for authors and to expand the features to support the changes being made in the publication process. It obsoletes RFC 7749.

7. Publication Format Documents

7.1. HTML

[RFC7992] describes the semantic HTML that will be produced by the RFC Editor from the xml2rfc v3 files.

Key points regarding the HTML output:

- The HTML will be rendered from the XML file; it will not be derived from the plain-text publication format.
- The body of the document will use a subset of HTML. The documents will include Cascading Style Sheets (CSS) for default visual presentation; it can be overwritten by a local CSS file.
- SVG is supported and will be included in the HTML file.
- Text will be reflowable.
- JavaScript will be supported on a limited basis. It will not be permitted to overwrite or change any text present in the rendered HTML. It may, on a limited basis, add text that provides post-publication metadata or pointers, if warranted. All such text will be clearly marked as additional.

7.2. PDF

[RFC7995] describes the tags and profiles that will be used to create the new PDF format, including both the internal structure and the visible layout of the file. A review of the different versions of PDF is offered, with a recommendation of what PDF standard should apply to RFCs.

Key points regarding the PDF output:

- The PDF file will be rendered from the XML file; it will not be derived from the plain-text publication format.
- The PDF publication format will conform to the PDF/A-3 standard and will embed the canonical XML source.
- The PDF will look more like the HTML publication format than the plain-text publication format.
- The PDF will include a rich set of tags and metadata within the document.

- SVG is supported and will be included in the PDF file.

7.3. Plain Text

[RFC7994] describes the details of the plain-text format; in particular, it focuses on what is changing from the existing plain-text output.

Key points regarding the plain-text output:

- The plain-text document will no longer be the canonical version of an RFC.
- The plain-text format will be UTF-8 encoded; non-ASCII characters will be allowed.
- A Byte Order Mark (BOM) will be added at the start of each file.
- Widow and orphan control [TYPOGRAPHY] for the plain-text publication format will not have priority for the developers creating the rendering code.
- Authors may choose to have pointers to line art in other publication formats in place of ASCII art in the .txt file.
- An unpaginated plain-text file will be created.
- Running headers and footers will not be used.

7.4. Potential Future Publication Formats

7.4.1. EPUB

This format is intended for use by ebook readers and will be available for RFCs after the requirements have been defined. No document on this topic is currently available.

8. Figures and Artwork

8.1. SVG

[RFC7996] describes the profile for SVG line art. SVG is an XML-based vocabulary for creating line drawings; SVG information will be embedded within the canonical XML at the time of publication.

9. Content and Page Layout

9.1. Non-ASCII Characters

There are security and readability implications to moving outside the ASCII range of characters. [RFC7997] focuses on exactly where and how non-ASCII characters may be used in an RFC, with an eye towards keeping the documents as secure and readable as possible, given the information that needs to be expressed.

9.2. Style Guide

The RFC Style Guide [RFC7322] was revised to remove as much page formatting information as possible, focusing instead on grammar, structure, and content of RFCs. Some of the changes recommended, however, informed the XML v3 vocabulary.

9.3. CSS Requirements

[RFC7993] describes how the CSS classes mentioned in "HyperText Markup Language Request for Comments Format" should be used to create an accessible and responsive design for the HTML format.

10. Transition Plan

10.1. Statement of Work and RFP for Tool Development

Existing tools for the creation of RFCs will need to be updated, and new tools created, to implement the updated format. As the requirements-gathering effort, described in the various documents described earlier in this document, finishes the bulk of the work, the Tools Development Team of the IETF will work with the RSE to develop Statements of Work (SoWs). Those SoWs will first be reviewed within the Tools Development Team and the Tools Management Committee, and it will then go out for a public comment period. After public review, the SoWs will be attached to an RFP and posted as per the IETF Administrative Support Activity (IASA) bid process [IASA-RFP].

Once bids have been received, reviewed, and awarded, coding will begin.

10.2. Testing and Transition

During the I-D review and approval process, authors and stream-approving bodies will select drafts to run through the proposed new publication process. The RFC Editor will process these documents after they have been approved for publication using xml2rfc v2 and will simultaneously test the selected I-Ds with the xml2rfc v3

process and tools. While the final RFCs published during this time will continue as plain text and immutable once published, the feedback process is necessary to bootstrap initial testing. These early tests will target finding issues with the proposed xml2rfc v3 vocabulary that result in poorly formed publication formats as well as issues that prevent proper review of submitted documents.

