

Network Working Group
Internet-Draft
Obsoletes: [RFC5336](#) (if approved)
Updates: [RFC5321](#) and 5322
(if approved)
Intended status: Standards Track
Expires: October 13, 2011

J. Yao
W. Mao
CNNIC
April 11, 2011

**SMTP Extension for Internationalized Email Address
draft-ietf-eai-rfc5336bis-09.txt**

Abstract

This document specifies an SMTP extension for transport and delivery of email messages with internationalized email addresses or header information.

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 October 13, 2011.

Copyright Notice

Copyright (c) 2011 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.

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.

Table of Contents

1.	Introduction	4
1.1.	Role of This Specification	4
1.2.	Terminology	4
1.3.	Updates to Other Specifications	5
2.	Overview of Operation	5
3.	Mail Transport-Level Protocol	5
3.1.	Framework for the Internationalization Extension	5
3.2.	The UTF8SMTPbis Extension	6
3.3.	Extended Mailbox Address Syntax	7
3.4.	MAIL Command Parameter Usage	10
3.5.	Non-ASCII addresses and Reply-codes	10
3.6.	Body Parts and SMTP Extensions	11
3.7.	Additional ESMTP Changes and Clarifications	11
3.7.1.	The Initial SMTP Exchange	11
3.7.2.	Mail eXchangers	12
3.7.3.	Trace Information	12
3.7.4.	UTF-8 Strings in Replies	14
4.	IANA Considerations	15
5.	Security Considerations	17
6.	Acknowledgements	18
7.	Change History	18
7.1.	draft-yao-eai-rfc5336bis: Version 00	18
7.2.	draft-ietf-eai-rfc5336bis: Version 00	19
7.3.	draft-ietf-eai-rfc5336bis: Version 01	19
7.4.	draft-ietf-eai-rfc5336bis: Version 02	19
7.5.	draft-ietf-eai-rfc5336bis: Version 03	19
7.6.	draft-ietf-eai-rfc5336bis: Version 04	19
7.7.	draft-ietf-eai-rfc5336bis: Version 05	19
7.8.	draft-ietf-eai-rfc5336bis: Version 06	19
7.9.	draft-ietf-eai-rfc5336bis: Version 07	19
7.10.	draft-ietf-eai-rfc5336bis: Version 08	19
7.11.	draft-ietf-eai-rfc5336bis: Version 09	20
8.	References	20
8.1.	Normative References	20
8.2.	Informative References	21
	Authors' Addresses	22

1. Introduction

The Simple Mail Transfer Protocol [[RFC5321](#)] provides a negotiation mechanism about service extension by which SMTP clients can discover SMTP server capabilities and make decisions for further processing. This document uses this mechanism and specifies an SMTP extension to permit internationalized email addresses (see [Section 1.2](#)) in the SMTP envelope, and Unicode characters encoded in UTF-8 [[RFC3629](#)] in the headers. An extended overview of the extension model for internationalized email addresses and the email header appears in [[RFC4952bis](#)], referred to as "the framework document" or just as "framework" elsewhere in this specification.

[[anchor1: Note in Draft and to RFC Editor: The keyword represented in this document by "UTF8SMTPbis" (and in the XML source byUTF8SMTPbis) is a placeholder. The actual keyword will be assigned when the standards track SMTP extension in this series [RFC5336bis-SMTP] is approved for publication and should be substituted here. This paragraph should be treated as normative reference to that SMTP extension draft, creating a reference hold until it is approved by the IESG. This paragraph should be removed before RFC publication.]]

1.1. Role of This Specification

The framework document [[RFC4952bis](#)] specifies the requirements for, and describes components of, full internationalization of electronic mail. A thorough understanding of the information in that document and in the base Internet email specifications [[RFC5321](#)] [[RFC5322](#)] is necessary to understand and implement this specification.

This document specifies an element of the email internationalization work, specifically the definition of an SMTP extension for internationalized email address transport delivery.

1.2. Terminology

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 [RFC 2119](#) [[RFC2119](#)].

The terms "UTF-8 string" or "UTF-8 character" are used to refer to Unicode characters encoded in UTF-8. All other specialized terms used in this specification are defined in the framework document or in the base Internet email specifications. In particular, the terms "ASCII address", "internationalized email address", "non-ASCII address", "UTF8SMTPbis", "internationalized message", and "message" are used in this document according to the definitions in the framework document.

Non-ASCII characters or strings referred in this document MUST be expressed in UTF-8, a standard Unicode Encoding Form.

