Network Working Group
Request for Comments: 1521
Obsoletes: 1341
Category: Standards Track

N. Borenstein
Bellcore
N. Freed

September 1993

MIME (Multipurpose Internet Mail Extensions) Part One:

Mechanisms for Specifying and Describing the Format of Internet Message Bodies

Status of this Memo

This RFC specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Abstract

STD 11, RFC 822 defines a message representation protocol which specifies considerable detail about message headers, but which leaves the message content, or message body, as flat ASCII text. This document redefines the format of message bodies to allow multi-part textual and non-textual message bodies to be represented and exchanged without loss of information. This is based on earlier work documented in RFC 934, STD 11, and RFC 1049, but extends and revises that work. Because RFC 822 said so little about message bodies, this document is largely orthogonal to (rather than a revision of) RFC 822.

In particular, this document is designed to provide facilities to include multiple objects in a single message, to represent body text in character sets other than US-ASCII, to represent formatted multi-font text messages, to represent non-textual material such as images and audio fragments, and generally to facilitate later extensions defining new types of Internet mail for use by cooperating mail agents.

This document does NOT extend Internet mail header fields to permit anything other than US-ASCII text data. Such extensions are the subject of a companion document [RFC -1522].

This document is a revision of RFC 1341. Significant differences from RFC 1341 are summarized in Appendix H.

THIS PAGE INTENTIONALLY LEFT BLANK.

The table of contents should be inserted after this page.

1 Introduction

Since its publication in 1982, RFC 822 [RFC-822] has defined the standard format of textual mail messages on the Internet. Its success has been such that the RFC 822 format has been adopted, wholly or partially, well beyond the confines of the Internet and the Internet SMTP transport defined by RFC 821 [RFC-821]. As the format has seen wider use, a number of limitations have proven increasingly restrictive for the user community. RFC 822 was intended to specify a format for text messages. As such, non-text messages, such as multimedia messages that might include audio or images, are simply not mentioned. Even in the case of text, however, RFC 822 is inadequate for the needs of mail users whose languages require the use of character sets richer than US ASCII [US-ASCII]. Since RFC 822 does not specify mechanisms for mail containing audio, video, Asian language text, or even text in most European languages, additional specifications are needed.

One of the notable limitations of RFC 821/822 based mail systems is the fact that they limit the contents of electronic mail messages to relatively short lines of seven-bit ASCII. This forces users to convert any non-textual data that they may wish to send into seven-bit bytes representable as printable ASCII characters before invoking a local mail UA (User Agent, a program with which human users send and receive mail). Examples of such encodings currently used in the Internet include pure hexadecimal, uuencode, the 3-in-4 base 64 scheme specified in RFC 1421, the Andrew Toolkit Representation [ATK], and many others.

The limitations of RFC 822 mail become even more apparent as gateways are designed to allow for the exchange of mail messages between RFC 822 hosts and X.400 hosts. X.400 [X400] specifies mechanisms for the inclusion of non-textual body parts within electronic mail messages. The current standards for the mapping of X.400 messages to RFC 822 messages specify either that X.400 non-textual body parts must be converted to (not encoded in) an ASCII format, or that they must be discarded, notifying the RFC 822 user that discarding has occurred. This is clearly undesirable, as information that a user may wish to receive is lost. Even though a user's UA may not have the capability of dealing with the non-textual body part, the user might have some mechanism external to the UA that can extract useful information from the body part. Moreover, it does not allow for the fact that the message may eventually be gatewayed back into an X.400 message handling system (i.e., the X.400 message is "tunneled" through Internet mail), where the non-textual information would definitely become useful again.

This document describes several mechanisms that combine to solve most of these problems without introducing any serious incompatibilities with the existing world of RFC 822 mail. In particular, it describes:

1. A MIME-Version header field, which uses a version number to declare a message to be conformant with this specification and allows mail processing agents to distinguish between such messages and those generated by older or non-conformant software, which is presumed to lack such a field.

