Final Report on the

Request for Comments Editor Task

01 January 2002 thru 31 March 2007

INFORMATION SCIENCES INSTITUTE
UNIVERSITY OF SOUTHERN CALIFORNIA
LOS ANGELES, CALIFORNIA 90089-1147

ISI Technical Contact: Bob Braden
USC/Information Sciences Institute
4676 Admiralty Way
Marina del Rey, California 90292
(310) 822-1511, FAX: (310) 823-6714, braden@isi.edu

1. Introduction

The series of archival documents on computer communication known as Request for Comments (RFCs) was begun in 1969 as a computerized record of design notes for the ARPAnet. For the past 20 years, RFCs have been the vehicle for official publication of Internet standards and other protocol documents by the Internet Engineering Task Force (IETF). The series currently includes 4900 documents, totaling more than 145,000 pages. RFCs are published online, and in fact RFCs formed the first online publication series, since "online" (e.g., FTP) was invented by the authors of the earliest RFCs.

The person or organization responsible for editing and publishing RFCs is called the RFC Editor. Since 1978, the home of the RFC Editor has been the USC Information Sciences Institute (ISI), a leading Computer Science research center attached to USC. Originally the RFC Editor was one person, Jon Postel, who began editing the RFC series in 1970, moved to ISI in 1978, and continued as the RFC Editor until his death in October 1998. Postel was joined by Joyce Reynolds and later by a few other staff editors.

Postel's passing marked the end of an era for RFC publication; we may refer to the 9 years since then as the "modern" era for RFCs. ISI immediately reorganized its RFC Editor project and staffing to provide continuity of RFC publication. ISI also began a vigorous program to update the RFC Editor operations and services, under the leadership of Joyce Reynolds and Bob Braden.

From 1970 through 1988, the US government supported the RFC Editor function as the primary technical interchange vehicle for the ARPAnet and Internet research projects. In the late 1980s, the Internet research project began to evolve into today's global and commercial Internet, but the US government continued to support the RFC Editor until 1998. Beginning on July 1, 1998, the Internet Society (ISOC) took over RFC Editor support. Under a series of contracts and agreements, ISOC funding of the RFC Editor at ISI has continued until today.

The early contractual relationship between the RFC Editor and ISOC followed the model that was customary at ISI, designed for research contracts with the US government. A new contractual vehicle specific to the RFC Editor function was negotiated between the parties and took effect on January 1, 2002. It continued in force for more than five years, through a succession of supplements with modified SoWs and budgets, terminating finally on 31

March 2007. This document is the final report this contract, covering the period January 2002- March 2007.

2. RFC Editor Functions

The primary function of the RFC Editor is to edit and publish RFCs. This process is summarized in the following section.

However, the RFC Editor performs many additional tasks related to RFCs. The RFC Editor:

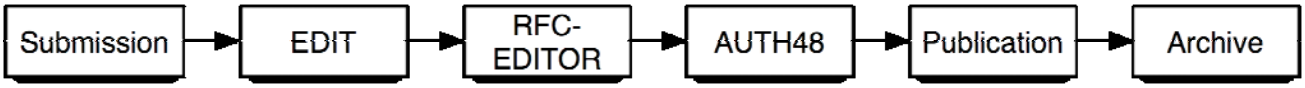
- Maintains a web page that shows a range of information, including:
  - retrieval of index entries for RFCs,
  - a comprehensive search engine for the RFC index,
  - a mirror of the IETF Internet Drafts directory and a search engine for retrieving these Internet Drafts,
  - access to compressed files containing collections of RFCs,
  - a display of the current RFC publication queue,
  - news of significant changes in service or policy,
  - hints to RFC authors about format and content,
  - RFC Editor reports presented to the IETF, and
  - a variety of historical data on publication performance.
- Manages editorial review of documents in the independent submission stream.
- Coordinates with the IESG, the IAB, and the IANA.
- Provides a liaison to attend IAB teleconferences and meetings as well as IESG teleconferences.
- Reports to the plenary session at each IETF meeting (three times a year) on the status of the RFC process.
- Staffs a "help" desk at IETF meetings with two senior editors.
- Presents a tutorial on RFCs and RFC authoring at IETF meetings.
- Experiments with new procedures and policies.
- Responds to all email sent to rfc-editor@rfc-editor.org.