This specification uses Augmented BNF (ABNF) rules [[RFC5234](#)], with some modifications. The modified rules are defined in this specification. When a new rule has a name starting with "u", it is a small modification to an older rule. Rules that are undefined here can be found from [[RFC5234](#)] or [[RFC5321](#)] or [[RFC5322](#)] under the same names.

1.3. Updates to Other Specifications

This specification modifies [RFC 5321](#) by permitting internationalized email address in the envelope. It also updates some syntax rules defined in [RFC 5321](#). It modifies [RFC 5322](#) by permitting data formats defined in [[RFC5335bis](#)]. It does not modify the 8BITMIME specification [[RFC6152](#)] in any way, but it does require that the 8BITMIME extension be announced by the EAI-aware SMTP server and used with "BODY=8BITMIME" by the SMTP client.

2. Overview of Operation

This specification describes an optional extension to the email transport mechanism that permits non-ASCII characters in both the envelope and header fields of messages, which are encoded in UTF-8. The extension is identified with the token "UTF8SMTPbis".

The EAI UTF-8 header specification [[RFC5335bis](#)] provides the details of email header features enabled by this extension

3. Mail Transport-Level Protocol

3.1. Framework for the Internationalization Extension

The following service extension is defined:

1. The name of the SMTP service extension is "Email Address Internationalization".
2. The EHLO keyword value associated with this extension is "UTF8SMTPbis".
3. No parameter values are defined for this EHLO keyword value. In order to permit future (although unanticipated) extensions, the EHLO response MUST NOT contain any parameters for this keyword. The EAI-aware SMTP client MUST ignore any parameters if they appear for this keyword; that is, the EAI-aware SMTP client MUST behave as if the parameters do not appear. If an SMTP server includes UTF8SMTPbis in its EHLO response, it MUST be fully

- compliant with this version of this specification.
4. One OPTIONAL parameter "UTF8SMTPbis" is added to the MAIL command. The parameter has no value. If this parameter is set in the MAIL command, it indicates that the SMTP client is EAI-aware and asserts that the envelope includes the non-ASCII address or the message being sent is internationalized message or the message being sent needs the UTF8SMTPbis support.
 5. The maximum length of a MAIL command line is increased by 12 characters by the possible addition of the UTF8SMTPbis parameter. [[anchor6: RFC Editor: the number '12' will be replaced by the new number (1 space + length of the new keyword supposed to replace "UTF8SMTPbis").]]
 6. One OPTIONAL parameter "UTF8SMTPbis" is added to the VRFY and EXPN commands. The parameter UTF8SMTPbis has no value. The parameter indicates that the SMTP client can accept Unicode characters in UTF-8 encoding in replies from the VRFY and EXPN commands.
 7. No additional SMTP verbs are defined by this extension.
 8. Servers offering this extension MUST provide support for, and announce, the 8BITMIME extension [[RFC6152](#)].
 9. The reverse-path and forward-path of the SMTP MAIL and RCPT commands are extended to allow Unicode characters encoded in UTF-8 in mailbox names (addresses).
 10. The mail message body is extended as specified in [[RFC5335bis](#)].
 11. The UTF8SMTPbis extension is valid on the submission port [[RFC4409](#)], and can be used with LMTP [[RFC2033](#)].

3.2. The UTF8SMTPbis Extension

An SMTP server that announces this UTF8SMTPbis extension MUST be prepared to accept a UTF-8 string [[RFC3629](#)] in any position in which [RFC 5321](#) specifies that a <mailbox> can appear. Although the characters in the <local-part> are permitted to contain non-ASCII characters, actual parsing of the <local-part>, and the delimiters used, are unchanged from the base email specification [[RFC5321](#)]. Any domain names to be looked up in the DNS MUST allow for [[RFC5890](#)] behavior. When doing lookups, the EAI-aware SMTP server MUST either use a Unicode aware DNS library, or transform it to A-label defined in [[RFC5890](#)].

An SMTP client that receives the UTF8SMTPbis extension keyword in response to the EHLO command MAY transmit mailbox names within SMTP commands as internationalized strings in UTF-8 form. It MAY send a UTF-8 header [[RFC5335bis](#)] (which may also include mailbox names in UTF-8). It MAY transmit the domain parts of mailbox names within SMTP commands or the message header as A-labels or U-labels [[RFC5890](#)]. The presence of the UTF8SMTPbis extension does not change [RFC 5321](#) server relaying behaviors.

