

Network Working Group  
Internet-Draft  
Intended status: Experimental  
Expires: April 28, 2010

R. Gellens  
QUALCOMM Incorporated  
C. Newman  
Sun Microsystems  
October 25, 2009

**POP3 Support for UTF-8**  
**draft-ietf-eai-pop-09.txt**

Status of this Memo

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

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

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>.

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

This Internet-Draft will expire on April 28, 2010.

Copyright Notice

Copyright (c) 2009 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 in effect on the date of publication of this document (<http://trustee.ietf.org/license-info>). Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

## Abstract

This specification extends the Post Office Protocol version 3 (POP3) to support un-encoded international characters in user names, passwords, mail addresses, message headers, and protocol-level textual error strings.

## Table of Contents

<a href="#">1.</a>	<a href="#">Introduction</a>	<a href="#">3</a>
<a href="#">1.1.</a>	<a href="#">Conventions Used in this Document</a>	<a href="#">3</a>
<a href="#">1.2.</a>	<a href="#">Change History</a>	<a href="#">4</a>
<a href="#">1.2.1.</a>	<a href="#">Changes from -08 to -09</a>	<a href="#">4</a>
<a href="#">1.2.2.</a>	<a href="#">Changes from -07 to -08</a>	<a href="#">4</a>
<a href="#">1.2.3.</a>	<a href="#">Changes from -06 to -07</a>	<a href="#">5</a>
<a href="#">1.2.4.</a>	<a href="#">Changes from -05 to -06</a>	<a href="#">5</a>
<a href="#">1.2.5.</a>	<a href="#">Changes from -04 to -05</a>	<a href="#">6</a>
<a href="#">1.2.6.</a>	<a href="#">Changes from -03 to -04</a>	<a href="#">6</a>
<a href="#">1.2.7.</a>	<a href="#">Changes from -02 to -03</a>	<a href="#">6</a>
<a href="#">1.2.8.</a>	<a href="#">Changes from -01 to -02</a>	<a href="#">7</a>
<a href="#">1.2.9.</a>	<a href="#">Changes from -00 to -01</a>	<a href="#">7</a>
<a href="#">1.2.10.</a>	<a href="#">Changes from <a href="#">draft-newman-ima-pop</a></a>	<a href="#">7</a>
<a href="#">1.3.</a>	<a href="#">Open Issues</a>	<a href="#">8</a>
<a href="#">2.</a>	<a href="#">LANG Capability</a>	<a href="#">8</a>
<a href="#">3.</a>	<a href="#">UTF8 Capability</a>	<a href="#">10</a>
<a href="#">3.1.</a>	<a href="#">The UTF8 Command</a>	<a href="#">11</a>
<a href="#">3.2.</a>	<a href="#">USER Argument to UTF8 Capability</a>	<a href="#">13</a>
<a href="#">4.</a>	<a href="#">Issues with UTF-8 Header maildrop</a>	<a href="#">13</a>
<a href="#">5.</a>	<a href="#">IANA Considerations</a>	<a href="#">14</a>
<a href="#">6.</a>	<a href="#">Security Considerations</a>	<a href="#">14</a>
<a href="#">7.</a>	<a href="#">References</a>	<a href="#">14</a>
<a href="#">7.1.</a>	<a href="#">Normative References</a>	<a href="#">14</a>
<a href="#">7.2.</a>	<a href="#">Informative References</a>	<a href="#">15</a>
<a href="#">Appendix A.</a>	<a href="#">Design Rationale</a>	<a href="#">16</a>
<a href="#">Appendix B.</a>	<a href="#">Acknowledgments</a>	<a href="#">16</a>
	<a href="#">Authors' Addresses</a>	<a href="#">16</a>



## **1. Introduction**

This document forms part of the Email Address Internationalization (EAI) experiment described in the EAI Framework [\[RFC4952\]](#) document (for background, please see the charter of the EAI working group) and should be evaluated within the context of EAI. As part of the overall EAI work, email messages may be transmitted and delivered containing un-encoded UTF-8 characters, and mail drops accessed using POP3 [\[RFC1939\]](#) might natively store UTF-8.

