Internet Engineering Task Force (IETF)
Request for Comments: 6531
Obsoletes: 5336
Category: Standards Track
ISSN: 2070-1721

J. Yao
W. Mao
CNNIC

February 2012

SMTP Extension for Internationalized Email

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 is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6531.

Table of Contents

1. Introduction 3
1.1. Terminology 3
1.2. Changes Made to Other Specifications 3
2. Overview of Operation 4
3. Mail Transport-Level Protocol 4
3.1. Framework for the Internationalization Extension 4
3.2. The SMTPUTF8 Extension 5
3.3. Extended Mailbox Address Syntax 7
3.4. MAIL Command Parameter Usage 8
3.5. Non-ASCII Addresses and Reply-Codes 9
3.6. Body Parts and SMTP Extensions 9
3.7. Additional ESMTP Changes and Clarifications 10
3.7.1. The Initial SMTP Exchange 10
3.7.2. Mail eXchangers 10
3.7.3. Trace Information 11
3.7.4. UTF-8 Strings in Replies 11
4. IANA Considerations 13
4.1. SMTP Service Extensions Registry 13
4.2. SMTP Enhanced Status Code Registry 13
4.3. WITH Protocol Types Sub-Registry of the Mail Transmission Types Registry 15
5. Security Considerations 15
6. Acknowledgements 16
7. References 16
7.1. Normative References 16
7.2. Informative References 18

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 [RFC6532].

An extended overview of the extension model for internationalized email addresses and the email header appears in RFC 6530 [RFC6530], 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.

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, in UTF-8 [RFC3629], a standard Unicode Encoding Form. 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", "SMTPUTF8", "internationalized message", and "message" are used in this document according to the definitions in the framework document [RFC6530].

Strings referred to in this document, including ASCII strings, MUST be expressed in UTF-8.

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 RFC 5234 [RFC5234], RFC 5321 [RFC5321], RFC 5890 [RFC5890], or RFC 6532 [RFC6532].

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 and in trace fields, but it does not modify RFC 5321. It permits data formats defined in RFC 6532 [RFC6532], but it does not modify RFC 5322. It does require that the 8BITMIME extension [RFC6152] be announced by the SMTPUTF8-aware SMTP server and used with "BODY=8BITMIME" by the SMTPUTF8-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 RFC 6530 [RFC6530] describes the changes in approach between RFC 5336 [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 "SMTPUTF8".

The internationalized email headers specification [RFC6532] 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 "SMTPUTF8".

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 SMTPUTF8-aware SMTP client MUST ignore any parameters if they appear for this keyword; that is, the SMTPUTF8-aware SMTP client MUST behave as if the parameters do not appear. If an SMTP server includes SMTPUTF8 in its EHLO response, it MUST be fully compliant with this version of this specification.

4. One OPTIONAL parameter, SMTPUTF8, is added to the MAIL command. The parameter does not accept a value. If this parameter is set in the MAIL command, it indicates that the SMTP client is SMTPUTF8-aware. Its presence also asserts that the envelope includes the non-ASCII address, the message being sent is an internationalized message, or the message being sent needs the SMTPUTF8 support.

5. The maximum length of a MAIL command line is increased by 10 characters to accommodate the possible addition of the SMTPUTF8 parameter.

6. One OPTIONAL parameter, SMTPUTF8, is added to the VERIFY (VRFY) and EXPAND (EXPN) commands. The SMTPUTF8 parameter does not accept a 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 RFC 6532 [RFC6532].

11. The SMTPUTF8 extension is valid on the submission port [RFC6409]. It may also be used with the Local Mail Transfer Protocol (LMTP) [RFC2033]. When these protocols are used, their use should be reflected in the trace field WITH keywords as appropriate [RFC3848].

3.2. The SMTPUTF8 Extension

An SMTP server that announces the SMTPUTF8 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, the 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 Internationalizing Domain Names in Applications (IDNA) [RFC5890]. When doing lookups, the SMTPUTF8-aware SMTP client or server MUST either use a Unicode-aware DNS library, or transform the internationalized domain name to A-label form (i.e., a fully-qualified domain name that contains one or more A-labels but no U-labels) as specified in RFC 5890 [RFC5890].