Feedback will result in regular iteration of the basic code and XML vocabulary. In order to limit the amount of time the RFC Production Center (RPC) spends on testing and quality assurance (QA), their priority will be to edit and publish documents; therefore, community assistance will be necessary to help move this stage along. A mailing list and experimental source directory on the RFC Editor website will be created for community members willing to assist in the detailed review of the XML and publication formats. Editorial checks of the publication formats by the community are out of scope; the focus will be the QA of each available output, checking for inconsistencies across formats.

The purpose of the testing phase is to work with the community to identify and fix bugs in the process and the code before producing canonical, immutable XML, and to collect additional feedback on the usability of the new publication formats.

Any modifications to the document review process, up to and including AUTH48, will happen with the community and the stream-approving bodies as we learn more about the features and outputs of the new publication tools. Defining those processes is out of scope for this document.

Success will be measured by the closure of all bugs identified by the RPC and the Tools Development Team as fatal in addition to reaching rough consensus with the community on the readiness of the XML vocabulary and final output files for publication. The actual rendering engine can go through further review and iteration, as the publication formats may be republished as needed.

Authors are not required to submit their approved drafts to the RFC Editor in an XML format, though they are strongly encouraged to do so; plain text will also remain an option for the foreseeable future. However, documents submitted as plain text cannot include such features as SVG artwork. The RPC will generate an XML file if necessary for basic processing and subsequent rendering into the approved output formats.

A known risk at this point of the transition is the difficulty in quantifying the resources required from the RPC. This phase will require more work on the part of the RPC to support both old and new

publication processes for at least six months. There is potential for confusion as consumers of RFCs find some documents published at this time with a full set of outputs, while older documents only have plain text. There may be a delay in publication as new bugs are found that must be fixed before the files can be converted into the canonical format and associated publication formats.

10.3. Completion

Authors may submit XML (preferred) or plain-text files. The XML files submitted for publication will be converted to canonical XML format and published with all available publication formats. All authors will be expected to review the final documents as consistent with the evolving procedures for reviewing documents.

Success for this phase will be measured by a solid understanding by the RSE and the IAOC of the necessary costs and resources required for long-term support of the new format model.

11. Security Considerations

Changing the format for RFCs involves modifying a great number of components to publication. Understanding those changes and the implications for the entire tool chain is critical so as to avoid unintended bugs that would allow unintended changes to text. Unintended changes to text could in turn corrupt a standard, practice, or critical piece of information about a protocol.

12. References

12.1. Normative References

[RFC6949] Flanagan, H. and N. Brownlee, "RFC Series Format Requirements and Future Development", RFC 6949, DOI 10.17487/RFC6949, May 2013, <http://www.rfc-editor.org/info/rfc6949>.

[RFC7749] Reschke, J., "The "xml2rfc" Version 2 Vocabulary", RFC 7749, DOI 10.17487/RFC7749, February 2016, <http://www.rfc-editor.org/info/rfc7749>.

[RFC7991] Hoffman, P., "The "xml2rfc" Version 3 Vocabulary", RFC 7991, DOI 10.17487/RFC7991, December 2016, <http://www.rfc-editor.org/info/rfc7991>.

[RFC7992] Hildebrand, J., Ed. and P. Hoffman, "HTML Format for RFCs", RFC 7992, DOI 10.17487/RFC7992, December 2016, <http://www.rfc-editor.org/info/rfc792>.

[RFC7993] Flanagan, H., "Cascading Style Sheets (CSS) Requirements for RFCs", RFC 7993, DOI 10.17487/RFC7993, December 2016, <http://www.rfc-editor.org/info/rfc7993>.

[RFC7994] Flanagan, H., "Requirements for Plain-Text RFCs", RFC 7994, DOI 10.17487/RFC7994, December 2016, <http://www.rfc-editor.org/info/rfc7994>.

[RFC7995] Hansen, T., Ed., Masinter, L., and M. Hardy, "PDF Format for RFCs", RFC 7995, DOI 10.17487/RFC7995, December 2016, <http://www.rfc-editor.org/info/rfc7995>.

[RFC7996] Brownlee, N., "SVG Drawings for RFCs: SVG 1.2 RFC", RFC 7996, DOI 10.17487/RFC7996, December 2016, <http://www.rfc-editor.org/info/rfc7996>.

[RFC7997] Flanagan, H., Ed., "The Use of Non-ASCII Characters in RFCs", RFC 7997, DOI 10.17487/RFC7997, December 2016, <http://www.rfc-editor.org/info/rfc7997>.

