Overview and Framework for Internationalized Email

Status of This Memo

Copyright Notice

Abstract

   Full use of electronic mail throughout the world requires that people
   be able to use their own names, written correctly in their own
   languages and scripts, as mailbox names in email addresses. This
   document introduces a series of specifications that define mechanisms
   and protocol extensions needed to fully support internationalized
   email addresses. These changes include an SMTP extension and
   extension of email header syntax to accommodate UTF-8 data. The
   document set also includes discussion of key assumptions and issues
   in deploying fully internationalized email.

Table of Contents

1. Introduction

   In order to use internationalized email addresses, we need to
   internationalize both the domain part and the local part of email
   addresses. The domain part of email addresses is already
   internationalized [RFC3490], while the local part is not. Without
   the extensions specified in this document, the mailbox name is
   restricted to a subset of 7-bit ASCII [RFC2821]. Though MIME
   [RFC2045] enables the transport of non-ASCII data, it does not
   provide a mechanism for internationalized email addresses. In RFC
   2047 [RFC2047], MIME defines an encoding mechanism for some specific
   message header fields to accommodate non-ASCII data. However, it
   does not permit the use of email addresses that include non-ASCII
   characters. Without the extensions defined here, or some equivalent
   set, the only way to incorporate non-ASCII characters in any part of
   email addresses is to use RFC 2047 coding to embed them in what RFC
   2822 [RFC2822] calls the "display name" (known as a "name phrase" or
   by other terms elsewhere) of the relevant headers. Information coded
   into the display name is invisible in the message envelope and, for
   many purposes, is not part of the address at all.

1.1. Role of This Specification

   This document presents the overview and framework for an approach to
   the next stage of email internationalization. This new stage
   requires not only internationalization of addresses and headers, but
   also associated transport and delivery models.

   This document provides the framework for a series of experimental
   specifications that, together, provide the details for a way to
   implement and support internationalized email. The document itself
   describes how the various elements of email internationalization fit
   together and how the relationships among the various documents are
   involved.

1.2. Problem Statement

   Internationalizing Domain Names in Applications (IDNA) [RFC3490]
   permits internationalized domain names, but deployment has not yet
   reached most users. One of the reasons for this is that we do not
   yet have fully internationalized naming schemes. Domain names are
   just one of the various names and identifiers that are required to be
   internationalized. In many contexts, until more of those identifiers
   are internationalized, internationalized domain names alone have
   little value.

   Email addresses are prime examples of why it is not good enough to
   just internationalize the domain name. As most of us have learned

from experience, users strongly prefer email addresses that resemble
names or initials to those involving seemingly meaningless strings of
letters or numbers. Unless the entire email address can use familiar
characters and formats, users will perceive email as being culturally
unfriendly. If the names and initials used in email addresses can be
expressed in the native languages and writing systems of the users,
the Internet will be perceived as more natural, especially by those
whose native language is not written in a subset of a Roman-derived
script.

Internationalization of email addresses is not merely a matter of
changing the SMTP envelope; or of modifying the From, To, and Cc
headers; or of permitting upgraded Mail User Agents (MUAs) to decode
a special coding and respond by displaying local characters. To be
perceived as usable, the addresses must be internationalized and
handled consistently in all of the contexts in which they occur.
This requirement has far-reaching implications: collections of
patches and workarounds are not adequate. Even if they were
adequate, a workaround-based approach may result in an assortment of
implementations with different sets of patches and workarounds having
been applied with consequent user confusion about what is actually
usable and supported. Instead, we need to build a fully
internationalized email environment, focusing on permitting efficient
communication among those who share a language or other community.
That, in turn, implies changes to the mail header environment to
permit the full range of Unicode characters where that makes sense,
an SMTP Extension to permit UTF-8 [RFC3629] mail addressing and
delivery of those extended headers, and (finally) a requirement for
support of the 8BITMIME SMTP extension [RFC1652] so that all of these
can be transported through the mail system without having to overcome
the limitation that headers do not have content-transfer-encodings.

1.3. Terminology

This document assumes a reasonable understanding of the protocols and
terminology of the core email standards as documented in [RFC2821]
and [RFC2822].

Much of the description in this document depends on the abstractions
of "Mail Transfer Agent" ("MTA") and "Mail User Agent" ("MUA").
However, it is important to understand that those terms and the
underlying concepts postdate the design of the Internet's email
architecture and the application of the "protocols on the wire"
principle to it. That email architecture, as it has evolved, and the
"wire" principle have prevented any strong and standardized
distinctions about how MTAs and MUAs interact on a given origin or
destination host (or even whether they are separate).