An SMTP client that receives the SMTPUTF8 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 [RFC6532] (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 SMTPUTF8 extension does not change the server-relaying behaviors described in RFC 5321.

If the SMTPUTF8 SMTP extension is not offered by the SMTP server, the SMTPUTF8-aware SMTP client MUST NOT transmit an internationalized email address and MUST NOT transmit a mail message containing internationalized mail headers as described in RFC 6532 [RFC6532] at any level within its MIME structure [RFC2045]. (For this paragraph, the internationalized domain name in A-label form as specified in IDNA definitions [RFC5890] is not considered to be "internationalized".) Instead, if an SMTPUTF8-aware SMTP client (sender) attempts to transfer an internationalized message and encounters an SMTP server that does not support the extension, the best action for it to take depends on other conditions. In particular:

- If it is a Message Submission Agent (MSA) [RFC6409] [RFC5598], it MAY choose its own way to deal with this scenario using the wide discretion for changing addresses or otherwise fixing up and transforming messages allowed by RFC 6409. As long as the resulting message conforms to the requirements of RFC 5321 (i.e., without the SMTPUTF8 extension), the details of that transformation are outside the scope of this document.

- If it is not an MSA or is an MSA and does not choose to transform the message to one that does not require the SMTPUTF8 extension, it SHOULD reject the message. As usual, this can be done either by generating an appropriate reply during the SMTP transaction or by accepting the message and then generating and transmitting a non-delivery notification. If the latter choice is made, the notification process MUST conform to the requirements of RFC 5321, RFC 3464 [RFC3464], and RFC 6533 [RFC6533].

- As specified in Section 2.2.3 of RFC 5321, an SMTP client with additional information and/or knowledge of special circumstances MAY choose to requeue the message and try later and/or try an alternate MX host as specified in that section.

This document applies when an SMTPUTF8-aware SMTP client or server supports the SMTPUTF8 extension. For all other cases, and for addresses and messages that do not require an SMTPUTF8 extension, SMTPUTF8-aware SMTP clients and servers do not change the behavior specified in RFC 5321 [RFC5321].

If an SMTPUTF8-aware SMTP server advertises the Delivery Status Notification (DSN) [RFC3461] extension, it MUST implement RFC 6533 [RFC6533].

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:

- The <Mailbox> ABNF rule is imported from RFC 5321 and updated in order to support the internationalized email address. Other related rules are imported from RFC 5321, RFC 5234, RFC 5890, and RFC 6532, or are extended in this document.

- The definition of <sub-domain> is extended to permit both the RFC 5321 definition and a UTF-8 string in a DNS label that conforms with IDNA definitions [RFC5890].

- The definition of <atext> is extended to permit both the RFC 5321 definition and a UTF-8 string. That string MUST NOT contain any of the ASCII graphics or control characters.

The following ABNF rules imported from RFC 5321, Section 4.1.2, are updated directly or indirectly by this document:

- <Mailbox>

- <Local-part>

- <Dot-string>

- <Quoted-string>

- <QcontentSMTP>

- <Domain>

- <Atom>

The following ABNF rule will be imported from RFC 6532, Section 3.1, directly:

- <UTF8-non-ascii>

The following ABNF rule will be imported from RFC 5234, Appendix B.1, directly:

- <DQUOTE>

The following ABNF rule will be imported from RFC 5890, Section 2.3.2.1, directly:

- <U-label>

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

```
sub-domain   =/  U-label
 ; extend the definition of sub-domain in RFC 5321, Section 4.1.2

atext   =/  UTF8-non-ascii
 ; extend the implicit definition of atext in
 ; RFC 5321, Section 4.1.2, which ultimately points to
 ; the actual definition in RFC 5322, Section 3.2.3

qtextSMTP  =/ UTF8-non-ascii
 ; extend the definition of qtextSMTP in RFC 5321, Section 4.1.2

esmtp-value  =/ UTF8-non-ascii
 ; extend the definition of esmtp-value in RFC 5321, Section 4.1.2
```

3.4. MAIL Command Parameter Usage

If the envelope or message being sent requires the capabilities of the SMTPUTF8 extension, the SMTPUTF8-aware SMTP client MUST supply the SMTPUTF8 parameter with the MAIL command. If this parameter is provided, it MUST not accept a value. If the SMTPUTF8-aware SMTP client is aware that neither the envelope nor the message being sent requires any of the SMTPUTF8 extension capabilities, it SHOULD NOT supply the SMTPUTF8 parameter with the MAIL command.