2.1. The Publication Process

2.1.1 Primary Process

The publication process for an RFC contains a series of stages, which are labeled as a set of state transitions devised by ISI. The states were devised to help both the RFC Editor and the community to track the progress of documents in the publication queue.

Appendix A contains a state complete diagram. The diagram below represents the major stages changes for a document as it moves through during the RFC publication..



- Submission

  The RFC Editor at ISI currently publishes five document streams: IETF standards-track, IETF Informational/Experimental, non-IETF Informational/Experimental RFCs (known as independent submissions), IAB (Internet Architecture Board), and IRTF (Internet Research Task Force) documents. Each has its own submission and approval procedure.

- EDIT State

  In the editing process, the RFC Editor applies a checklist of some 30 bullets, to maintain consistency and clarity. An editor looks for:

  - Typographic errors (spelling, capitalization, punctuation) or inconsistencies within the document and other documents on the same subject,
  - Grammar errors and malformed sentences,
  - Excessively long, tortured, or ambiguous sentences,
  - Formatting inconsistent with established guidelines,
  - Inconsistency between citations and references, and
  - Errors in formal languages (e.g., MIBs, ABNF, XML, and ASN.1).

  The editor communicates with the author(s) about any questions that arise in this process, to resolve issues before AUTH48 state.

- RFC-EDITOR State

  When editing is complete, the document enters a final quality control stage, in which the many RFC-specific and IETF-specific rules are checked. At the end of this stage, an RFC number is assigned and inserted in relevant places in the document.

  The RFC Editor works closely with the Internet Assigned Numbers Authority (IANA) to register and insert necessary protocol parameters into documents, prior to publication. ISI has long-standing relations with the IANA and has established communication methods to ensure timely processing. The RFC Editor roles include checking for parameter registration requirements that have not been revealed by the IANA Considerations section and inserting the registered values from IANA into the text. The RFC Editor also checks purported registrations against the IANA web site, as authors sometimes try to assign different values than IANA has chosen.

- AUTH48 State - Authors' Final Review

  When an RFC is ready for publication, the author is asked to review and approve the final text. Ideally, changes during this stage should be small editorial corrections, not extensive edits or technical changes. However, in practice, problems found at this stage range from trivial editorial changes to significant technical fixes. For the latter, the Area Directors and perhaps the working group become involved and must agree. Editing staff are expected to recognize changes that are not solely editorial and require AD approval.

- Publication

  When all responsible parties have agreed, the document is published, which includes putting the publication-format document(s) online, updating necessary index files, notifying IANA of the RFC number for reference purposes (if necessary), and archiving all final source and text files. At this point, the document is announced to the community.

    - RFCs are published on the RFC Editor website. This site includes hyperlinked access to several indices as well as a convenient search engine. The search engine will return a catalog ("index") entry for one or more RFCs, matching on title, author, or number. The RFC Editor also provides access to individual RFCs and to collections of RFCs using SMTP, FTP, and RSync.

2.1.2. Exceptional Cases

The RFC Editor strives to move documents through the above process as quickly as possible, while maintaining a high level of quality. However, there are a number of reasons for significant delay, which also greatly increase the complexity of the editorial task.

- *Normative Reference Hold* - A "normative" reference (i.e., a reference to another standards document) in an RFC must refer to a document that was previously published or must be published concurrently. When a set of related RFCs contain references to each other, all must be held up until they are completely edited and approved, so that they can be published simultaneously. This strict rule resulted from many years of experience with unexpected publication delays that resulted in "dangling" normative references to unpublished documents.

- *Set Hold* - Sometimes authors or working groups request that a set of documents be published simultaneously, even when they are not tied together by normative references.

- *IESG Hold* - The IESG may temporarily suspend or withdraw an IETF document from publication to allow further discussion, clarification, or to remand it to a working group.

- *Author Hold* - The RFC Editor may require action by the author, for a variety of reasons (technical and/or editorial). For example, the editorial process may have revealed some technical issue or discrepancy, or some change in format may be required that only the author can provide.

