

Network Working Group  
Internet-Draft  
Obsoletes: [5336](#) (if approved)  
Intended status: Standards Track  
Expires: May 13, 2012

J. Yao  
W. Mao  
CNNIC  
November 10, 2011

**SMTP Extension for Internationalized Email  
draft-ietf-eai-rfc5336bis-16.txt**

Abstract

This document specifies an SMTP extension for transport and delivery of email messages with internationalized email addresses or header information. This specification replaces [RFC 5336](#).

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 May 13, 2012.

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

<a href="#">1.</a>	<a href="#">Introduction</a>	<a href="#">4</a>
<a href="#">1.1.</a>	<a href="#">Terminology</a>	<a href="#">4</a>
<a href="#">1.2.</a>	<a href="#">Changes Made to Other Specifications</a>	<a href="#">5</a>
<a href="#">2.</a>	<a href="#">Overview of Operation</a>	<a href="#">5</a>
<a href="#">3.</a>	<a href="#">Mail Transport-Level Protocol</a>	<a href="#">5</a>
<a href="#">3.1.</a>	<a href="#">Framework for the Internationalization Extension</a>	<a href="#">5</a>
<a href="#">3.2.</a>	<a href="#">The UTF8SMTPbis Extension</a>	<a href="#">6</a>
<a href="#">3.3.</a>	<a href="#">Extended Mailbox Address Syntax</a>	<a href="#">7</a>
<a href="#">3.4.</a>	<a href="#">MAIL Command Parameter Usage</a>	<a href="#">9</a>
<a href="#">3.5.</a>	<a href="#">Non-ASCII addresses and Reply-codes</a>	<a href="#">9</a>
<a href="#">3.6.</a>	<a href="#">Body Parts and SMTP Extensions</a>	<a href="#">10</a>
<a href="#">3.7.</a>	<a href="#">Additional ESMTP Changes and Clarifications</a>	<a href="#">11</a>
<a href="#">3.7.1.</a>	<a href="#">The Initial SMTP Exchange</a>	<a href="#">11</a>
<a href="#">3.7.2.</a>	<a href="#">Mail eXchangers</a>	<a href="#">11</a>
<a href="#">3.7.3.</a>	<a href="#">Trace Information</a>	<a href="#">11</a>
<a href="#">3.7.4.</a>	<a href="#">UTF-8 Strings in Replies</a>	<a href="#">12</a>
<a href="#">4.</a>	<a href="#">IANA Considerations</a>	<a href="#">14</a>
<a href="#">4.1.</a>	<a href="#">SMTP Service Extensions Registry</a>	<a href="#">14</a>
<a href="#">4.2.</a>	<a href="#">SMTP Enhanced Status Code Registry</a>	<a href="#">14</a>
<a href="#">4.3.</a>	<a href="#">WITH protocol types sub-registry of the Mail Transmission Types Registry</a>	<a href="#">15</a>
<a href="#">5.</a>	<a href="#">Security Considerations</a>	<a href="#">16</a>
<a href="#">6.</a>	<a href="#">Acknowledgements</a>	<a href="#">17</a>
<a href="#">7.</a>	<a href="#">Change History</a>	<a href="#">17</a>
<a href="#">7.1.</a>	<a href="#">draft-yao-eai-rfc5336bis: Version 00</a>	<a href="#">17</a>
<a href="#">7.2.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 00</a>	<a href="#">17</a>
<a href="#">7.3.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 01</a>	<a href="#">17</a>
<a href="#">7.4.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 02</a>	<a href="#">17</a>
<a href="#">7.5.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 03</a>	<a href="#">18</a>
<a href="#">7.6.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 04</a>	<a href="#">18</a>
<a href="#">7.7.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 05</a>	<a href="#">18</a>
<a href="#">7.8.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 06</a>	<a href="#">18</a>
<a href="#">7.9.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 07</a>	<a href="#">18</a>
<a href="#">7.10.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 08</a>	<a href="#">18</a>
<a href="#">7.11.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 09</a>	<a href="#">18</a>
<a href="#">7.12.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 10</a>	<a href="#">18</a>
<a href="#">7.13.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 11</a>	<a href="#">18</a>
<a href="#">7.14.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 12</a>	<a href="#">19</a>
<a href="#">7.15.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 13</a>	<a href="#">19</a>
<a href="#">7.16.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 14</a>	<a href="#">19</a>
<a href="#">7.17.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 15</a>	<a href="#">19</a>
<a href="#">7.18.</a>	<a href="#">draft-ietf-eai-rfc5336bis: Version 16</a>	<a href="#">19</a>
<a href="#">8.</a>	<a href="#">References</a>	<a href="#">19</a>
<a href="#">8.1.</a>	<a href="#">Normative References</a>	<a href="#">19</a>
<a href="#">8.2.</a>	<a href="#">Informative References</a>	<a href="#">21</a>
	<a href="#">Authors' Addresses</a>	<a href="#">21</a>