However, the term "final delivery MTA" is used in this document in a
fashion equivalent to the term "delivery system" or "final delivery
system" of RFC 2821. This is the SMTP server that controls the
format of the local parts of addresses and is permitted to inspect
and interpret them. It receives messages from the network for
delivery to mailboxes or for other local processing, including any
forwarding or aliasing that changes envelope addresses, rather than
relaying. From the perspective of the network, any local delivery
arrangements such as saving to a message store, handoff to specific
message delivery programs or agents, and mechanisms for retrieving
messages are all "behind" the final delivery MTA and hence are not
part of the SMTP transport or delivery process.

In this document, an address is "all-ASCII", or just an "ASCII
address", if every character in the address is in the ASCII character
repertoire [ASCII]; an address is "non-ASCII", or an "i18n-address",
if any character is not in the ASCII character repertoire. Such
addresses may be restricted in other ways, but those restrictions are
not relevant to this definition. The term "all-ASCII" is also
applied to other protocol elements when the distinction is important,
with "non-ASCII" or "internationalized" as its opposite.

The umbrella term to describe the email address internationalization
specified by this document and its companion documents is "UTF8SMTP".
For example, an address permitted by this specification is referred
to as a "UTF8SMTP (compliant) address".

Please note that, according to the definitions given here, the set of
all "all-ASCII" addresses and the set of all "non-ASCII" addresses
are mutually exclusive. The set of all UTF8SMTP addresses is the
union of these two sets.

An "ASCII user" (i) exclusively uses email addresses that contain
ASCII characters only, and (ii) cannot generate recipient addresses
that contain non-ASCII characters.

An "i18mail user" has one or more non-ASCII email addresses. Such a
user may have ASCII addresses too; if the user has more than one
email account and a corresponding address, or more than one alias for
the same address, he or she has some method to choose which address
to use on outgoing email. Note that under this definition, it is not
possible to tell from an ASCII address if the owner of that address
is an i18mail user or not. (A non-ASCII address implies a belief
that the owner of that address is an i18mail user.) There is no such
thing as an "i18mail message"; the term applies only to users and
their agents and capabilities.

A "message" is sent from one user (sender) using a particular email
address to one or more other recipient email addresses (often
referred to just as "users" or "recipient users").

A "mailing list" is a mechanism whereby a message may be distributed
to multiple recipients by sending it to one recipient address. An
agent (typically not a human being) at that single address then
causes the message to be redistributed to the target recipients.
This agent sets the envelope return address of the redistributed
message to a different address from that of the original single
recipient message. Using a different envelope return address
(reverse-path) causes error (and other automatically generated)
messages to go to an error handling address.

As specified in RFC 2821, a message that is undeliverable for some
reason is expected to result in notification to the sender. This can
occur in either of two ways. One, typically called "Rejection",
occurs when an SMTP server returns a reply code indicating a fatal
error (a "5yz" code) or persistently returns a temporary failure
error (a "4yz" code). The other involves accepting the message
during SMTP processing and then generating a message to the sender,
typically known as a "Non-delivery Notification" or "NDN". Current
practice often favors rejection over NDNs because of the reduced
likelihood that the generation of NDNs will be used as a spamming
technique. The latter, NDN, case is unavoidable if an intermediate
MTA accepts a message that is then rejected by the next-hop server.

The pronouns "he" and "she" are used interchangeably to indicate a
human of indeterminate gender.

The key words "MUST", "SHALL", "REQUIRED", "SHOULD", "RECOMMENDED",
and "MAY" in this document are to be interpreted as described in RFC
2119 [RFC2119].

2. Overview of the Approach

   This set of specifications changes both SMTP and the format of email
   headers to permit non-ASCII characters to be represented directly.
   Each important component of the work is described in a separate
   document. The document set, whose members are described in the next
   section, also contains informational documents whose purpose is to
   provide implementation suggestions and guidance for the protocols.