- *IANA Hold* - The RFC Editor may request action or clarification by the IANA per instructions in the IANA Considerations section of the document.

These delays, which are generally outside the control of the RFC Editor, add to the complexity of managing the editorial workflow.

2.1.3. Submission, Publication, and Archival Formats

ASCII remains the primary format for RFC publication and archiving. The published text is created by the RFC Editor (and can be re-created in the future) using the venerable Unix markup tool nroff. Authors of RFCs are free to use any text preparation method. The published ASCII version and the nroff source are archived by the RFC Editor. The RFC Editor may also publish and archive (but not edit) a subsidiary version of a document in PostScript or PDF, but the primary version is ASCII text.

The following table summarizes the current formats. During the subject contract, the RFC Editor began to experiment with the use of the xml2rfc tool by authors.

Function	Formats
Submission	.txt or .nroff or .xml
Editing	.nroff or .xml followed by .nroff
Publication	.txt (from .nroff source), .txt.pdf*, .pdf**, .ps**
Archiving	Publication formats, .nroff source, and .xml (if submitted)

* The .txt.pdf is published for each RFC, to help Windows users.

** .pdf and .ps are optional. They are created by the author from the final RFC, so that complex diagrams and charts can be included.

2.2 Editorial Policy

The RFC Editor and the IETF have complementary and synergistic goals. The function of the IETF is to produce first-class technical specifications, while the RFC Editor's goal to ensure that the published specifications are expressed in clear, correct, and consistent English prose and symbolism and in a consistent and readable format.

The world of technical publishing has generally accepted standards for the typographic rules for "correct" grammar, punctuation, capitalization, sentence length and complexity, parallelism, etc. The RFC Editor at ISI follows these accepted standards, but with particular exceptions. There are a few specific rule variants that have been imposed on RFCs to avoid ambiguity in complex technical prose and to handle mixtures of text and computer languages. There is also a prime directive that must rule over typographic conventions: do not change the intended meaning of the text.

On the other hand, the RFC Editor strives to respect the long history of individuality in the IETF community. We generally allow variant typography, as long as it is used consistently.

Similarly, we allow either British or American English, but if the usage is inconsistent, we will prefer American English. Thus, although we try to be tolerant of carefully crafted and deliberate alternatives, we have our own preference for the "standard" usage.

Furthermore, the RFC Editor at ISI aims to avoid purely "stylistic" changes that, while formally preferable by general standards, do not advance the primary goals of correct English, accuracy, clarity, and consistency. Examples of such "stylistic" changes might be replacing the conjunction "as" with "because", or removing first-person references. In any case, in the end, RFC authors and the IESG rule. The ISI editors are experienced at maintaining this balance.

Over the past few years, ISI has collected editorial guidelines into the document "Instructions for RFC Authors". This instruction manual helps guide the community through the format, editorial, and publication process. We will continue to expand this document as necessary.

2.3 Quality of Published Documents

The RFC publication process is a balance between quality – freedom from errors – and the significant throughput demands. The IETF community has indicated some willingness to accept a higher errata rate than would be generally acceptable in the publishing world (and perhaps higher than some other standards organizations find acceptable), in the interest of expediency.

However, ISI strongly believes that it is important to minimize the number of errors in published documents, to the extent possible, while maintaining aggressive throughput goals.

The RFC Editor cannot generally be responsible for technical errors, of course. Fortunately, however, correcting purely editorial errors sometimes lead to authors' discovering and fixing content errors. For example, untangling a tortured sentence sometimes leads to the realization that the original sentence was semantically ambiguous, resulting in the replacement with a sentence that is clear and unambiguous.

ISI instituted a number of management techniques to maintain document quality.

- The Authors' Final Review (AUTH48) step described earlier allows authors to reread the edited document and request changes, reducing both editorial and technical errors in published documents.

- The reporting of errata items by readers provides a vital measure of the error rate. This feedback is used to alert the editorial staff to specific editorial issues.

- We created some relatively simple but effective heuristic tools for automating the checking of particular editorial issues. An example is the *matchref* program, which matches citations against references.