## **1. Introduction**

The document defines a Simple Mail Transfer Protocol [[RFC5321](#)] extension so servers can advertise the ability to accept and process internationalized email addresses (see [section 1.1](#)), and internationalized email headers [[RFC5335bis](#)].

An extended overview of the extension model for internationalized email addresses and the email header appears in [[RFC4952bis](#)], referred to as "the framework document" in this specification. 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.

[[anchor1: Note in Draft and to RFC Editor: The keyword represented in this document by "UTF8SMTPbis" (and in the XML source by "UTF8SMTPbis") is a placeholder. The actual keyword will need to be assigned after document approval by a process to be worked out between the responsible AD, WG co-chairs, and IANA. The assigned keyword should be substituted here. This paragraph should be removed before RFC publication.]]

### **1.1. 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, which may or may not be members of the ASCII subset, 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](#)]. Some basic rules in this document are identified in [Section 3.3](#) as being defined (under the same names) in [[RFC5234](#)], [[RFC5321](#)], [[RFC5890](#)] or [[RFC5335bis](#)].



## **1.2. Changes Made to Other Specifications**

This specification extends some syntax rules defined in [RFC 5321](#) and permits internationalized email addresses in the envelope, but it does not modify [RFC 5321](#). It permits data formats defined in [\[RFC5335bis\]](#), but it does not modify [RFC 5322](#). It does require that the 8BITMIME extension [\[RFC6152\]](#) be announced by the UTF8SMTPbis-aware SMTP server and used with "BODY=8BITMIME" by the UTF8SMTPbis-aware SMTP client, but it does not modify the 8BITMIME specification in any way.

This specification replaces an earlier, experimental, approach to the same problem [\[RFC5336\]](#). Section 6 of [\[RFC4952bis\]](#) describes the changes in approach between [\[RFC5336\]](#) and this specification. Anyone trying to convert an implementation from the experimental specification to the specification in this document will need to review those changes carefully.

## **2. Overview of Operation**

This document specifies an element of the email internationalization work, specifically the definition of an SMTP extension for internationalized email. The extension is identified with the token "UTF8SMTPbis".

The internationalized email headers 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 "Internationalized Email".
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 UTF8SMTPbis-aware SMTP client MUST ignore any parameters if they appear for this keyword; that is, the UTF8SMTPbis-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 UTF8SMTPbis-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 13 characters by the possible addition of the UTF8SMTPbis parameter. [[anchor5: RFC Editor: the number '13' will be replaced by the new number (2 spaces + 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](#)]. It may also be used with LMTP [[RFC2033](#)]. When these protocols are used, their use should be reflected in trace field WITH keywords as appropriate [[RFC3848](#)].