If the UTF8SMTPbis SMTP extension is not offered by the SMTP server, the EAI-aware SMTP client MUST NOT transmit an internationalized email address and MUST NOT transmit a mail message containing internationalized mail headers as described in [\[RFC5335bis\]](#) at any level within its MIME structure [\[RFC2045\]](#) and [\[RFC2047\]](#). (For this paragraph, the internationalized domain name in the form of A-labels as specified in IDNA definitions [\[RFC5890\]](#) is not considered to be "internationalized".) Instead, if an EAI-aware SMTP client (EAI-aware SMTP sender) attempts to transfer an internationalized message and encounters an SMTP server that does not support the extension, it MUST make one of the following three choices and the priority order is 1, 2 and 3.

1. It MAY either reject the message during the SMTP transaction or accept the message and then generate and transmit a notification of non-deliverability. Such notification MUST be done as specified in [RFC 5321](#) [\[RFC5321\]](#), [RFC 3464](#) [\[RFC3464\]](#), and the EAI delivery status notification (DSN) specification [\[RFC5337bis\]](#).
2. If and only if the EAI-aware SMTP client (sender) is a Message Submission Agent ("MSA") [\[RFC4409\]](#) [\[RFC5598\]](#), it MAY rewrite the envelope, headers, or message material to make them entirely ASCII [\[ASCII\]](#) and consistent with the provisions of [RFC 5321](#) [\[RFC5321\]](#) and [RFC 5322](#) [\[RFC5322\]](#).
3. It MAY find an alternate route to the destination that permits UTF8SMTPbis. That route MAY be discovered by trying alternate Mail eXchanger (MX) hosts (using preference rules as specified in [RFC 5321](#)) or using other means available to the EAI-aware SMTP client.

This document applies only when an EAI-aware SMTP client is trying to send an internationalized message to an EAI-aware SMTP server. For all other cases, and for addresses and messages that do not require an UTF8SMTPbis extension, EAI-aware SMTP clients and servers do not change the behavior specified in [\[RFC5321\]](#).

An EAI-aware MUA/MSA sending to a legacy SMTP server [\[RFC5321\]](#) and [\[RFC5322\]](#) MAY convert an ASCII@U-label [\[RFC5890\]](#) address into the format of ASCII@A-label [\[RFC5890\]](#) if the email address is in the format of ASCII@U-label.

[3.3. Extended Mailbox Address Syntax](#)

[RFC 5321, Section 4.1.2](#), defines the syntax of a <mailbox> entirely in terms of ASCII characters.

The key changes made by this specification include:

- o Change the definition of <Domain> to permit both the [RFC 5321](#) definition and a UTF-8 string representing a DNS label that is conforming with IDNA definitions [[RFC5890](#)].
- o Change the definition of <Local-part> to permit both the [RFC 5321](#) definition and a UTF-8 string. That string MUST NOT contain any of the ASCII characters (either graphics or controls) that are not permitted in <atext>; it is otherwise unrestricted.

According to the description above, the syntax of an internationalized email mailbox name (address) is defined in ABNF [[RFC5234](#)] as follows.

uMailbox = uLocal-part "@" (uDomain / address-literal)
; Replace Mailbox in [RFC 5321, Section 4.1.2](#)

address-literal = <Defined in [Section 4.1.2 of RFC 5321](#)>

uLocal-part = uDot-string / uQuoted-string
; MAY be case-sensitive
; Replace Local-part in [RFC 5321, Section 4.1.2](#)

uDot-string = uAtom *("." uAtom)
; Replace Dot-string in [RFC 5321, Section 4.1.2](#)

uAtom = 1*ucharacter
; Replace Atom in [RFC 5321, Section 4.1.2](#)

ucharacter = atext / UTF8-non-ascii

atext = <Defined in [Section 3.2.3 of RFC 5322](#)>
; Same definition with atext in [RFC 5321, Section 4.1.2](#)

uQuoted-string = DQUOTE *uQcontentSMTP DQUOTE
; Replace Quoted-string in [RFC 5321, Section 4.1.2](#)

DQUOTE = <Defined in [appendix B.1 of RFC 5234](#)>

uQcontentSMTP = qtextSMTP / quoted-pairSMTP / UTF8-non-ascii

qtextSMTP = <Defined in [Section 4.1.2 of RFC 5321](#)>

quoted-pairSMTP = <Defined in [Section 4.1.2 of RFC 5321](#)>

uDomain = sub-udomain *("." sub-udomain)
; Replace Domain in [RFC 5321, Section 4.1.2](#)

sub-udomain = uLet-dig [uLdh-str]
; Replace sub-domain in [RFC 5321, Section 4.1.2](#)
; Note that uDomain may not be resolvable if sub-udomain is not a
; valid U-Label or LDH Label as defined in [RFC 5890](#)

uLet-dig = Let-dig / UTF8-non-ascii

Let-dig = <Defined in [Section 4.1.2 of RFC 5321](#)>

uLdh-str = *(ALPHA / DIGIT / "-" / UTF8-non-ascii) uLet-dig
; Replace Ldh-str in [RFC 5321, Section 4.1.2](#)