3. Document Plan

   In addition to this document, the following documents make up this
   specification and provide advice and context for it.

o SMTP extensions. This document [EAI-SMTPext] provides an SMTP
   extension for internationalized addresses, as provided for in RFC
   2821.

o Email headers in UTF-8. This document [EAI-UTF8] essentially
   updates RFC 2822 to permit some information in email headers to be
   expressed directly by Unicode characters encoded in UTF-8 when the
   SMTP extension described above is used. This document, possibly
   with one or more supplemental ones, will also need to address the
   interactions with MIME, including relationships between UTF8SMTP
   and internal MIME headers and content types.

o In-transit downgrading from internationalized addressing with the
   SMTP extension and UTF-8 headers to traditional email formats and
   characters [EAI-downgrade]. Downgrading either at the point of
   message origination or after the mail has successfully been
   received by a final delivery SMTP server involve different
   constraints and possibilities; see Section 4.3 and Section 5,
   below. Processing that occurs after such final delivery,
   particularly processing that is involved with the delivery to a
   mailbox or message store, is sometimes called "Message Delivery"
   processing.

o Extensions to the IMAP protocol to support internationalized
   headers [EAI-imap].

o Parallel extensions to the POP protocol [EAI-pop].

o Description of internationalization changes for delivery
   notifications (DSNs) [EAI-DSN].

o Scenarios for the use of these protocols [EAI-scenarios].

4. Overview of Protocol Extensions and Changes

4.1. SMTP Extension for Internationalized Email Address

   An SMTP extension, "UTF8SMTP" is specified as follows:

o Permits the use of UTF-8 strings in email addresses, both local
   parts and domain names.

o Permits the selective use of UTF-8 strings in email headers (see
   Section 4.2).

o Requires that the server advertise the 8BITMIME extension
   [RFC1652] and that the client support 8-bit transmission so that
   header information can be transmitted without using special
   content-transfer-encoding.

o Provides information to support downgrading mechanisms.

Some general principles affect the development decisions underlying
this work.

1. Email addresses enter subsystems (such as a user interface) that
    may perform charset conversions or other encoding changes. When
    the left hand side of the address includes characters outside the
    US-ASCII character repertoire, use of punycode on the right hand
    side is discouraged to promote consistent processing of
    characters throughout the address.

2. An SMTP relay must

    * Either recognize the format explicitly, agreeing to do so via
       an ESMTP option,

    * Select and use an ASCII-only address, downgrading other
       information as needed (see Section 4.3), or

    * Reject the message or, if necessary, return a non-delivery
       notification message, so that the sender can make another
       plan.

    If the message cannot be forwarded because the next-hop system
    cannot accept the extension and insufficient information is
    available to reliably downgrade it, it MUST be rejected or a non-
    delivery message generated and sent.

3. In the interest of interoperability, charsets other than UTF-8
    are prohibited in mail addresses and headers. There is no
    practical way to identify them properly with an extension similar
    to this without introducing great complexity.

Conformance to the group of standards specified here for email
transport and delivery requires implementation of the SMTP Extension
specification, including recognition of the keywords associated with
alternate addresses, and the UTF-8 Header specification. Support for
downgrading is not required, but, if implemented, MUST be implemented
as specified. Similarly, if the system implements IMAP or POP, it
MUST conform to the i18n IMAP or POP specifications respectively.

4.2. Transmission of Email Header Fields in UTF-8 Encoding

   There are many places in MUAs or in a user presentation in which
   email addresses or domain names appear. Examples include the
   conventional From, To, or Cc header fields; Message-ID and
   In-Reply-To header fields that normally contain domain names (but
   that may be a special case); and in message bodies. Each of these
   must be examined from an internationalization perspective. The user
   will expect to see mailbox and domain names in local characters, and
   to see them consistently. If non-obvious encodings, such as
   protocol-specific ASCII-Compatible Encoding (ACE) variants, are used,
   the user will inevitably, if only occasionally, see them rather than
   "native" characters and will find that discomfiting or astonishing.
   Similarly, if different codings are used for mail transport and
   message bodies, the user is particularly likely to be surprised, if
   only as a consequence of the long-established "things leak"
   principle. The only practical way to avoid these sources of
   discomfort, in both the medium and the longer term, is to have the
   encodings used in transport be as similar to the encodings used in
   message headers and message bodies as possible.

   When email local parts are internationalized, it seems clear that
   they should be accompanied by arrangements for the email headers to
   be in the fully internationalized form. That form should presumably
   use UTF-8 rather than ASCII as the base character set for the
   contents of header fields (protocol elements such as the header field
   names themselves will remain entirely in ASCII). For transition
   purposes and compatibility with legacy systems, this can done by
   extending the encoding models of [RFC2045] and [RFC2231]. However,
   our target should be fully internationalized headers, as discussed in
   [EAI-UTF8].