### **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 name to be looked up in the DNS MUST conform to and be processed as specified for IDNA [[RFC5890](#)]. When doing lookups, the UTF8SMTPbis-aware SMTP client or server MUST either use a Unicode aware DNS library, or transform the internationalized domain name to the form of A-label as described 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 UTF8SMTPbis-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]. (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 a UTF8SMTPbis-aware SMTP client (UTF8SMTPbis-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 internationalized delivery status and disposition notifications specification [RFC5337bis].
2. If and only if the UTF8SMTPbis-aware SMTP client (sender) is a Message Submission Agent ("MSA") [RFC4409] [RFC5598], it MAY choose its own way to deal with this scenario according to the provisions of [RFC4409] or its future versions. But the detailed specification of this process and its results are outside the scope of this document.
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 UTF8SMTPbis-aware SMTP client.

This document applies when a UTF8SMTPbis-aware SMTP client or server supports the UTF8SMTPbis extension. For all other cases, and for addresses and messages that do not require a UTF8SMTPbis extension, UTF8SMTPbis-aware SMTP clients and servers do not change the behavior specified in [RFC5321].

If a UTF8SMTPbis-aware SMTP server advertises the Delivery Status Notification (DSN) [RFC3461] extension, it MUST implement [RFC5337bis].

### **3.3. Extended Mailbox Address Syntax**

RFC 5321, section 4.1.2, defines the syntax of a <Mailbox> entirely in terms of ASCII characters. This document extends <Mailbox> to add



support of non-ASCII characters.

The key changes made by this specification include:

- o In order to update the <Mailbox> to support the internationalized email address, the <Mailbox> ABNF rule will be imported from [RFC 5321](#) directly, and other related rules are imported from [RFC 5321](#), [RFC 5234](#), [RFC 5890](#) or RFC 5335bis, or are extended in this document.
- o Extend the definition of <sub-domain> to permit both the [RFC 5321](#) definition and a UTF-8 string in a DNS label that is conforming with IDNA definitions [[RFC5890](#)].
- o Extend the definition of <atext> to permit both the [RFC 5321](#) definition and a UTF-8 string. That string MUST NOT contain any of the ASCII graphics or controls characters.

The following ABNF rules imported from [RFC 5321, section 4.1.2](#) are updated directly or indirectly by this document:

- o <Mailbox>
- o <Local-part>
- o <Dot-string>
- o <Quoted-string>
- o <QcontentSMTP>
- o <Domain>
- o <Atom>

The following ABNF rule will be imported from RFC 5335bis, [section 3.1](#) directly:

- o <UTF8-non-ascii>

The following ABNF rule will be imported from [RFC 5234, appendix B.1](#) directly:

- o <DQUOTE>

The following ABNF rule will be imported from [RFC 5890, section 2.3.2.1](#) directly:

- o <U-label>

The following rules are extended in ABNF [[RFC5234](#)] as follows.



```
sub-domain    =/  U-label
; extend the definition of sub-domain in RFC5321, section 4.1.2

atext         =/  UTF8-non-ascii
; extend the implicit definition of atext in
; RFC5321, Section 4.1.2, which ultimately points to
; the actual definition in RFC5322, Section 3.2.3

qtextSMTP     =/  UTF8-non-ascii
; extend the definition of qtextSMTP in RFC5321, section 4.1.2

esmtplib-value =/  UTF8-non-ascii
; extend the definition of esmtplib-value in RFC5321, section 4.1.2
```

### **[3.4.](#) MAIL Command Parameter Usage**

If the envelope or message being sent requires the capabilities of the UTF8SMTPbis extension, the UTF8SMTPbis-aware SMTP client MUST supply the UTF8SMTPbis parameter with the MAIL command. If this parameter is provided, it MUST have no value. If the UTF8SMTPbis-aware 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.

### **[3.5.](#) Non-ASCII addresses and Reply-codes**

A UTF8SMTPbis-aware SMTP client MUST NOT send an internationalized message to an SMTP server that does not support UTF8SMTPbis. If the SMTP server does not support this option, then the UTF8SMTPbis-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 UTF8SMTPbis-aware SMTP server supports enhanced mail system status codes [[RFC3463](#)], reply-code "X.6.7" [[RFC5248](#)] (see [section 4](#)) 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 a message is rejected after the final "." of the DATA command because one or more recipient is unable to accept and process a message with internationalized email headers, the reply-code "554" is used with the meaning "Transaction failed". If the UTF8SMTPbis-aware SMTP server supports enhanced mail system status codes [[RFC3463](#)], reply code "X.6.9" [[RFC5248](#)] (see [section 4](#)) is used to indicate this condition, meaning that "UTF-8 header message can not be transmitted to one or more recipients, so the message MUST be rejected".