2. A Content-Type header field, generalized from RFC 1049 [RFC-1049], which can be used to specify the type and subtype of data in the body of a message and to fully specify the native representation (encoding) of such data.

   2.a. A "text" Content-Type value, which can be used to represent textual information in a number of character sets and formatted text description languages in a standardized manner.

   2.b. A "multipart" Content-Type value, which can be used to combine several body parts, possibly of differing types of data, into a single message.

   2.c. An "application" Content-Type value, which can be used to transmit application data or binary data, and hence, among other uses, to implement an electronic mail file transfer service.

   2.d. A "message" Content-Type value, for encapsulating another mail message.

   2.e An "image" Content-Type value, for transmitting still image (picture) data.

   2.f. An "audio" Content-Type value, for transmitting audio or voice data.

   2.g. A "video" Content-Type value, for transmitting video or moving image data, possibly with audio as part of the composite video data format.

3. A Content-Transfer-Encoding header field, which can be used to specify an auxiliary encoding that was applied to the data in order to allow it to pass through mail transport mechanisms which may have data or character set limitations.

4. Two additional header fields that can be used to further describe the data in a message body, the Content-ID and Content-Description header fields.

MIME has been carefully designed as an extensible mechanism, and it is expected that the set of content-type/subtype pairs and their associated parameters will grow significantly with time. Several other MIME fields, notably including character set names, are likely to have new values defined over time. In order to ensure that the set of such values is developed in an orderly, well-specified, and public manner, MIME defines a registration process which uses the Internet Assigned Numbers Authority (IANA) as a central registry for such values. Appendix E provides details about how IANA registration is accomplished.

Finally, to specify and promote interoperability, Appendix A of this document provides a basic applicability statement for a subset of the above mechanisms that defines a minimal level of "conformance" with this document.

> *HISTORICAL NOTE: Several of the mechanisms described in this document may seem somewhat strange or even baroque at first reading. It is important to note that compatibility with existing standards AND*

robustness across existing practice were two of the highest priorities of the working group that developed this document. In particular, compatibility was always favored over elegance.

MIME was first defined and published as RFCs 1341 and 1342 [RFC-1341] [RFC-1342]. This document is a relatively minor updating of RFC 1341, and is intended to supersede it. The differences between this document and RFC 1341 are summarized in Appendix H. Please refer to the current edition of the "IAB Official Protocol Standards" for the standardization state and status of this protocol. Several other RFC documents will be of interest to the MIME implementor, in particular [RFC 1343], [RFC-1344], and [RFC-1345].

2 Notations, Conventions, and Generic BNF Grammar

This document is being published in two versions, one as plain ASCII text and one as PostScript[1] . The latter is recommended, though the textual contents are identical. An Andrew-format copy of this document is also available from the first author (Borenstein).

Although the mechanisms specified in this document are all described in prose, most are also described formally in the modified BNF notation of RFC 822. Implementors will need to be familiar with this notation in order to understand this specification, and are referred to RFC 822 for a complete explanation of the modified BNF notation.

Some of the modified BNF in this document makes reference to syntactic entities that are defined in RFC 822 and not in this document. A complete formal grammar, then, is obtained by combining the collected grammar appendix of this document with that of RFC 822 plus the modifications to RFC 822 defined in RFC 1123, which specifically changes the syntax for ‘return’, ‘date’ and ‘mailbox’.

The term CRLF, in this document, refers to the sequence of the two ASCII characters CR (13) and LF (10) which, taken together, in this order, denote a line break in RFC 822 mail.

The term "character set" is used in this document to refer to a method used with one or more tables to convert encoded text to a series of octets. This definition is intended to allow various kinds of text encodings, from simple single-table mappings such as ASCII to complex table switching methods such as those that use ISO 2022’s techniques. However, a MIME character set name must fully specify the mapping to be performed.

The term "message", when not further qualified, means either the (complete or "top-level") message being transferred on a network, or a message encapsulated in a body of type "message".

---

[1] PostScript is a trademark of Adobe Systems Incorporated.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.