UTF8-non-ascii = <Defined in [Section 4.1](#) of RFC5335bis>

3.4. MAIL Command Parameter Usage

If the envelope or message being sent requires the capabilities of the UTF8SMTPbis extension, the SMTP client **MUST** supply the UTF8SMTPbis parameter with the MAIL command. If this parameter is provided, it **MUST** have no value. If the SMTP client is aware that neither the envelope nor the message being sent requires any of the UTF8SMTPbis extension capabilities, it **SHOULD NOT** supply the UTF8SMTPbis parameter with the MAIL command.

Because there is no guarantee that a next-hop SMTP server will support the UTF8SMTPbis extension, use of the UTF8SMTPbis extension always carries a risk of transmission failure. In fact, during the early stages of deployment for the UTF8SMTPbis extension, the risk will be quite high. Hence there is a distinct near-term advantage for ASCII-only messages to be sent without using this extension. The long-term advantage of casting ASCII [[ASCII](#)] characters(0x7f and below) as UTF-8 form is that it permits pure-Unicode environments.

This specification does not require that the EAI-aware SMTP client inspect the message or otherwise go to extraordinary lengths to assure itself whether the UTF8SMTPbis extension is **REQUIRED** for the particular message.

3.5. Non-ASCII addresses and Reply-codes

An EAI-aware SMTP client **MUST** only send an internationalized message to an SMTP server that supports UTF8SMTPbis. If the SMTP server does not support this option, then the EAI-aware SMTP client has three choices according to [section 3.2](#) of this specification.

The three-digit Reply-codes used in this section are based on their meanings as defined in [RFC 5321](#).

When messages are rejected because the RCPT command requires an ASCII address, the reply-code 553 is returned with the meaning "mailbox name not allowed". When messages are rejected because the MAIL command requires an ASCII address, the reply-code 550 is returned with the meaning "mailbox unavailable". When the EAI-aware SMTP server supports enhanced mail system status codes [[RFC3463](#)], reply-code "X.6.7" [[RFC5248](#)] is used, meaning that "non-ASCII addresses not permitted for that sender/recipient".

When messages are rejected for other reasons, the server follows the model of the base email specifications in [RFC 5321](#); this extension does not change those circumstances or reply messages.

If the reply-code is issued after the final "." of the DATA command,

the reply-code "554" is used with the meaning "Transaction failed". When the EAI-aware SMTP server supports enhanced mail system status codes [[RFC3463](#)], reply-code "X.6.9" [[RFC5248](#)] is used, meaning that "UTF-8 header message can not be transmitted to one or more recipients, so the message MUST be rejected".

[3.6.](#) Body Parts and SMTP Extensions

The MAIL command parameter UTF8SMTPbis asserts that a message is an internationalized message or the message being sent needs the UTF8SMTPbis support. The message being sent via the MAIL command with the UTF8SMTPbis parameter has still a chance of that the message transmitted is not an internationalized message. An EAI-aware SMTP client or server that requires accurate knowledge of whether a message is internationalized needs to parse all message header fields and MIME header fields [[RFC2045](#)] and [[RFC2047](#)] in the message body. However, this specification does not require that the SMTP client or server inspects the message.

While this specification requires that EAI-aware SMTP servers support the 8BITMIME extension [[RFC6152](#)] to ensure that servers have adequate handling capability for 8-bit data and to avoid a number of complex encoding problems, the use of internationalized email addresses obviously does not require non-ASCII body parts in the MIME message in [RFC 2045](#) and [RFC 2047](#). The UTF8SMTPbis extension MAY be used with the BODY=8BITMIME parameter [[RFC6152](#)] if that is appropriate given the body content or, with the BODY=BINARYMIME parameter, if the SMTP server advertises BINARYMIME [[RFC3030](#)] and that is appropriate.

[3.7.](#) Additional ESMTP Changes and Clarifications

The information carried in the mail transport process involves addresses ("mailboxes") and domain names in various contexts in addition to the MAIL and RCPT commands and extended alternatives to them. In general, the rule is that, when [RFC 5321](#) specifies a mailbox, this SMTP extension requires UTF-8 form to be used for the entire string; when [RFC 5321](#) specifies a domain name, the name SHOULD be in the form of A-label if this domain name is an internationalized domain name[RFC5890].