12.2. Informative References

[RFC4845] Daigle, L., Ed. and Internet Architecture Board, "Process for Publication of IAB RFCs", RFC 4845, DOI 10.17487/RFC4845, July 2007, <http://www.rfc-editor.org/info/rfc4845>.

[RFC6635] Kolkman, O., Ed., Halpern, J., Ed., and IAB, "RFC Editor Model (Version 2)", RFC 6635, DOI 10.17487/RFC6635, June 2012, <http://www.rfc-editor.org/info/rfc6635>.

[RFC7322] Flanagan, H. and S. Ginoza, "RFC Style Guide", RFC 7322, DOI 10.17487/RFC7322, September 2014, <http://www.rfc-editor.org/info/rfc7322>.

[RFC7841] Halpern, J., Ed., Daigle, L., Ed., and O. Kolkman, Ed., "RFC Streams, Headers, and Boilerplates", RFC 7841, DOI 10.17487/RFC7841, May 2016, <http://www.rfc-editor.org/info/rfc7841>.

[ASCII] American National Standards Institute, "Coded Character Set - 7-bit American Standard Code for Information Interchange", ANSI X3.4-1986, 1986.

[GEN-ART] IETF, "General Area Review Team (Gen-ART)", <http://www.ietf.org/iesg/directorate/gen-art.html>.

[IASA-RFP] IETF Administrative Support Activity, "RFPs and RFIs", <http://iaoc.ietf.org/rfps-rfis.html>.

[IETF84] Flanagan, H., "IETF 84 Proceedings: RFC Format (rfcform)", July 2012, <http://www.ietf.org/proceedings/84/rfcform.html>.

[IETF85] Flanagan, H., "IETF 85 Proceedings: RFC Format (rfcform)", November 2012, <http://www.ietf.org/proceedings/85/rfcform.html>.

[IETF88] Flanagan, H., "IETF 88 Proceedings: RFC Format (rfcform)", November 2013, <http://www.ietf.org/proceedings/88/rfcform.html>.

[IETF89] Flanagan, H., "IETF 89 Proceedings: RFC Format (rfcform)", March 2014, <http://www.ietf.org/proceedings/89/rfcform.html>.

[IETF90] Flanagan, H., "IETF 90 Proceedings: RFC Format (rfcform)", July 2014, <http://www.ietf.org/proceedings/90/rfcform.html>.

[ISTATS] "Internet Live Stats", <http://www.internetlivestats.com/internet-users/>.

[NDN2014] "28th NORDUnet Conference 2014", 2014, <https://events.nordu.net/display/NORDU2014/BoF%27s+and+side+meetings>.

[RFC-INTEREST] RFC Editor, "rfc-interest -- A list for discussion of the RFC series and RFC Editor functions.", <https://www.rfc-editor.org/mailman/listinfo/rfc-interest>.

[RNC] Clark, J., "RELAX NG Compact Syntax", OASIS , November 2002, <http://www.oasis-open.org/committees/relax-ng/compact-20021121.html>.

[RSOC] IAB, "RFC Editor Program: The RSOC", <http://www.iab.org/activities/programs/rfc-editor-program/>.

[TNC2014] Flanagan, H., "IETF Update - 'What's Hot?' - RFC Update", 2014, <https://tnc2014.terena.org/core/presentation/84>.

[STM] STM, "The global voice of scholarly publishing", <http://www.stm-assoc.org/>.

[TYPOGRAPHY] Butterick, M., "Butterick's Practical Typography", <http://practicaltypography.com/widow-and-orphan-control.html>.

[XML-ANNOUNCE] Flanagan, H., "Subject: [rfc-i] Direction of the RFC Format Development effort", message to the rfc-interest mailing list, May 2013, <http://www.rfc-editor.org/pipermail/rfc-interest/2013-May/005584.html>.

IAB Members at the Time of Approval

The IAB members at the time this memo was approved were (in alphabetical order):

Jari Arkko
Ralph Droms
Ted Hardie
Joe Hildebrand
Russ Housley
Lee Howard
Erik Nordmark
Robert Sparks
Andrew Sullivan
Dave Thaler
Martin Thomson
Brian Trammell
Suzanne Woolf

Acknowledgements

With many thanks to the RFC Format Design Team for their efforts in making this transition successful: Nevil Brownlee (ISE), Tony Hansen, Joe Hildebrand, Paul Hoffman, Ted Lemon, Julian Reschke, Adam Roach, Alice Russo, Robert Sparks (Tools Team liaison), and Dave Thaler.

Author's Address

Heather Flanagan
RFC Editor

Email: rse@rfc-editor.org
URI: http://orcid.org/0000-0002-2647-2220