2.4 Management and Staffing Experience

Editing is a human-intensive operation that cannot be totally automated. Editing is difficult and at times intellectually challenging, but it is often tedious, requiring close attention for many hours at a time. Proof reading a dense 120-page technical document and catching a serious typographic or consistency error on page 107 requires great discipline as well as skill. ISI has experience in recruiting, training, and motivating a competent and efficient editorial staff.

It takes 2 to 6 months to train even an experienced editor in the editorial rules and conventions specific to RFCs and to the Internet technology. ISI has devoted significant effort to training new people and to upgrading the skills of the current staff. The editorial process can be divided into fairly discrete steps with varying skill levels. For example, we commonly partition the process into three successive phases: (1) copy editing -- marking up the documents using general editorial standards and without specific knowledge about the field, (2) inserting formatting directives, and (3) general editing -- filtering through and inserting edits while applying all the rules and Internet-specific language knowledge. The experienced senior editors can of course

do all of these, and ISI performs flexible assignment of phases according to the skills of available staff.

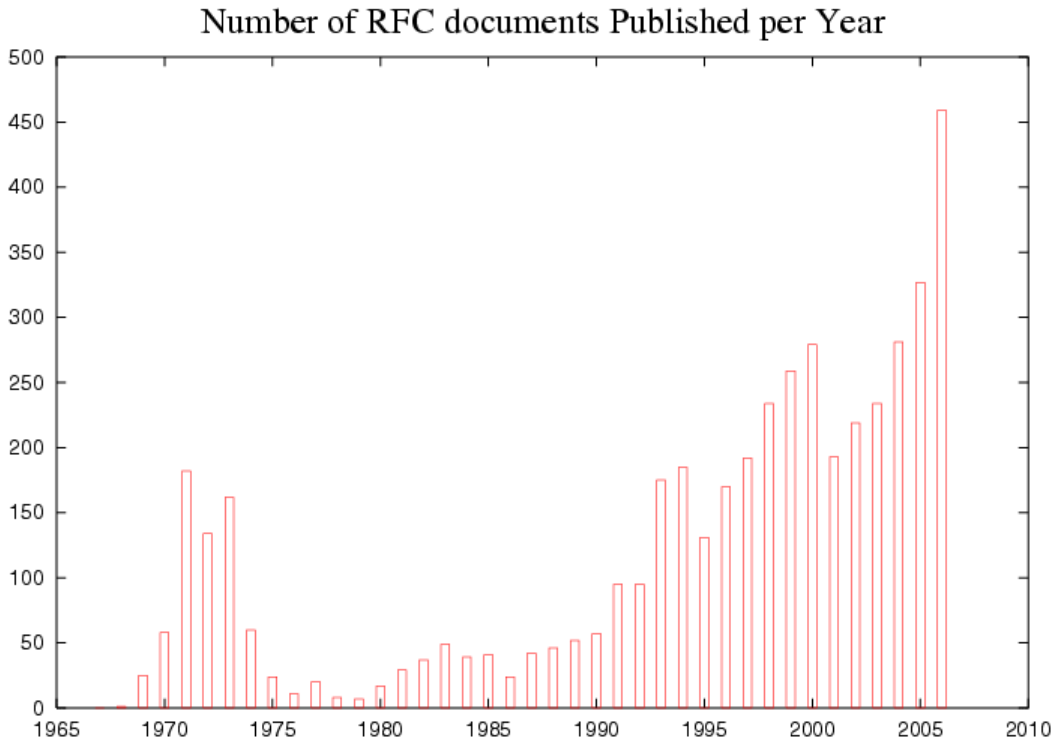
One solution used by ISI to bridge personnel gaps and handling temporary overloads has been to hire editors as temporary workers. Since it is not generally possible to train temp editors in the specifics of RFC rules or Internet conventions, ISI has hired "commodity" copy editors who have no specific knowledge about the field (step (1) above). Later, one of ISI's highly trained senior editors goes over each marked-up document and makes just those changes that are appropriate according to Section 2.2, considering Internet-specific and RFC-specific rules and conventions.