4.3. Downgrading Mechanism for Backward Compatibility

   As with any use of the SMTP extension mechanism, there is always the
   possibility of a client that requires the feature encountering a
   server that does not support the required feature. In the case of
   email address and header internationalization, the risk should be
   minimized by the fact that the selection of submission servers are
   presumably under the control of the sender's client and the selection
   of potential intermediate relays is under the control of the
   administration of the final delivery server.

   For situations in which a client that needs to use UTF8SMTP
   encounters a server that does not support the extension UTF8SMTP,
   there are two possibilities:

o Reject the message or generate and send a non-delivery message,
   requiring the sender to resubmit it with traditional-format
   addresses and headers.

o Figure out a way to downgrade the envelope or message body in
   transit. Especially when internationalized addresses are
   involved, downgrading will require that all-ASCII addresses be
   obtained from some source. An optional extension parameter is
   provided as a way of transmitting an alternate address. Downgrade
   issues and a specification are discussed in [EAI-downgrade].

(The client can also try an alternate next-hop host or requeue the
message and try later, on the assumption that the lack of UTF8SMTP is
a transient failure; since this ultimately resolves to success or
failure, it doesn't change the discussion here.)

The first of these two options, that of rejecting or returning the
message to the sender MAY always be chosen.

If a UTF8SMTP capable client is sending a message that does not
require the extended capabilities, it SHOULD send the message whether
or not the server announces support for the extension. In other
words, both the addresses in the envelope and the entire set of
headers of the message are entirely in ASCII (perhaps including
encoded words in the headers). In that case, the client SHOULD send
the message whether or not the server announces the capability
specified here.

5. Downgrading before and after SMTP Transactions

   In addition to the in-transit downgrades discussed above, downgrading
   may also occur before or during the initial message submission or
   after the delivery to the final delivery MTA. Because these cases
   have a different set of available information from in-transit cases,
   the constraints and opportunities may be somewhat different too.
   These two cases are discussed in the subsections below.

5.1. Downgrading before or during Message Submission

   Perhaps obviously, the most convenient time to find an ASCII address
   corresponding to an internationalized address is at the originating
   MUA. This can occur either before the message is sent or after the
   internationalized form of the message is rejected. It is also the
   most convenient time to convert a message from the internationalized
   form into conventional ASCII form or to generate a non-delivery
   message to the sender if either is necessary. At that point, the
   user has a full range of choices available, including contacting the
   intended recipient out of band for an alternate address, consulting

appropriate directories, arranging for translation of both addresses
and message content into a different language, and so on. While it
is natural to think of message downgrading as optimally being a
fully-automated process, we should not underestimate the capabilities
of a user of at least moderate intelligence who wishes to communicate
with another such user.

In this context, one can easily imagine modifications to message
submission servers (as described in [RFC4409]) so that they would
perform downgrading, or perhaps even upgrading, operations, receiving
messages with one or more of the internationalization extensions
discussed here and adapting the outgoing message, as needed, to
respond to the delivery or next-hop environment it encounters.

5.2. Downgrading or Other Processing After Final SMTP Delivery

When an email message is received by a final delivery SMTP server, it
is usually stored in some form. Then it is retrieved either by
software that reads the stored form directly or by client software
via some email retrieval mechanisms such as POP or IMAP.

The SMTP extension described in Section 4.1 provides protection only
in transport. It does not prevent MUAs and email retrieval
mechanisms that have not been upgraded to understand
internationalized addresses and UTF-8 headers from accessing stored
internationalized emails.

Since the final delivery SMTP server (or, to be more specific, its
corresponding mail storage agent) cannot safely assume that agents
accessing email storage will always be capable of handling the
extensions proposed here, it MAY either downgrade internationalized
emails or specially identify messages that utilize these extensions,
or both. If this is done, the final delivery SMTP server SHOULD
include a mechanism to preserve or recover the original
internationalized forms without information loss to support access by
UTF8SMTP-aware agents.

6. Additional Issues

This section identifies issues that are not covered as part of this
set of specifications, but that will need to be considered as part of
deployment of email address and header internationalization.

6.1. Impact on URIs and IRIs

The mailto: schema defined in [RFC2368] and discussed in the
Internationalized Resource Identifier (IRI) specification [RFC3987]
may need to be modified when this work is completed and standardized.