The following subsections list and discuss all of the relevant cases.

[3.7.1.](#) The Initial SMTP Exchange

When an SMTP connection is opened, the SMTP server sends a "greeting" response consisting of the 220 reply-code and some information. The SMTP client then sends the EHLO command. Since the SMTP client cannot know whether the SMTP server supports UTF8SMTPbis until after

it receives the response from EHLO, the EAI-aware SMTP client MUST send only ASCII (LDH label or A-label [[RFC5890](#)]) domains in the EHLO command and that, if the EAI-aware SMTP server provides domain names in the EHLO response, they MUST be in the form of LDH labels or A-labels.

[3.7.2.](#) Mail eXchangers

If multiple DNS MX records are used to specify multiple servers for a domain in [section 5 of \[RFC5321\]](#), it is strongly advised that all or none of them SHOULD support the UTF8SMTPbis extension. Otherwise, surprising rejections can happen during temporary or permanent failures, which users might perceive as serious reliability issues. In order to avoid the possible surprising rejections, the EAI-aware email system may also implement the advice in EAI addresses advice document [EAI addresses] and EAI deployment advice document [EAI Deployment].

[3.7.3.](#) Trace Information

For the trace information [[RFC5321](#)], this memo updates the time stamp line and the return path line [[RFC5321](#)] formally defined as follows:

uReturn-path-line = "Return-Path:" FWS uReverse-path CRLF
; Replaces Return-path-line in [Section 4.4 of RFC 5321](#)

uReverse-path = uPath / "<>"
; Replace Reverse-path in [RFC 5321, section 4.1.2](#)

uPath = "<" [A-d-l ":"] uMailbox ">"
; Replace Path in [RFC 5321, section 4.1.2](#)
; uMailbox is defined in [section 3.3](#) of this document

A-d-l = <Defined in [section 4.1.2 of RFC 5321](#)>

uTime-stamp-line = "Received:" FWS uStamp CRLF
; Replaces Time-stamp-line in [Section 4.4 of RFC 5321](#)

uStamp = From-domain By-domain uOpt-info [CFWS] ";" FWS date-time
; Replaces Stamp in [Section 4.4 of RFC 5321](#)

From-domain = <Defined in [section 4.4 of RFC 5321](#)>

By-domain = <Defined in [section 4.4 of RFC 5321](#)>

date-time = <Defined in [section 3.3 of RFC 5322](#)>
; Same definition with date-time in [Section 4.4 of RFC 5321](#)

uOpt-info = [Via] [With] [ID] [uFor]
[Additional-Registered-Clauses]
; Replaces Opt-info in [Section 4.4 of RFC 5321](#)
; The protocol value for With will allow a UTF8SMTPbis value

Via = <Defined in [section 4.4 of RFC 5321](#)>

With = <Defined in [section 4.4 of RFC 5321](#)>

ID = <Defined in [section 4.4 of RFC 5321](#)>

Additional-Registered-Clauses = <Defined in [section 4.4 of RFC 5321](#)>

uFor = CFWS "FOR" FWS (uPath / uMailbox)
; Replaces For in [Section 4.4 of RFC 5321](#)
; uMailbox is defined in [section 3.3](#) of this document

Except in the 'uFor' clause and 'uReverse-path' value where internationalized domain name with the U-label form MAY be used, internationalized domain names in Received fields MUST be transmitted in the form of A-labels. The protocol value of the WITH clause when this extension is used is one of the UTF8SMTPbis values specified in the "IANA Considerations" section of this document.

3.7.4. UTF-8 Strings in Replies

3.7.4.1. MAIL Command

If an SMTP client follows this specification and sends any MAIL commands containing the UTF8SMTPbis parameter, the EAI-aware SMTP server is permitted to use UTF-8 characters in the email address associated with 251 and 551 reply-codes, and the SMTP client MUST be able to accept and process them. If a given MAIL command does not include the UTF8SMTPbis parameter, the EAI-aware SMTP server MUST NOT return a 251 or 551 response containing a non-ASCII mailbox. Instead, it MUST transform such responses into 250 or 550 responses that do not contain non-ASCII addresses.

