

Intended Status: Experimental

Internet Draft: Mailing Lists and Internationalized
Email Addresses

R. Gellens
Qualcomm

Document: [draft-ietf-eai-mailinglist-07.txt](#)

Expires: December 28, 2010

June 28, 2010

Mailing Lists and Internationalized Email Addresses

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

Copyright Notice

Copyright (c) 2010 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 BSD License.

Internet Draft

Mailing Lists and UTF-8 Mail Addresses

March 2010

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Abstract

This document describes considerations for mailing lists with the introduction of internationalized email addresses.

This document makes some specific recommendations on how mailing lists should act in various situations.

Internet Draft Mailing Lists and UTF-8 Mail Addresses March 2010

Table of Contents

1.	Introduction	3
2.	Conventions Used in this Document	5
3.	Scenarios Involving Mailing Lists	5
4.	Capabilities and Requirements	6
5.	List Header Fields	7
6.	Further Discussion	8
7.	IANA Considerations	9
8.	Security Considerations	9
9.	Acknowledgments	9
10.	Normative References	9
11.	Informative References	10
12.	Authors' Addresses	10
	Appendix A : Changes from Previous Version	10

[1.](#) Introduction

This document describes considerations for mailing lists with the introduction of internationalized email addresses.

Mailing lists are an important part of email usage and collaborative communications. The introduction of internationalized email addresses affects mailing lists in three main areas: (1) transport (receiving and sending messages); (2) message headers of received and retransmitted messages; and (3) mailing list operational policies.

A mailing list is a mechanism whereby a message may be distributed to multiple recipients by sending to one address. An agent (typically not a human being) at that single address receives the message and then causes the message to be redistributed to a list of recipients. This agent sets the envelope return address of the redistributed message to a different address from that of the

original message. Using a different envelope return address (reverse-path) directs error (and other automatically generated) messages to an error handling address associated with the mailing list. (This avoids having error and other automatic messages go to the original sender, who typically doesn't control the list and hence can't do anything about them.)

In most cases, the mailing list agent redistributes a received message to its subscribers as a new message, that is, conceptually it uses message submission [[submission](#)] (as did the sender of the original message). The exception, where the mailing list is not a separate agent that receives and redistributes messages in separate transactions, but is instead an expansion step within an SMTP transaction where one local address expands to multiple local or

non-local addresses, is out of scope for this document.

Some mailing lists alter message header fields, while others do not. A number of standardized list-related header fields have been defined, and many lists add one or more of these headers. Separate from these standardized list-specific header fields, and despite a history of interoperability problems from doing so, some lists alter or add header fields in an attempt to control where replies are sent. Such lists typically add or replace the "Reply-To" field and some add or replace the "Sender" field. Poorly-behaved lists may alter or replace other fields, including "From".

Among these list-specific header fields are those specified in [RFC 2369](#) ("The Use of URLs as Meta-Syntax for Core Mail List Commands and their Transport through Message Header Fields") [[List-*](#)] and [RFC 2919](#) ("List-Id: A Structured Field and Namespace for the Identification of Mailing Lists") [[List-ID](#)]. For more information, see [Section 5](#).

While the mail transport protocol does not differ between regular email recipients and mailing list recipients, lists have special considerations with internationalized email addresses because they retransmit messages composed by other agents to potentially many recipients.

There are considerations for internationalized email addresses in the envelope as well as in header fields of redistributed messages. In particular, an internationalized message cannot be downgraded

unless all envelope addresses are available in ASCII (that is, each address either is ASCII or has an alt-address [[UTF8SMTP](#)]).

With mailing lists, there are two different types of considerations: first, the purely technical ones involving message handling, error cases, downgrades, and the like, and second, those that arise from the fact that humans use mailing lists to communicate. As an example of the first, mailing lists might choose to reject all messages from internationalized addresses that lack an alt-address, or even all internationalized messages that can not be downgraded. As an example of the second, a user who sends a message to a list often is unaware of the list membership. In particular, the user often doesn't know if the members are UTF-8 mail users or not, and often neither the original sender nor the recipients personally know each other. As a consequence of this, remedies that may be readily available for one-to-one communication might not be appropriate when dealing with mailing lists. For example, if a user sends a message which is undeliverable, normally the telephone, instant messaging, or other forms of communication are available to obtain a working address. With mailing lists, the users may not have any recourse. Of course, with mailing lists, the original sender usually does not

know if the message was successfully received by any list members, or if it was undeliverable to some.

Conceptually, a mailing list's internationalization can be divided into three capabilities: First, does it have a UTF-8 submission or return-path address? Second, does it accept subscriptions to UTF-8 addresses? And third, does it accept [[UTF8SMTP](#)] messages? This is explored in [Section 4](#).