This specification extends POP3 [\[RFC1939\]](#) using the POP3 Extension Mechanism [\[RFC2449\]](#) to permit un-encoded UTF-8 [\[RFC3629\]](#) in headers as described in Internationalized Email Headers [\[RFC5335\]](#). It also adds a mechanism to support login names outside the ASCII character set, and a mechanism to support UTF-8 protocol-level error strings in a language appropriate for the user.

This document updates POP3 [\[RFC1939\]](#), and the fact that an Experimental specification updates a Standards-Track specification means that people who participate in the experiment have to consider the standard updated. Note that, as an Experimental document, there is no "Updates" header. If and when a version of this document moves to the standards track, an "Updates: 1939" header should be added.

Within this specification, the term down-conversion refers to the process of modifying a message containing UTF8 headers [\[RFC5335\]](#) or body parts with 8bit content-transfer-encoding as defined in MIME [section 2.8 \[RFC2045\]](#) into conforming 7-bit Internet Message Format [\[RFC5322\]](#) with Message Header Extensions for Non-ASCII Text [\[RFC2047\]](#) and other 7-bit encodings. Down-conversion is specified by Downgrading mechanism for Email Address Internationalization [\[RFC5504\]](#).

### **1.1. Conventions Used in this Document**

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in "Key words for use in RFCs to Indicate Requirement Levels" [\[RFC2119\]](#).

The formal syntax uses the Augmented Backus-Naur Form (ABNF) [\[RFC5234\]](#) notation including the core rules defined in [Appendix B of RFC 5234](#).

In examples, "C:" and "S:" indicate lines sent by the client and server respectively. If a single "C:" or "S:" label applies to multiple lines, then the line breaks between those lines are for editorial clarity only and are not part of the actual protocol



exchange.

Note that examples always use 7-bit ASCII characters due to limitations of this document format; in particular, some examples for the "LANG" command may appear silly as a result.

## **1.2. Change History**

This section describes the change history of this Internet draft and will be removed when/if this is published as an RFC.

### **1.2.1. Changes from -08 to -09**

- o Added new paragraph to start of Introduction to more clearly explain this document within the larger EAI context.
- o Added informative reference to EAI Framework ([RFC 4952](#)).
- o Removed "Updates: 1939" header and added a note that one should be added if and when it is published on the standards track.
- o Added clarifying text that the "language range" argument of [RFC 4647](#) is the "Basic Language Range".

### **1.2.2. Changes from -07 to -08**

- o Changed wording on applying SASLprep to APOP digest inputs.
- o Added mandatory rejection of user names or passwords which fail to comply with formal syntax of [RFC 3629](#).
- o Added text that, when applying SASLprep, servers MUST reject user names or passwords which contain characters listed in [section 2.3 of RFC 4013](#).
- o Added normative reference to [RFC 3629](#).
- o Changed SASLprep language so that both clients and servers MUST apply SASLprep to user names and passwords used to compute APOP digest, and servers SHOULD apply SASLprep to arguments of USER and PASS.
- o Fixed typo ("ACII" instead of "ASCII").
- o Clarified that size doesn't include byte-stuffing.
- o Added explanation to Introduction regarding updating [RFC 1939](#).



- o Added more prominent text on examples which are silly because they use 7-bit ASCII.
- o Replaced French examples with Spanish to try and make them slightly less embarrassing.
- o Replaced French-Canadian (fr-ca) example with Swedish, to try and avoid accented characters. Also added Swedish to language listing.
- o Added "Examples" to LANG command examples.
- o Added introductory text to sections on LANG and UTF8 capability tags.

#### **1.2.3. Changes from -06 to -07**

- o Added discussion about accuracy of size.
- o Added mention of potential buffer overflow problems because of inaccurate sizes to the Security Considerations.
- o Added informative reference to SASL for POP3 ([RFC 5034](#)).
- o Removed text making changes to AUTH, as this is handled by POP3 SASL.
- o Fixed typo ("depricated" instead of "deprecated").
- o Reworded Design Rationale appendix.

#### **1.2.4. Changes from -05 to -06**

- o Removed LIST and TOP as possible arguments to the UTF8 tag in the CAPA response.
- o Clarified that the UTF8 command has no parameters.
- o Changed "arguments" to "arguments with CAPA tag" to clarify that these are possible arguments to the tag in the CAPA response and not command parameters.
- o Clarified use of "argument" to refer to CAPA tag and "parameter" to refer to commands.
- o Clarified that free-form text is non-standard.