3.7.4.2. VRFY and EXPN Commands and the UTF8SMTPbis Parameter

If the VRFY and EXPN commands are transmitted with the parameter "UTF8SMTPbis", it indicates the SMTP client can accept UTF-8 strings in replies to those commands. This parameter for the VRFY and EXPN commands SHOULD only be used after the SMTP client sees the EHLO response with the UTF8SMTPbis keyword. This allows the EAI-aware SMTP server to use UTF-8 strings in mailbox names and full names that occur in replies without concern that the SMTP client might be confused by them. An SMTP client that conforms to this specification MUST accept and correctly process replies from the VRFY and EXPN commands that contain UTF-8 strings. However, the EAI-aware SMTP server MUST NOT use UTF-8 strings in replies if the SMTP client does not specifically allow such replies by transmitting this parameter. Most replies do not require that a mailbox name be included in the returned text, and therefore UTF-8 string is not needed in them. Some replies, notably those resulting from successful execution of the VRFY and EXPN commands, do include the mailbox.

VERIFY (VRFY) and EXPAND (EXPN) command syntaxes are changed to:

```
vrfy = "VRFY" SP uString  
      [ SP "UTF8SMTPbis" ] CRLF
```

```
expn = "EXPN" SP uString  
      [ SP "UTF8SMTPbis" ] CRLF
```

```
uString = uAtom / uQuoted-string  
; uAtom and uQuoted-string are defined in  
; Section 3.3 of this document.
```

The "UTF8SMTPbis" parameter does not use a value. If the reply to a VERIFY (VRFY) or EXPAND (EXPN) command requires UTF-8 string, but the SMTP client did not use the "UTF8SMTPbis" parameter, then the EAI-

aware SMTP server MUST use either the reply-code 252 or 550. Reply-code 252, defined in [[RFC5321](#)], means "Cannot VRFY user, but will accept the message and attempt the delivery". Reply-code 550, also defined in [[RFC5321](#)], means "Requested action not taken: mailbox unavailable". When the EAI-aware SMTP server supports enhanced mail system status codes [[RFC3463](#)], the enhanced reply-code as specified below is used. Using the "UTF8SMTPbis" parameter with a VERIFY (VRFY) or EXPAND (EXPN) command enables UTF-8 replies for that command only.

If a normal success response (i.e., 250) is returned, the response MAY include the full name of the user and MUST include the mailbox of the user. It MUST be in either of the following forms:

User Name <uMailbox>

- ; uMailbox is defined in [Section 3.3](#) of this document.
- ; User Name can contain non-ASCII characters.

uMailbox

- ; uMailbox is defined in [Section 3.3](#) of this document.

If the SMTP reply requires UTF-8 strings, but UTF-8 string is not allowed in the reply, and the EAI-aware SMTP server supports enhanced mail system status codes [[RFC3463](#)], the enhanced reply-code is "X.6.8" [[RFC5248](#)], meaning "A reply containing a UTF-8 string is REQUIRED to show the mailbox name, but that form of response is not permitted by the SMTP client".

If the SMTP client does not support the UTF8SMTPbis extension, but receives a UTF-8 string in a reply, it may not be able to properly report the reply to the user, and some clients might crash. Internationalized messages in replies are only allowed in the commands under the situations described above. Under any other circumstances, UTF-8 string MUST NOT appear in the reply.

Although UTF-8 form is needed to represent email addresses in responses under the rules specified in this section, this extension does not permit the use of UTF-8 string for any other purposes. EAI-aware SMTP servers MUST NOT include non-ASCII characters in replies except in the limited cases specifically permitted in this section.

[4.](#) IANA Considerations

IANA is requested to add a new value "UTF8SMTPbis" to the SMTP Service Extension subregistry of the Mail Parameters registry, according to the following data:

Keywords	Description	Reference
UTF8SMTPbis	Internationalized email address	[RFCXXXX]

This document updates the values to the SMTP Enhanced Status Code subregistry of the Mail Parameters registry, following the guidance in Sections [3.5](#) and [3.7.4.2](#) of this document, and being based on [\[RFC5248\]](#). The registration data is as follows:

Code: X.6.7
 Sample Text: non-ASCII addresses not permitted
 for that sender/recipient
 Associated basic status code: 550, 553
 Description: This indicates the reception of a MAIL or RCPT
 command that non-ASCII addresses are not permitted
 Defined: RFC XXXX (Standard track)
 Submitter: Jiankang YAO
 Change controller: ima@ietf.org

Code: X.6.8
 Sample Text: UTF-8 string reply is required,
 but not permitted by the SMTP client
 Associated basic status code: 252, 550, 553
 Description: This indicates that a reply containing a UTF-8
 string is required to show the mailbox name,
 but that form of response is not
 permitted by the SMTP client.
 Defined: RFC XXXX (Standard track)
 Submitter: Jiankang YAO
 Change controller: ima@ietf.org

Code: X.6.9
 Sample Text: UTF-8 header message can not be transferred
 to one or more recipient so the message
 must be rejected
 Associated basic status code: 550
 Description: This indicates that transaction failed
 after the final "." of the DATA command.
 Defined: RFC XXXX (Standard track)
 Submitter: Jiankang YAO
 Change controller: ima@ietf.org

Code: X.6.10
 Description: This is a duplicate of X.6.8 and
 is thus deprecated.