A brief discussion on a few additional considerations for mailing list operation is in [Section 6](#).

[2](#). Conventions Used in this Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [[KEYWORDS](#)].

[3](#). Scenarios Involving Mailing Lists

Generally (and exclusively within the scope of this document), an original message is sent to a mailing list as a completely separate and independent transaction from the mailing list agent sending the retransmitted message to one or more list recipients. In both cases, the message might have only one recipient, or might have multiple recipients. That is, the original message might be sent to additional recipients as well as the mailing list agent, and the mailing list might choose to send the retransmitted message to each list recipient in a separate message submission [[submission](#)] transaction, or might choose to include multiple recipients per transaction. (Often, mailing lists are constructed to work in cooperation with, rather than include the functionality of, a message submission server [[submission](#)], and hence the list transmits to a single submission server one copy of the retransmitted message, with all list recipients specified in the SMTP envelope. The submission server then decides which recipients to include in which transaction.)

The retransmitted message sent by the mailing list to its subscribers might need to be downgraded [[EAI-Downgrade](#)]. In order for a downgrade to be possible, the return path set by the mailing list agent must be an ASCII address or have alt-address [[UTF8SMTP](#)] specified. In addition, the recipient addresses need to have ASCII addresses available. It may be advisable for mailing list operators to pre-obtain an alt-address for all its internationalized member addresses.

In the case where a member or non-member with an internationalized email address is sending to a mailing list, no alt-address [[UTF8SMTP](#)] is specified, and a downgrade is required, the message cannot be delivered. To protect against this, a UTF8SMTP-aware [[UTF8SMTP](#)] mailing list might prefer to reject submissions from internationalized email addresses that lack an alt-address.

(Note that the situation is not unique to mailing lists. Mail relays that are UTF8SMTP-aware will potentially encounter the same situation.) Further discussions are included in [section 6](#) of this document.

[4.](#) Capabilities and Requirements

There are three primary internationalization capabilities of mailing lists: First, does it have a UTF-8 submission or return-path address? Second, does it allow subscriptions from UTF-8 addresses? And third, does it accept [\[UTF8SMTP\]](#) messages?

In theory, any list can support any combination of these. In practice, only some offer any benefit. For example, neither allowing UTF-8 addresses to subscribe, nor accepting UTF8SMTP messages, makes much sense without the other (an all-ASCII address might or might not be capable of receiving UTF8SMTP messages, but a UTF-8 address of necessity needs to accept UTF8SMTP messages). Likewise, there is no real benefit to a list in using a UTF-8 submission address unless it also accepts UTF8SMTP messages and permits UTF-8 addresses to subscribe.

However, requirements for lists can be discussed separately for each of the three capabilities.

1. If the list uses a UTF-8 submission or return-path address, it SHOULD specify an alt-address [\[UTF8SMTP\]](#) for it. Clearly, it needs to sit behind a UTF8SMTP-enabled final-delivery SMTP server [\[UTF8SMTP\]](#) and delivery agent. Likewise, if a list uses a UTF-8 return-path address, then its MSA [\[submission\]](#) needs to support UTF8SMTP.

The list's return-path address is usually separate from its submission address (so that delivery reports and other automatically-generated messages are not sent to the submission address). For reliability in receiving delivery status notifications, a list MAY choose to use an all-ASCII return-path even if it uses a UTF-8 submission address. If the list does use a UTF-8 return path, it MUST specify an alt-address [\[UTF8SMTP\]](#) (or else there is a high risk of being unable to receive non-delivery reports).

There are also implications for the List-* headers (see below).

2. If it allows UTF-8 addresses to subscribe, it MAY require an alt-address [\[UTF8SMTP\]](#) to be specified for each UTF-8 subscriber.

Naturally, if it permits UTF-8 addresses to subscribe, it needs a mechanism to accept subscription requests from such addresses

(preferably specified in the form <utf8@utf8 <ascii@ascii>>). In order to send email to its subscribers who have UTF-8 addresses, its MSA needs to support [[UTF8SMTP](#)].