Because there is no guarantee that a next-hop SMTP server will support the SMTPUTF8 extension, use of the SMTPUTF8 extension always carries a risk of transmission failure. In fact, during the early stages of deployment for the SMTPUTF8 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

An SMTPUTF8-aware SMTP client MUST NOT send an internationalized message to an SMTP server that does not support SMTPUTF8. If the SMTP server does not support this option, then the SMTPUTF8-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 SMTPUTF8-aware SMTP server supports enhanced mail system status codes [RFC3463], reply-code "X.6.7" [RFC5248] (see Section 4) is used, meaning "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 specification 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 recipients are unable to accept and process a message with internationalized email headers, the reply-code "554" is used with the meaning "Transaction failed". If the SMTPUTF8-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 "UTF-8 header message cannot be transmitted to one or more recipients, so the message must be rejected".

The SMTPUTF8-aware SMTP servers are encouraged to detect that recipients cannot 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 SMTPUTF8 asserts that a message is an internationalized message or the message being sent needs the SMTPUTF8 support. There is still a chance that a message being sent via the MAIL command with the SMTPUTF8 parameter is not an internationalized message. An SMTPUTF8-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 SMTPUTF8-aware SMTP client or server inspects the message.

Although this specification requires that SMTPUTF8-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 as specified in RFC 2045. The SMTPUTF8 extension MAY be used as follows (assuming it is appropriate given the body content):

- with the BODY=8BITMIME parameter [RFC6152], or

- with the BODY=BINARYMIME parameter, if the SMTP server advertises BINARYMIME [RFC3030].

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 U-label form if the SMTPUTF8 extension is supported; otherwise, it SHOULD be in A-label form.

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 SMTPUTF8 until after it receives the response to the EHLO, the SMTPUTF8-aware SMTP client MUST send only ASCII (LDH label or A-label [RFC5890]) domains in the EHLO command. If the SMTPUTF8-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 (as described in Section 5 of RFC 5321 [RFC5321]), it is strongly advised that all or none of them SHOULD support the SMTPUTF8

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 Section 4.4 of RFC 5321 [RFC5321]. This document updates <Mailbox> and <Domain> to support non-ASCII characters. When the SMTPUTF8 extension is used, the 'Reverse-path' clause of the Return-path-line may include an internationalized domain name that uses the U-label form. Also, 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 SMTPUTF8-aware SMTP client or relay server without the SMTPUTF8 parameter included in the MAIL commands, trace field values must conform to RFC 5321 regardless of the SMTP server's capability.

When an SMTPUTF8-aware SMTP server adds a trace field to a message that was or will be transmitted with the SMTPUTF8 parameter included in the MAIL commands, that server SHOULD use the U-label form for internationalized domain names in the new trace field.

The protocol value of the 'WITH' clause when this extension is used is one of the SMTPUTF8 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 SMTPUTF8 parameter, the SMTPUTF8-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 SMTPUTF8 parameter, the SMTPUTF8-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 SMTPUTF8 Parameter

If the SMTPUTF8 parameter is transmitted with the VRFY and EXPN commands, it indicates that the SMTP client can accept UTF-8 strings in replies to those commands. The parameter with the VRFY and EXPN

commands SHOULD only be used after the SMTP client sees the EHLO response with the SMTPUTF8 keyword. This allows an SMTPUTF8-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 to the VRFY and EXPN commands that contain UTF-8 strings. However, an SMTPUTF8-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 with the VRFY and EXPN commands.