The following entries SHOULD be updated or added in the "Mail Transmission Types" registry under the Mail Parameters registry.

WITH protocol types	Description	Reference
UTF8SMTP	ESMTP with UTF8SMTP	[RFCXXXX]
UTF8SMTPA	ESMTP with UTF8SMTP and SMTP AUTH	[RFC4954] [RFCXXXX]
UTF8SMTPS	ESMTP with UTF8SMTP and STARTTLS	[RFC3207] [RFCXXXX]
UTF8SMTPSA	ESMTP with UTF8SMTP and both STARTTLS and SMTP AUTH	[RFC3207] [RFC4954] [RFCXXXX]
UTF8LMTP	LMTP with UTF8SMTP	[RFCXXXX]
UTF8LMTPA	LMTP with UTF8SMTP and SMTP AUTH	[RFC4954] [RFCXXXX]
UTF8LMTPS	LMTP with UTF8SMTP and STARTTLS	[RFC3207] [RFCXXXX]
UTF8LMTPSA	LMTP with UTF8SMTP and both STARTTLS and LMTP AUTH	[RFC3207] [RFC4954] [RFCXXXX]

5. Security Considerations

The extended security considerations discussion in the framework document [RFC4952bis] will be applied here.

More security considerations are discussed below:

Beyond the use inside the email global system (in SMTP envelopes and message headers), internationalized email addresses will also show up inside other cases, in particular:

- o the logging systems of SMTP transactions and other logs to monitor the email systems;
- o the trouble ticket systems used by Security Teams to manage security incidents, when an email address is involved;

In order to avoid problems that could cause loss of data, this will

likely require extending these systems to support full UTF-8, or to require to provide an adequate mechanisms for mapping non-ASCII strings to ASCII.

Another security aspect to be considered is related to the ability by security team members to quickly understand, read and identify email addresses from the logs, when they are tracking an incident. Mechanisms to automatically and quickly provide the origin or ownership of an internationalized email address SHALL be implemented for use also by log readers which cannot read easily non-ASCII information.

The SMTP commands VRFY and EXPN are sometimes used in SMTP transactions where there is no message to transfer (by tools used to take automated actions in case potential spam messages are identified). [RFC 5321 section 3.5](#) and 7.3 give some detailed description of use and possible behaviours. Implementation of internationalized addresses can affect also logs and actions by these tools.

6. Acknowledgements

This document revised the [[RFC5336](#)] document based on the EAI WG's discussion result. Many EAI WG members did some tests and implementations to move this document to the Standard Track document. Significant comments and suggestions were received from Xiaodong LEE, Nai-Wen Hsu, Yangwoo KO, Yoshiro YONEYA, and other members of the JET team and were incorporated into the specification. Additional important comments and suggestions, and often specific text, were contributed by many members of the WG and design team. Those contributions include material from John C Klensin, Charles Lindsey, Dave Crocker, Harald Tveit Alvestrand, Marcos Sanz, Chris Newman, Martin Duerst, Edmon Chung, Tony Finch, Kari Hurtta, Randall Gellens, Frank Ellermann, Alexey Melnikov, Pete Resnick, S. Moonesamy, Soobok Lee, Shawn Steele, Alfred Hoenes, Miguel Garcia, Magnus Westerlund, and Lars Eggert. Of course, none of the individuals are necessarily responsible for the combination of ideas represented here.

7. Change History

[[anchor14: RFC Editor: Please remove this section.]]

7.1. [draft-yao-eai-rfc5336bis](#): Version 00

Applied errata suggested by Alfred Hoenes.

7.2. [draft-ietf-eai-rfc5336bis](#): Version 00

Applied the changes suggested by the EAI new charter.

7.3. [draft-ietf-eai-rfc5336bis](#): Version 01

Applied the changes suggested by 78 IETF EAI meeting.