The UTF8SMTPbis-aware SMTP servers are encouraged to detect that recipients can not accept internationalized messages and generate an error after the RCPT command rather than waiting until after the DATA command to issue an error.

### **[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. There is still a chance that a message being sent via the MAIL command with the UTF8SMTPbis parameter is not an internationalized message. A UTF8SMTPbis-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](#)] in the message body. However, this specification does not require that the UTF8SMTPbis-aware SMTP client or server inspects the message.

Although this specification requires that UTF8SMTPbis-aware SMTP servers support the 8BITMIME extension [[RFC6152](#)] to ensure that servers have adequate handling capability for 8-bit data, it does not require non-ASCII body parts in the MIME message in [RFC 2045](#). 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 internationalized domain name SHOULD be in the form of U-label if the UTF8SMTPbis extension is supported; otherwise, it SHOULD be in the form of A-label.

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 UTF8SMTPbis-aware SMTP client MUST send only ASCII (LDH label or A-label [[RFC5890](#)] ) domains in the EHLO command and that, if the UTF8SMTPbis-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, unexpected rejections can happen during temporary or permanent failures, which users might perceive as serious reliability issues.

#### **3.7.3. Trace Information**

The trace information <Return-path-line>, <Time-stamp-line> and their related rules are defined in in [section 4.4 of RFC 5321](#) [[RFC5321](#)]. This document updates <Mailbox> and <Domain> to support non-ASCII characters. When the UTF8SMTPbis extension is used, the 'Reverse-path' clause of the Return-path-line may include an internationalized domain name that uses the U-label form; The 'Stamp' clause of the Time-stamp-line may include an internationalized domain name that uses the U-label form.

If the messages that include trace fields are sent by an UTF8SMTPbis-aware SMTP client or relay server without the UTF8SMTPbis parameter at MAIL commands, trace field values must conform to [RFC 5321](#)



regardless of the SMTP server's capability.

When a UTF8SMTPbis-aware SMTP server adds a trace field to a message that was or will be transmitted with the UTF8SMTPbis parameter at MAIL commands, that server SHOULD use the U-label form for internationalized domain names in that new trace field.

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 UTF8SMTPbis-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 UTF8SMTPbis-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 UTF8SMTPbis-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 UTF8SMTPbis-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 String
      [ SP "UTF8SMTPbis" ] CRLF
      ; String may include Non-ASCII characters