3. If it accepts UTF8SMTP messages, the MTAs and MDA in its inbound path need to support UTF8SMTP.

5. List Header Fields

A number of header fields specifically for mailing lists have been introduced in [RFC 2369](#) and [RFC 2919](#). These include, for example:

```
List-Id: List Header Mailing List <list-header.example.com>
List-Help: <mailto:list@example.com?subject=help>
List-Unsubscribe: <mailto:list@example.com?subject=unsubscribe>
List-Subscribe: <mailto:list@example.com?subject=subscribe>
List-Post: <mailto:list@example.com>
List-Owner: <mailto:listmom@example.com> (Contact Person for Help)
List-Archive: <mailto:archive@example.com?subject=index%20list>
```

As described in [RFC 2369](#), "The contents of the list header fields mostly consist of angle-bracket ('<', '>') enclosed URLs, with internal whitespace being ignored." [List-*] Whereas [RFC2919](#) specifies that, "The list identifier will, in most cases, appear like a host name in a domain of the list owner." [[List-ID](#)]

These mailing list header fields contain URLs. The most common schemes are generally HTTP, HTTPS, mailto, and FTP. Schemes which permit both URI and IRI [[IRI](#)] forms should use the URI-encoded form described in [[IRI](#)]. Future work may extend these header fields or define replacements to directly support non-encoded UTF-8 in IRIs (for example, [[mailto-bis](#)]), but in the absence of such extension or replacement, non-ASCII characters can only appear within when encoded as ASCII. Note that discussion on whether internationalized domain names should be percent-encoded or puny-coded, is ongoing; see [[IRI-bis](#)]

Even without these header fields being extended to support UTF-8, some special provisions may be helpful when downgrading. In particular, if a List-* header contains a UTF-8 mailto (even encoded in ASCII) followed by an ASCII mailto, it may be advisable to not

only copy and preserve the original header as usual (ENCAPSULATION method of [[EAI-Downgrade](#)]), but also to edit the header to remove the UTF-8 address. Otherwise, a client might run into trouble if the decoded mailto results in a non-ASCII address.

When mailing lists use a UTF-8 form of a List-* header, an ASCII form SHOULD also be used. These headers are vital to good operations and use of mailing lists; caution is called for when considering how to form and use these headers in a non-ASCII environment.

The most commonly-used URI schemes in List-* headers tend to be HTTP and mailto. The current specification for mailto does not permit unencoded UTF-8 characters, although work has been proposed to extend or more likely replace mailto in order to permit this. For mailto URIs, a separate consideration is how to include an alternate ASCII address (alt-address) [[UTF8SMTP](#)] for a UTF-8 address. Note that the existing ability to specify multiple URLs within each List-* header field provides one solution.

[List-*] says:

A list of multiple, alternate, URLs MAY be specified by a comma-separated list of angle-bracket enclosed URLs. The URLs have order of preference from left to right. The client application should use the left most protocol that it supports, or knows how to access by a separate application.

When a UTF-8 mailto is used in a List-* header field, an alt-address [[UTF8SMTP](#)], if available, SHOULD be supplied.

The List-ID header field uniquely identifies a list. The intent is that the value of this header remain constant, even if the machine or system used to operate or host the list changes. This header field is often used in various filters and tests, such as client-side filters, Sieve filters, and so forth. Such filters and tests may not properly compare a non-ASCII value which has been encoded into ASCII. In addition to these comparison considerations, it is generally desirable that this header field contain something meaningful that users can type in. However, ASCII encodings of non-ASCII characters are unlikely to be meaningful to users or easy for them to accurately type.

6. Further Discussion

While mailing lists do not create a significant additional burden to the deployment of internationalized email address functionalities, there are some specific areas that need to be considered when the operator of a mailing list or of a final delivery MTA that serves a

mailing list upgrades to internationalized mail.

Mailing lists face additional complexity since they redistribute messages composed by other agents. Hence, they may be asked to accept a message with non-ASCII headers composed by a UTF8SMTP-aware user agent [[UTF8SMTP](#)], and redistribute it to UTF-8 mail and all-ASCII mail users via systems that are not UTF8SMTP-aware.

1. Obtaining Downgrade Information -- for a mailing list, or mail relay server for that matter, that is UTF8SMTP-aware, receiving mail from an internationalized email address, the alt-address [[UTF8SMTP](#)] is not required from the sending MTA for the transport to be complete. When the mailing list then retransmits the message to its subscribers, it may encounter paths where a downgrade is needed (if a relay or final MSA does not supports UTF8SMTP). In order to mitigate this situation, the mailing list might perhaps decide to reject all incoming mail from an internationalized email address that lacks an alt-address. However, note that in general, downgrades are not expected to be the normal case.

2. Downgrading Considerations for mailto URLs -- UTF-8 addresses in mailto links in List-* headers will be easier to downgrade if they contain an alt-address [[UTF8SMTP](#)].

[7.](#) IANA Considerations

None.

[8.](#) Security Considerations

Security considerations are discussed in the Framework document [[EAI-Framework](#)]. No further security considerations are raised by this document.

[9.](#) Acknowledgments

Edmon Chung of Afilias wrote the original version of this document. Thanks to Harald Alvestrand for his extensive comments. Ted Hardie contributed helpful text on IRIs. Last-call comments from S. Moonesamy and Amanda Baber, plus shepherd review by Pete Resnick,

improved the document.