- o Removed open issue (downgrading).
- o Added discussion of downgrading to [Appendix A](#).
- o Updated downgrade reference to [RFC 5504](#).
- o Tweaked [RFC 2119](#) text to satisfy I-D nit checker.

#### **1.2.5. Changes from -04 to -05**

- o Downgrading is back to an informative, not normative reference, and is suggested as a good idea but explicitly not required.
- o Language listing now specifies that the human-readable description of a language is in the language itself.
- o Updated 2822 reference to 5322, made text "Internet Message Format".
- o Updated reference to utf8headers draft to [RFC5335](#).
- o Updated reference to [RFC4234](#) to [RFC5234](#).

#### **1.2.6. Changes from -03 to -04**

- o Specified that it is an error to issue STLS after UTF8.
- o Removed prior open issues.
- o Downgrading added as open issue.

#### **1.2.7. Changes from -02 to -03**

- o Updated references.
- o Replaced US-ASCII with ASCII.
- o Added comment to language listing failure example.
- o Replaced RET8, LST8, and TOP8 commands with a single mode-switch UTF8 command issued before authentication. This simplifies the protocol, and allows servers to optionally down-convert a cache of the maildrop prior to issuing the +OK response entering TRANSACTION state.
- o Removed most up-conversion material.



- o Removed definition of up-conversion.
- o Removed IMAP4 reference.
- o Added AUTH command to those affected by UTF8 capability.
- o Removed LST8 and TOP8 capability parameters and commands.
- o Removed NO-RETR capability. POP servers are now unconditionally required to support down-conversion of UTF8-native maildrops.
- o Added sentence about modifying authentication code to Security Considerations.
- o eai-downgrade draft is now normative and required.
- o Deleted references to RFCs 1341, 1847, 2049, 2183, 3501, 3516, and 3490.

#### **1.2.8. Changes from -01 to -02**

- o Minor grammatical tweaks.
- o Add passwords to Abstract.
- o Removed new editor's name from Acknowledgments.

#### **1.2.9. Changes from -00 to -01**

- o Update references

#### **1.2.10. Changes from [draft-newman-ima-pop](#)**

- o Change title to make this a WG document.
- o Add LANG command and extension.
- o Rename RET8 capability to UTF8 and add sub-sections for arguments.
- o Add TOP8 command.
- o Add definition of up-conversion and down-conversion.
- o Some grammar fix-ups and section re-ordering based on RFC editor style.



### **1.3. Open Issues**

1. none

## **2. LANG Capability**

Per the POP3 Extension Mechanism [[RFC2449](#)], this document adds a new capability response tag to indicate support for a new command: LANG. The capability tag and new command are described below.

CAPA tag:

LANG

Arguments with CAPA tag:

none

Added Commands:

LANG

Standard commands affected:

All

Announced states / possible differences:

both / no

Commands valid in states:

AUTHENTICATION, TRANSACTION

Specification reference:

this document

Discussion:

POP3 allows most +OK and -ERR server responses to include human-readable text that in some cases needs to be presented to the user. But that text is limited to ASCII by the POP3 specification [[RFC1939](#)]. The LANG capability and command permit a POP3 client to negotiate which language the server should use when sending human-readable text.

A server that advertises the LANG extension MUST use the language "i-default" as described in [[RFC2277](#)] as its default language until another supported language is negotiated by the client. A server MUST include "i-default" as one of its supported languages.

The LANG command requests that human-readable text included in all subsequent +OK and -ERR responses be localized to a language matching



the language range argument (the "Basic Language Range" as described by [RFC4647](#)). If the command succeeds, the server returns a +OK response followed by a single space, the exact language tag selected, another space, and the rest of the line is human-readable text in the appropriate language. This and subsequent protocol-level human readable text is encoded in the UTF-8 charset.

If the command fails, the server returns an -ERR response and subsequent human-readable response text continues to use the language that was previously active (typically i-default).

The special "\*" language range argument indicates a request to use a language designated as preferred by the server administrator. The preferred language MAY vary based on the currently active user.