3. RFC Editor Accomplishments During this Period

3.1. Publication

The primary accomplishment under the subject contract is the publication of RFCs. During the period of the subject contract, ISI published 1297 documents, totaling approximately 43400 pages. The average document size has remained remarkably constant at around 30 pages per document, although there is a very wide distribution of sizes.

The following graph shows the number of RFCs published in each year, for the entire history of the RFC series through 2006. The RFC series includes more than 4900 technical documents totaling over 145,000 pages.



3.2. Procedure and Policy Improvements

In addition to publication, during the contract period time the RFC Editor at ISI acted to replace or remove obsolete mechanisms and conventions, updating the RFC Editor function while retaining the essential features. While Jon Postel was RFC Editor, his procedures and policies included many historical aspects that had accumulated since the series began in 1970. Since 1998, the RFC Editor at ISI made many changes to modernize and fine-tune the RFC publication effort. These changes were made carefully and incrementally, but the net result was a very large shift towards efficiency and transparency of the operation. Many of these changes were in response to suggestions and requests from members of the Internet technical community.

This section samples the wide range of changes that ISI has made since 1998.

- **Improved Transparency**
  - ISI completely revamped the RFC Editor web site, to include convenient access to search engines, alternate views of the RFC archive and index, instructions to users, policies, and news.
  - ISI formalized the state diagram for the publication process (see appendix A) and updated it to more accurately track document progress.

- ISI created a user-friendly search engine for RFCs and Internet-Drafts in the ISI repository. This repository is primary for RFCs and mirrors the IETF site for Internet Drafts. The RFC search engine shows STDs and BCPs as well as RFCs, and it shows obsoleted documents in a distinct font.

- ISI installed automated email to authors (cc'ing working group chairs and area directors), so that whenever the publication state changes, state-dependent messages are sent. This provides authors with relevant state change information and automates routine email messages.

- ISI implemented an authors' 48 hours notice message that is cc'ed to working group chairs and area directors. Reminders are sent to authors on a weekly basis, and a message goes to ADs (cc'ing authors and WG chairs) when there has been no response.

- ISI created a script to produce a daily summary of all documents in the RFC Editor queue. This report is used by ISI staff to track documents, and it is supplied weekly to the IANA and to the IETF and IAB chairs.

    - Each document that requires IANA processing is shown with a "*A" flag.
    - Each document that has unpublished normative references is shown with a "*R" flag.

- ISI created scripts to make HTML and XML versions of the RFC index file, in addition to the historic rfc-index.txt file. All three files are available on the RFC Editor web site.

- The RFC Editor at ISI began to send senior members of the team to staff a "help desk" at each IETF meeting. This has allowed users to ask direct questions about status and process, and provided personal interaction with members of the community.

- The RFC Editor began to present tutorials on the RFC series and process, during the Sunday afternoon tutorial schedule at each IETF meeting. These Powerpoint presentations were made available to all on the RFC Editor web site.

- **Improved Services**

    - ISI promoted the Abstract into a first-class part of every document. It may be displayed in the search engine, for example.

- ISI created an archive of errata for published RFCs. This list is on the web site and is linked to the search engine, so that search results include hyperlinks to any errata items.

- ISI instituted the use of htmlwdiff and sends its output to authors to highlight editorial changes.

- ISI created a secondary archive of RFC documents for the convenience of Windows users: PDF facsimiles of each ASCII RFC. They exist in the archive with file names of the form: rfcnnnn.txt.pdf.

- ISI experimented with the use of xml2rfc in the editing process, and we now accept the XML as a submission format (along with the corresponding .txt file) and use it to improve editing efficiency when possible.

- ISI collaborated with the IESG to conduct an experiment in "early" editing of documents while they are still in the working group process.

- The RFC Editor at ISI added additional useful information, such as the category ("status") as originally published, the source Internet Draft file name, and the source IETF Working Group name (if any), to the index database.

- The RFC Editor added additional information to the XML version of the index; this included the draft file name, the Abstract, and the Keywords for each RFC. This new information was not included in the user file rfc-index.txt, because some users have scripts that parse that file to extract data, so we felt we should not change its format. However, we are free to evolve the contents of the XML version of the index.