10. Normative References

[EAI-Framework] J. Klensin and Y. Ko, "Overview and Framework for Internationalized Email", [RFC 4952](#), July 2007.

[KEYWORDS] "Key words for use in RFCs to Indicate Requirement Levels", S. Bradner, [RFC 2119](#), [BCP 14](#), March 1997.

[List-*] G. Neufeld and J. Baer, "The Use of URLs as Meta-Syntax for Core Mail List Commands and their Transport through Message Header Fields", [RFC 2369](#), July 1998

[List-ID] R. Chandhok and G. Wenger, "List-Id: A Structured Field and Namespace for the Identification of Mailing Lists", [RFC 2919](#), March 2001

[submission] R. Gellens and J. Klensin, "Message Submission for Mail", [RFC 4409](#), April 2006.

[UTF8SMTP] J. Yao and W. Mao, "SMTP Extension for Internationalized Email Addresses", [RFC 5336](#), September 2008.

11. Informative References

[mailto-bis] M. Duerst and L. Masinter, "The mailto URI scheme", [draft-duerst-mailto-bis-xx](#) (work in progress).

[IRI] M. Duerst and M. Suignard, "Internationalized Resource Identifiers (IRIs)", [RFC 3987](#), January 2005

[IRI-bis], Internationalized Resource Identifiers IETF Working Group, <http://datatracker.ietf.org/wg/iri/>

[EAI-Downgrade] K. Fujiwara and Y. Yoneya, "Downgrading Mechanism for Email Address Internationalization", [RFC 5504](#), March 2009.

[12.](#) Authors' Addresses

Randall Gellens
QUALCOMM Incorporated
5775 Morehouse Drive
San Diego, CA 92121
rg+iETF@qualcomm.com

R. Gellens

Expires September 2010

[Page 10]

Internet Draft

Mailing Lists and UTF-8 Mail Addresses

March 2010

Appendix A: Changes from Previous Version

THIS SECTION TO BE REMOVED PRIOR TO PUBLICATION.

Changes made from version -06 to -07:

- o Added explanatory text in [section 2](#) about why a list server's MSA needs to support UTF8SMTP.
- o Changed [section 5](#) to say that an alt-address should "be supplied" instead of "immediately follow" the UTF-8 address.
- o Moved [List-*], [[List-ID](#)], [[submission](#)], and [[UTF8SMTP](#)] references from Informative to Normative.
- o Switched former sections [1](#) ("Conventions used in this document") and [2](#) ("Introduction").
- o Added new sentence to start of Introduction.
- o Changed "the message header" to "message header fields" in Introduction.
- o Added reference to [RFC 5336](#) after some uses of "alt-address".
- o Made case consistent in "alt-address".
- o Used sanctioned instead of ad-hoc example domains.
- o Replaced "i18mail" with "UTF-8 mail".
- o Added additional people to Acknowledgments.

Changes made from version -05 to -06:

- o Reworded IRI text per Ted Hardie's suggestion.
- o Added reference to IRI WG.
- o Deleted remaining fEditor's Note.

Changes made from version -04 to -05:

- o Added [[submission](#)] as informative reference.

- o Normalized "UTF-8" for descriptive terms except as part of hyphenated compounds.
- o Reworded Introduction text on envelope addresses and ASCII.
- o Added informative reference to [RFC 5336](#).
- o Reworded discussion of List-ID in "List Header Fields" section to more clearly explain the risks of using ASCII encodings of non-ASCII characters.
- o Added text to "Further Discussion" clarifying that downgrade might be needed if an MAA between the list and a subscriber does not support UTF8SMTP.
- o Corrected some references.

Changes made from version -03 to -04:

- o Rewrote text in [Section 5](#) on List-* headers. Added new text specifically on List-ID. Noted that currently, IRIs in List-* headers must be encoded as ASCII.

Changes made from version -02 to -03:

- o Deleted [Section 6](#).
- o Restored missing text in third paragraph from the end of [Section 3.1](#).
- o Deleted broken suggestion in [Section 5](#).
- o Additional text fixes.
- o Reworked text on List-* header fields.
- o Removed most Editor's Notes, including deletion of all text that had been followed by an Editor's Note asking if it was useful.
- o Modified Abstract.
- o Edited Sections [3](#), [4](#), and [5](#).

Changes made from version -01 to -02:

- o Significant changes throughout the document. Sorry.

Changes made from version -00 to -01:

- o Fixed SMTP envelope versus message header confusion.
- o Fixed erroneous mailing list operation text.
- o Removed references to ATOMIC.
- o Removed unneeded scenarios.
- o Added discussion of human considerations which arise with lists.
- o Fixed some typos.

[R.](#) Gellens

Expires September 2010

[Page 12]