If no argument is given and the POP3 server issues a positive response, then the response given is multi-line. After the initial +OK, for each language tag the server supports, the POP3 server responds with a line for that language. This line is called a "language listing".

In order to simplify parsing, all POP3 servers are required to use a certain format for language listings. A language listing consists of the language tag [RFC4646](#) of the message, optionally followed by a single space and a human readable description of the language in the language itself, using the UTF-8 charset.

Examples:

< Note that some examples do not include the correct character accents due to limitations of this document format. >

< The server defaults to using English i-default responses until the client explicitly changes the language. >

```
C: USER karen
S: +OK Hello, karen
C: PASS password
S: +OK karen's maildrop contains 2 messages (320 octets)
```

< Client requests deprecated MUL language. Server replies with -ERR response >

```
C: LANG MUL
S: -ERR invalid language MUL
```

< A LANG command with no parameters is a request for a language listing. >





```
C: LANG
S: +OK Language listing follows:
S: en English
S: en-boont English Boontling dialect
S: de Deutsch
S: it Italiano
S: es Espanol
S: sv Svenska
S: i-default Default language
S: .
```

< A request for a language listing might fail >

```
C: LANG
S: -ERR Server is unable to list languages
```

< Once the client changes the language, all responses will be in that language starting with the response to the LANG command.

```
C: LANG es
S: +OK es Idioma cambiado
```

< If a server does not support the requested primary language, responses will continue to be returned in the current language the server is using. >

```
C: LANG uga
S: -ERR es Idioma <<UGA>> no es conocido
```

```
C: LANG sv
S: +OK sv Kommandot "LANG" lyckades
```

```
C: LANG *
S: +OK es Idioma cambiado
```

### Examples

## 3. UTF8 Capability

Per the POP3 Extension Mechanism [[RFC2449](#)], this document adds a new capability response tag to indicate support for new server functionality including a new command, UTF8. The capability tag and new command and functionality are described below.



CAPA tag:  
UTF8

Arguments with CAPA tag:  
USER

Added Commands:  
UTF8

Standard commands affected:  
USER, PASS, APOP, LIST, TOP, RETR

Announced states / possible differences:  
both / no

Commands valid in states:  
AUTHORIZATION

Specification reference:  
this document

Discussion:

This capability adds the "UTF8" command to POP3. The UTF8 command switches the session from ASCII to UTF8 mode.

### **3.1. The UTF8 Command**

The UTF8 command enables UTF8 mode. The UTF8 command has no parameters.

Maildrops can natively store UTF8 or be limited to ASCII. UTF8 mode has no effect on messages in an ASCII-only maildrop. Messages in native-UTF8 maildrops can be ASCII or UTF8 using internationalized headers [[RFC5335](#)] and/or 8bit content-transfer-encoding as defined in MIME [section 2.8](#) [[RFC2045](#)]. In UTF8 mode, both UTF8 and ASCII messages are sent to the client as-is (without conversion). When not in UTF8 mode, UTF8 messages in a native UTF8 maildrop MUST be down-converted (downgraded) to comply with unextended POP and Internet Mail Format. POP servers (unlike SMTP and Submit servers) are not required to use Downgrading mechanism for Email Address Internationalization [[RFC5504](#)].

Discussion: The main argument against a single required mechanism for downgrade by a POP server is that the only clients that have any use for a standardized downgraded message (because they wish to interpret downgrade headers, for example) are ones that can support UTF8 and hence will issue the UTF8 command in the first place. The counter



argument to this is that non-UTF8 clients might be upgraded in the future; it's desirable for an upgraded client to be capable of interpreting prior downgraded messages in the local mail store, which is most likely if the messages were downgraded using one standardized procedure.

Therefore, while POP servers are not required to use the Downgrading mechanism for Email Address Internationalization [[RFC5504](#)], there are advantages to them doing so.

Note that even in UTF8 mode, MIME binary content-transfer-encoding is still not permitted.

The octet count (size) of a message reported in a response to the LIST command SHOULD match the actual number of octets sent in a RETR response (not counting byte-stuffing). Sizes reported elsewhere, such as in STAT responses and non-standardized free-form text in positive status indicators (following "+OK") need not be accurate, but it is preferable if they are.