- **Improved Coordination**

  - ISI worked with IANA to clarify and improve synchronization of IANA protocol parameter assignment with editing. In particular, we modified our procedures to allow parallel processing with IANA assignment.

  - ISI reorganized publication states to clarify the impact of normative reference holds. This involved two additional states, MISSREF and REF. The MISSREF state contains documents that contain at least one normative reference to a document that is not yet submitted to the RFC Editor. The REF state contains documents that contain at least one normative reference whose editing is not yet complete.

  - ISI established an Editorial Board to advise the RFC Editor on independent submissions as well as general editorial policies.

- **Improved Efficiency**

  - ISI worked with the xml2rfc development community to make xml2rfc an effective tool for RFC publication. The xml2rfc source language was primarily designed for ease of document preparation, so it did not provide the fine control over formatting that is required for final markup for publication. However, ISI was able to make effective use of xml2rfc source to do the great majority of editing, but at the final stage prior to publication we convert the source from .xml to .nroff source to perform the fine-tuning of format. Meanwhile, we worked with the development community to follow the RFC formatting conventions.

    This proposal includes further efforts to maximize the power of xml2rfc for RFC publication, and we anticipate significant productivity gains as a result.

  - ISI created an online database of RFC reference entries, to speed up processing. This is also available to authors via the RFC Editor web site.

- **Editorial Procedures and Policies**

  - ISI improved the consistency and accuracy of the editorial process. For example, ISI:

    - Set editorial guidelines for abstracts, titles, Tables of Contents, and abbreviations.
    - Instituted formal language checking for MIBs and XML schemas as well as ABNF.
    - Instituted checking that references to IETF documents are the latest versions.
    - Created tools for routine checking of references and formatting.

  - The editorial staff changed its policy to be much more proactive about communicating with author(s) about editorial issues or things we do not understand, as early as possible in the process.

  - In collaboration with the IESG, ISI designed and implemented the division of references into "Normative" and “Informative”.

- **System Changes**

  - ISI replaced the historical “flat-file” RFC index database with a mySQL database.

- The html file that contains the current publication queue is now generated automatically from this database. This eliminates typographic errors that used to creep into the queue file.