7.4. [draft-ietf-eai-rfc5336bis](#): Version 02

remove the appendix since rfc4952bis has added this material

improve the text

remove the text about no body parameter

7.5. [draft-ietf-eai-rfc5336bis](#): Version 03

improve the text

7.6. [draft-ietf-eai-rfc5336bis](#): Version 04

update the abstract

improve the text

7.7. [draft-ietf-eai-rfc5336bis](#): Version 05

improve the text based on AD and Co-chairs

7.8. [draft-ietf-eai-rfc5336bis](#): Version 06

update the iana consideration

7.9. [draft-ietf-eai-rfc5336bis](#): Version 07

improve the iana consideration

7.10. [draft-ietf-eai-rfc5336bis](#): Version 08

improve the texts

add the mail parameter

add the new section about mail command parameter usage

update the security consideration

7.11. draft-ietf-eai-rfc5336bis: Version 09

improve the texts

8. References

8.1. Normative References

- [ASCII] American National Standards Institute (formerly United States of America Standards Institute), "USA Code for Information Interchange", ANSI X3.4-1968, 1968.
- [RFC2033] Myers, J., "Local Mail Transfer Protocol", [RFC 2033](#), October 1996.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC3463] Vaudreuil, G., "Enhanced Mail System Status Codes", [RFC 3463](#), January 2003.
- [RFC3464] Moore, K. and G. Vaudreuil, "An Extensible Message Format for Delivery Status Notifications", [RFC 3464](#), January 2003.
- [RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", [RFC 3629](#), November 2003.
- [RFC4409] Gellens, R. and J. Klensin, "Message Submission for Mail", [RFC 4409](#), April 2006.
- [RFC4952bis]
Klensin, J. and Y. Ko, "Overview and Framework for Internationalized Email", I-D rfc4952bis, September 2010.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, [RFC 5234](#), January 2008.
- [RFC5248] Hansen, T. and J. Klensin, "A Registry for SMTP Enhanced Mail System Status Codes", [RFC 5248](#), June 2008.
- [RFC5321] Klensin, J., "Simple Mail Transfer Protocol", [RFC 5321](#), October 2008.
- [RFC5322] Resnick, P., Ed., "Internet Message Format", [RFC 5322](#), October 2008.

[RFC5335bis]

Abel, Y. and S. Steel, "Internationalized Email Headers", I-D rfc5335bis, March 2011.

[RFC5337bis]

Hansen, T., Ed., Newman, C., and A. Melnikov, Ed., "Internationalized Delivery Status and Disposition Notifications", I-D 5337bis, October 2010.

[RFC5890] Klensin, J., "Internationalizing Domain Names in Applications (IDNA definitions)", [RFC 5890](#), June 2010.

[RFC5891] Klensin, J., "Internationalized Domain Names in Applications (IDNA): Protocol", [RFC 5891](#), August 2010.

[RFC6152] Klensin, J., Freed, N., Rose, M., and D. Crocker, "SMTP Service Extension for 8-bit MIME Transport", STD 71, [RFC 6152](#), March 2011.

[8.2.](#) Informative References

[EAI Deployment]

Yao, J., Lee, X., and S. Steel, "Advice for EAI deployment", draft eai-deployment, December 2010.

[EAI addresses]

Steel, S., Yao, J., and Mark. Davis, "Advice for non-ASCII & ASCII addresses", draft eai-address-advice, December 2010.

[RFC2045] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies", [RFC 2045](#), November 1996.

[RFC2047] Moore, K., "MIME (Multipurpose Internet Mail Extensions) Part Three: Message Header Extensions for Non-ASCII Text", [RFC 2047](#), November 1996.

[RFC3030] Vaudreuil, G., "SMTP Service Extensions for Transmission of Large and Binary MIME Messages", [RFC 3030](#), December 2000.

[RFC3207] Hoffman, P., "SMTP Service Extension for Secure SMTP over Transport Layer Security", [RFC 3207](#), February 2002.

[RFC4954] Siemborski, R. and A. Melnikov, "SMTP Service Extension for Authentication", [RFC 4954](#), July 2007.

[RFC5336] Yao, J. and W. Mao, "SMTP Extension for Internationalized Email Addresses", [RFC 5336](#), September 2008.

[RFC5598] Crocker, D., "Internet Mail Architecture", [RFC 5598](#), July 2009.

Authors' Addresses

Jiankang YAO
CNNIC
No.4 South 4th Street, Zhongguancun
Beijing

Phone: +86 10 58813007
Email: yaojk@cnnic.cn

Wei MAO
CNNIC
No.4 South 4th Street, Zhongguancun
Beijing

Phone: +86 10 58812230
Email: maowei_ietf@cnnic.cn