Discussion: Mail stores are either ASCII or native UTF-8, and clients either issue the UTF8 command or not. The message needs converting only when it is native UTF8 and the client has not issued the UTF8 command, in which case the server must downconvert it. The downconverted message may be larger. The server may choose various strategies regarding downconversion, which include when to downconvert, whether to cache or store the downconverted form of a message (and if so, for how long), and whether to calculate or retain the size of a downconverted message independently of the downconverted content. If the server does not have immediate access to the accurate downconverted size, it may be faster to estimate rather than calculate it. Servers are expected to normally follow the [RFC 1939](#) [[RFC1939](#)] text on using the "exact size" in a scan listing, but there may be situations with maildrops containing very large numbers of messages in which this might be a problem. If the server does estimate, reporting a scan listing size smaller than what it turns out to be could be a problem for some clients. In summary, it is better for servers to report accurate sizes, but if not, high guesses are better than small ones. Some POP servers include the message size in the non-standardized text response following "+OK" (the 'text' production of [RFC 2449](#) [[RFC2449](#)]), in a RETR or TOP response (possibly because some examples in POP3 [[RFC1939](#)] do so). There has been at least one known case of a client relying on this to know when it had received all of the message rather than following the POP3 [[RFC1939](#)] rule of looking for a line consisting of a termination octet (".") and a CRLF pair. While any such client is non-compliant, if a server does include the size in such text, it is better if it is accurate.



Clients MUST NOT issue the STLS command [[RFC2595](#)] after issuing UTF8; servers MAY (but are not required to) enforce this by rejecting with an "-ERR" response an STLS command issued subsequent to a successful UTF8 command. (Because this is a protocol error as opposed to a failure based on conditions, an extended response code [[RFC2449](#)] is not specified.)

### **[3.2.](#) USER Argument to UTF8 Capability**

If the USER argument is included with this capability, it indicates that the server accepts UTF-8 user names and passwords.

Servers which include the USER argument in the UTF8 capability response SHOULD apply SASLprep [[RFC4013](#)] to the arguments of the USER and PASS commands.

A client or server that supports APOP and permits UTF-8 in user names or passwords MUST apply SASLprep [[RFC4013](#)] to the user name and password used to compute the APOP digest.

When applying SASLprep [[RFC4013](#)], servers MUST reject UTF-8 user names or passwords which contain a Unicode character listed in [section 2.3](#) of SASLprep [[RFC4013](#)].

The client does not need to issue the UTF8 command prior to using UTF8 in authentication. However, clients MUST NOT use UTF8 in USER, PASS, or APOP commands unless the USER argument is included in the UTF8 capability response.

The server MUST reject UTF-8 user names or passwords which fail to comply with the formal syntax in UTF-8 [[RFC3629](#)].

Use of UTF8 in the AUTH command is governed by the POP3 SASL [[RFC5034](#)] mechanism.

## **[4.](#) Issues with UTF-8 Header maildrop**

When a POP3 server uses a UTF8-native maildrop, it is the responsibility of the server to comply with the POP3 base specification [[RFC1939](#)] and Internet Message Format [[RFC5322](#)] when not in UTF8 mode. Mechanisms for 7-bit downgrading to help comply with the standards are described in Downgrading mechanism for Email Address Internationalization [[RFC5504](#)].





## **5. IANA Considerations**

This adds two new capabilities ("UTF8" and "LANG") to the POP3 capability registry [[RFC2449](#)].

## **6. Security Considerations**

The security considerations of UTF-8 [[RFC3629](#)] and SASLprep [[RFC4013](#)] apply to this specification, particularly with respect to use of UTF-8 in user names and passwords.

The "LANG \*" command can reveal the existence and preferred language of a user to an active attacker probing the system if the active language changes in response to the USER, PASS, or APOP commands prior to validating the user's credentials. Servers MUST implement a configuration to prevent this exposure.

It is possible for a man-in-the-middle attacker to insert a LANG command in the command stream thus making protocol-level diagnostic responses unintelligible to the user. A mechanism to integrity protect the session, such as TLS [[RFC2595](#)] can be used to defeat such attacks.

Modifying server authentication code (in this case, to support UTF8) needs to be done with care to avoid introducing vulnerabilities (for example, in string parsing).