expn = "EXPN" SP String
      [ SP "UTF8SMTPbis" ] CRLF
      ; String may include Non-ASCII characters
```

The "UTF8SMTPbis" parameter does not have a value. If the reply to a VERIFY (VRFY) or EXPAND (EXPN) command requires a UTF-8 string, but the SMTP client did not use the "UTF8SMTPbis" parameter, then the UTF8SMTPbis-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 UTF8SMTPbis-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 <Mailbox>
      ; Mailbox is defined in section 3.3 of this document.
      ; User Name can contain non-ASCII characters.
```

```
Mailbox
      ; Mailbox 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 UTF8SMTPbis-aware SMTP server supports enhanced mail system status codes [\[RFC3463\]](#), the enhanced reply-code is "X.6.8" [\[RFC5248\]](#) (see [section 4](#)), 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 mishandle that reply. Internationalized messages in replies are only allowed in the commands under the situations described above.

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. UTF8SMTPbis-aware SMTP servers MUST NOT include non-ASCII characters in replies except in the limited cases specifically permitted in this section.

## **4. IANA Considerations**

### **4.1. SMTP Service Extensions Registry**

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

+-----+	+-----+	+-----+
Keywords	Description	Reference
+-----+	+-----+	+-----+
UTF8SMTPbis	Internationalized email address	[RFCXXXX]
+-----+	+-----+	+-----+

### **4.2. SMTP Enhanced Status Code Registry**

The new code definitions in this document replace those that now appear in 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.

#### **4.3. WITH protocol types sub-registry of the Mail Transmission Types Registry**

IANA is requested to update or add the following entries in the "Mail Transmission Types" registry under the Mail Parameters registry.

WITH protocol types	Description	Reference
UTF8SMTP	ESMTP with UTF8SMTPbis	[RFCXXXX]
UTF8SMTPA	ESMTP with UTF8SMTPbis and SMTP AUTH	[RFC4954]
UTF8SMTPS	ESMTP with UTF8SMTPbis and STARTTLS	[RFC3207]
UTF8SMTPSA	ESMTP with UTF8SMTPbis and both STARTTLS and SMTP AUTH	[RFC3207]
		[RFCXXXX]



UTF8LMTP	LMTP with UTF8SMTPbis	[RFCXXXX]	
UTF8LMTPA	LMTP with UTF8SMTPbis and SMTP	[ <a href="#">RFC4954</a> ]	
	AUTH	[RFCXXXX]	
UTF8LMTPS	LMTP with UTF8SMTPbis and	[ <a href="#">RFC3207</a> ]	
	STARTTLS	[RFCXXXX]	
UTF8LMTPSA	LMTP with UTF8SMTPbis and both	[ <a href="#">RFC3207</a> ]	
	STARTTLS and LMTP AUTH	[ <a href="#">RFC4954</a> ]	
		[RFCXXXX]	
+-----+-----+-----+-----+			

## 5. Security Considerations

The extended security considerations discussion in the framework document [[RFC4952bis](#)] will apply 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 Email Address Internationalization (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 Hurtt, Randall Gellens, Frank Ellermann, Alexey Melnikov, Pete Resnick, S. Moonesamy, Soobok Lee, Shawn Steele, Alfred Hoenes, Miguel Garcia, Magnus Westerlund, Joseph Yee and Lars Eggert. Of course, none of the individuals are necessarily responsible for the combination of ideas represented here.

Thanks a lot to Dave Crocker for his comments and helping of ABNF refinement.

## **7. Change History**

[[anchor15: 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

**7.12. [draft-ietf-eai-rfc5336bis](#): Version 10**

refine the ABNF definitions

improve the texts

**7.13. [draft-ietf-eai-rfc5336bis](#): Version 11**

remove the update of [RFC5321](#) and [RFC5322](#)

change the title from "SMTP Extension for Internationalized Email Address" to "SMTP Extension for Internationalized Email" based on



Ernie's comment

the trace field of [section 3.7.3](#) is updated to reflect the WG's conclusion

improve the texts

#### **7.14. [draft-ietf-eai-rfc5336bis](#): Version 12**

Update according to Chris Newman's comments

improve the texts

#### **7.15. [draft-ietf-eai-rfc5336bis](#): Version 13**

Update the esmpt-value syntax according to Chris Newman's comments

improve the texts

#### **7.16. [draft-ietf-eai-rfc5336bis](#): Version 14**

improve the texts

#### **7.17. [draft-ietf-eai-rfc5336bis](#): Version 15**

improve the texts

updates based on IESG members' comments

#### **7.18. [draft-ietf-eai-rfc5336bis](#): Version 16**

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.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC3461] Moore, K., "Simple Mail Transfer Protocol (SMTP) Service Extension for Delivery Status Notifications (DSNs)", [RFC 3461](#), January 2003.



- [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.
- [RFC3848] Newman, C., "ESMTP and LMTP Transmission Types Registration", [RFC 3848](#), July 2004.
- [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 2011.
- [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.
- [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**

- [RFC2033] Myers, J., "Local Mail Transfer Protocol", [RFC 2033](#), October 1996.
- [RFC2045] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies", [RFC 2045](#), 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