Most replies do not require that a mailbox name be included in the returned text, and therefore a 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 "SMTPUTF8" ] CRLF
 ; String may include Non-ASCII characters

expn = "EXPN" SP String
  [ SP "SMTPUTF8" ] CRLF
 ; String may include Non-ASCII characters
```

The SMTPUTF8 parameter does not accept a value. If the reply to a VRFY or EXPN command requires a UTF-8 string, but the SMTP client did not use the SMTPUTF8 parameter, then the SMTPUTF8-aware SMTP server MUST use either the reply-code 252 or 550. Reply-code 252, defined in RFC 5321 [RFC5321], means "Cannot VRFY user, but will accept the message and attempt the delivery". Reply-code 550, also defined in RFC 5321 [RFC5321], means "Requested action not taken: mailbox unavailable". When the SMTPUTF8-aware SMTP server supports enhanced mail system status codes [RFC3463], the enhanced reply-code as specified below is used. Using the SMTPUTF8 parameter with a VRFY or 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 a UTF-8 string is not allowed in the reply, and the SMTPUTF8-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 SMTPUTF8 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 strings are needed to represent email addresses in responses under the rules specified in this section, this extension does not permit the use of UTF-8 strings for any other purposes. SMTPUTF8-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 has added a new value "SMTPUTF8" to the "SMTP Service Extension" registry of the "Mail Parameters" registry, according to the following data:

Keywords	Description	Reference
SMTPUTF8	Internationalized email address	[RFC6531]

4.2. SMTP Enhanced Status Code Registry

The code definitions in this document replace those specified in RFC 5336, following the guidance in Sections 3.5 and 3.7.4.2 of this document, and based on RFC 5248 [RFC5248]. IANA has updated the "Simple Mail Transfer Protocol (SMTP) Enhanced Status Code Registry" with the following data:

```
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 6531 (Standards 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 6531 (Standards Track)
Submitter:   Jiankang YAO
Change controller: ima@ietf.org


Code:        X.6.9
Sample Text: UTF-8 header message cannot be transferred to one or
             more recipients, 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 6531 (Standards 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 has modified or added the following entries in the "WITH protocol types" sub-registry under the "Mail Transmission Types" registry.

WITH protocol types	Description	Reference
UTF8SMTP	ESMTP with SMTPUTF8	[RFC6531]
UTF8SMTPA	ESMTP with SMTPUTF8 and AUTH	[RFC4954] [RFC6531]
UTF8SMTPS	ESMTP with SMTPUTF8 and STARTTLS	[RFC3207] [RFC6531]
UTF8SMTPSA	ESMTP with SMTPUTF8 and both STARTTLS and AUTH	[RFC3207] [RFC4954] [RFC6531]
UTF8LMTP	LMTP with SMTPUTF8	[RFC6531]
UTF8LMTPA	LMTP with SMTPUTF8 and AUTH	[RFC4954] [RFC6531]
UTF8LMTPS	LMTP with SMTPUTF8 and STARTTLS	[RFC3207] [RFC6531]
UTF8LMTPSA	LMTP with SMTPUTF8 and both STARTTLS and AUTH	[RFC3207] [RFC4954] [RFC6531]

5. Security Considerations

The extended security considerations discussion in the framework document [RFC6530] applies 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:

- the logging systems of SMTP transactions and other logs to monitor the email systems;
- 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 require providing an adequate mechanism 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 by log readers that cannot easily read 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). Sections 3.5 and 7.3 of RFC 5321 give detailed descriptions of use and possible behaviors. Implementation of internationalized addresses can also affect logs and actions by these tools.

6. Acknowledgements

This document revises RFC 5336 [RFC5336] based on the result of the Email Address Internationalization (EAI) working group's discussion. Many EAI working group members did tests and implementations to move this document to the Standards Track. Significant comments and suggestions were received from Xiaodong LEE, Nai-Wen HSU, Yangwoo KO, Yoshiro YONEYA, and other members of JET and were incorporated into the specification. Additional important comments and suggestions, and often specific text, were contributed by many members of the working group 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, 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 with ABNF refinement.

7. References

7.1. 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.

[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.

[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.

[RFC6409] Gellens, R. and J. Klensin, "Message Submission for Mail", STD 72, RFC 6409, November 2011.

[RFC6530] Klensin, J. and Y. Ko, "Overview and Framework for Internationalized Email", RFC 6530, February 2012.

[RFC6532] Yang, A., Steele, S., and N. Freed, "Internationalized Email Headers", RFC 6532, February 2012.

[RFC6533] Hansen, T., Ed., Newman, C., and A. Melnikov, Ed., "Internationalized Delivery Status and Disposition Notifications", RFC RFC6533, February 2012.

7.2. 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
China

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

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

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