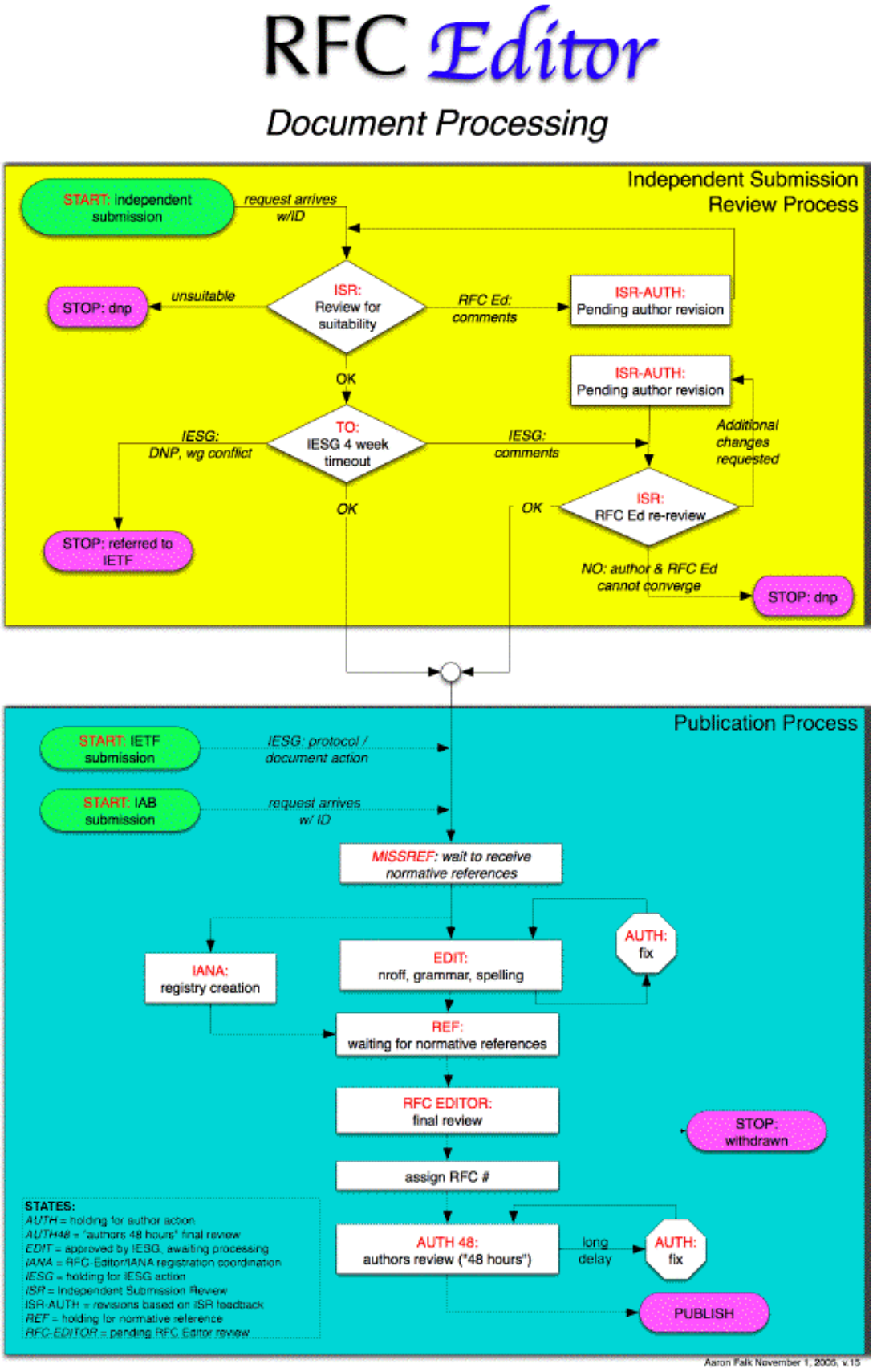
- **Managing Independent Submission Review**

  The RFC Editor at ISI has continued its historical function of publishing a stream of documents that are outside the formal IETF process. These are called "independent" submissions. In practice, the independent submissions span a great range of quality, relevance, and technical depth. To aid in the editorial review process, the RFC Editor established an editorial advisory board. This Editorial Board includes 11 senior people, including both deep Internet technical knowledge and significant experience with academic publication. We expect that it will be further expanded.

4. Conclusion

During the period January 2002 – March 31 2007, ISI performed the RFC Editor function under funding from the Internet Society. In this effort, the RFC Editor maintained historical standards for RFCs despite an ever-increasing publication load, while making substantial progress towards modernizing and streamlining the editing and publishing function.

Appendix A


RFC Editor
Document Processing
Independent Submission Review Process
START: independent submission
request arrives w/ID
ISR: Review for suitability
unsuitable
STOP: dnp
RFC Ed: comments
ISR-AUTH: Pending author revision
OK
ISR-AUTH: Pending author revision
TO: IESG 4 week timeout
IESG: DNP, wg conflict
IESG: comments
Additional changes requested
OK
ISR: RFC Ed re-review
OK
STOP: referred to IETF
NO: author & RFC Ed cannot converge
STOP: dnp
Publication Process
START: IETF submission
IESG: protocol / document action
START: IAB submission
request arrives w/ ID
MISSREF: wait to receive normative references
IANA: registry creation
EDIT: nroff, grammar, spelling
AUTH: fix
REF: waiting for normative references
RFC EDITOR: final review
STOP: withdrawn
assign RFC #
AUTH 48: authors review ("48 hours")
long delay
AUTH: fix
PUBLISH
STATES:
AUTH = holding for author action
AUTH48 = "authors 48 hours" final review
EDIT = approved by IESG, awaiting processing
IANA = RFC Editor/IANA registration coordination
IESG = holding for IESG action
ISR = Independent Submission Review
ISR-AUTH = revisions based on ISR feedback
REF = holding for normative reference
RFC-EDITOR = pending RFC Editor review
Aaron Falk November 1, 2006, v.15