The UTF8 Command ([Section 3.1](#)) description contains a discussion on reporting inaccurate sizes. An additional risk to doing so is that, if a client allocates buffers based on the reported size, it may overrun the buffer, crash, or have other problems if the message data is larger than reported.

## **7. References**

### **7.1. Normative References**

- [RFC1939] Myers, J. and M. Rose, "Post Office Protocol - Version 3", STD 53, [RFC 1939](#), May 1996.
- [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.

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC2277] Alvestrand, H., "IETF Policy on Character Sets and Languages", [BCP 18](#), [RFC 2277](#), January 1998.
- [RFC2449] Gellens, R., Newman, C., and L. Lundblade, "POP3 Extension Mechanism", [RFC 2449](#), November 1998.
- [RFC5322] Resnick, P., Ed., "Internet Message Format", [RFC 5322](#), October 2008.
- [RFC4646] Phillips, A. and M. Davis, "Tags for Identifying Languages", [BCP 47](#), [RFC 4646](#), September 2006.
- [RFC4647] Phillips, A. and M. Davis, "Matching of Language Tags", [BCP 47](#), [RFC 4647](#), September 2006.
- [RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", STD 63, [RFC 3629](#), November 2003.
- [RFC4013] Zeilenga, K., "SASLprep: Stringprep Profile for User Names and Passwords", [RFC 4013](#), February 2005.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, [RFC 5234](#), January 2008.
- [RFC5335] Abel, Y., "Internationalized Email Headers", [RFC 5335](#), September 2008.

## **[7.2.](#) Informative References**

- [RFC2595] Newman, C., "Using TLS with IMAP, POP3 and ACAP", [RFC 2595](#), June 1999.
- [RFC4952] Klensin, J. and Y. Ko, "Overview and Framework for Internationalized Email", [RFC 4952](#), July 2007.
- [RFC5034] Siemborski, R. and A. Menon-Sen, "The Post Office Protocol (POP3) Simple Authentication and Security Layer (SASL) Authentication Mechanism", [RFC 5034](#), July 2007.
- [RFC5504] Fujiwara, K. and Y. Yoneya, "Downgrading Mechanism for Email Address Internationalization", [RFC 5504](#), March 2009.



## [Appendix A.](#) Design Rationale

This non-normative section discusses the reasons behind some of the design choices in the above specification.

Having servers perform up-conversion so that, at a minimum, [RFC2047](#)-encoded words are decoded into UTF8 is tempting, since this is an area that clients often fail to correctly implement. However, after much discussion the group felt that the benefits did not justify the burden.

Due to interoperability problems with [RFC 2047](#) and limited deployment of [RFC 2231](#), it is hoped these 7-bit encoding mechanisms can be deprecated in the future when UTF-8 header support becomes prevalent.

USER is optional because the implementation burden of SASLprep [[RFC4013](#)] is not well understood and mandating such support in all cases could negatively impact deployment.

While it is possible to provide useful examples for language negotiation without support for non-ASCII characters, it is difficult to provide useful examples for commands specifically designed to use the UTF-8 charset un-encoded when the document format is limited to ASCII. As a result, there are no plans to provide examples for that part of the specification as long as this remains an experimental proposal. However, implementers of this specification are encouraged to provide examples to the document author for a future revision.

While down-conversion of native-UTF8 messages is mandatory in the absence of the UTF8 command, servers are not required to do so as specified in Downgrading Mechanism [[RFC5504](#)]. As clients are upgraded with UTF8 support and the ability to intelligently handle (e.g., display and reply to) UTF8 messages that were downgraded in transit, it is better if they are also able to handle messages in the local mail store that were downgraded by the POP server. This is more likely if the POP server downgrades messages using the same mechanism as an SMTP server.

## [Appendix B.](#) Acknowledgments

Thanks to John Klensin, Tony Hansen and other EAI working group participants who provided helpful suggestions and interesting debate that improved this specification.



Authors' Addresses

Randall Gellens  
QUALCOMM Incorporated  
5775 Morehouse Drive  
San Diego, CA 92651  
US

Email: [rg+ietf@qualcomm.com](mailto:rg+ietf@qualcomm.com)

Chris Newman  
Sun Microsystems  
800 Royal Oaks  
Monrovia, CA 91016-6347  
US

Email: [chris.newman@sun.com](mailto:chris.newman@sun.com)