6.2. Interaction with Delivery Notifications

   The advent of UTF8SMTP will make necessary consideration of the
   interaction with delivery notification mechanisms, including the SMTP
   extension for requesting delivery notifications [RFC3461], and the
   format of delivery notifications [RFC3464]. These issues are
   discussed in a forthcoming document that will update those RFCs as
   needed [EAI-DSN].

6.3. Use of Email Addresses as Identifiers

   There are a number of places in contemporary Internet usage in which
   email addresses are used as identifiers for individuals, including as
   identifiers to Web servers supporting some electronic commerce sites.
   These documents do not address those uses, but it is reasonable to
   expect that some difficulties will be encountered when
   internationalized addresses are first used in those contexts, many of
   which cannot even handle the full range of addresses permitted today.

6.4. Encoded Words, Signed Messages, and Downgrading

   One particular characteristic of the email format is its persistency:
   MUAs are expected to handle messages that were originally sent
   decades ago and not just those delivered seconds ago. As such, MUAs
   and mail filtering software, such as that specified in Sieve
   [RFC3028], will need to continue to accept and decode header fields
   that use the "encoded word" mechanism [RFC2047] to accommodate
   non-ASCII characters in some header fields. While extensions to both
   POP3 and IMAP have been proposed to enable automatic EAI-upgrade --
   including RFC 2047 decoding -- of messages by the POP3 or IMAP
   server, there are message structures and MIME content-types for which
   that cannot be done or where the change would have unacceptable side
   effects.

   For example, message parts that are cryptographically signed, using
   e.g., S/MIME [RFC3851] or Pretty Good Privacy (PGP) [RFC3156], cannot
   be upgraded from the RFC 2047 form to normal UTF-8 characters without
   breaking the signature. Similarly, message parts that are encrypted
   may contain, when decrypted, header fields that use the RFC 2047
   encoding; such messages cannot be 'fully' upgraded without access to
   cryptographic keys.

   Similar issues may arise if signed messages are downgraded in transit
   [EAI-downgrade] and then an attempt is made to upgrade them to the
   original form and then verify the signatures. Even the very subtle
   changes that may result from algorithms to downgrade and then upgrade
   again may be sufficient to invalidate the signatures if they impact

either the primary or MIME bodypart headers. When signatures are
present, downgrading must be performed with extreme care if at all.

6.5. Other Uses of Local Parts

Local parts are sometimes used to construct domain labels, e.g., the
local part "user" in the address user@domain.example could be
converted into a vanity host user.domain.example with its Web space
at <http://user.domain.example> and the catchall addresses
any.thing.goes@user.domain.example.

Such schemes are obviously limited by, among other things, the SMTP
rules for domain names, and will not work without further
restrictions for other local parts such as the <utf8-local-part>
specified in [EAI-UTF8]. Whether this issue is relevant to these
specifications is an open question. It may be simply another case of
the considerable flexibility accorded to delivery MTAs in determining
the mailbox names they will accept and how they are interpreted.

6.6. Non-Standard Encapsulation Formats

Some applications use formats similar to the application/mbox format
defined in [RFC4155] instead of the message/digest RFC 2046, Section
5.1.5 [RFC2046] form to transfer multiple messages as single units.
Insofar as such applications assume that all stored messages use the
message/rfc822 RFC 2046, Section 5.2.1 [RFC2046] format with US-ASCII
headers, they are not ready for the extensions specified in this
series of documents and special measures may be needed to properly
detect and process them.

7. Experimental Targets

In addition to the simple question of whether the model outlined here
can be made to work in a satisfactory way for upgraded systems and
provide adequate protection for un-upgraded ones, we expect that
actually working with the systems will provide answers to two
additional questions: what restrictions such as character lists or
normalization should be placed, if any, on the characters that are
permitted to be used in address local-parts and how useful, in
practice, will downgrading turn out to be given whatever restrictions
and constraints that must be placed upon it.

8. IANA Considerations

This overview description and framework document does not contemplate
any IANA registrations or other actions. Some of the documents in
the group have their own IANA considerations sections and
requirements.

9. Security Considerations

   Any expansion of permitted characters and encoding forms in email
   addresses raises some risks. There have been discussions on so
   called "IDN-spoofing" or "IDN homograph attacks". These attacks
   allow an attacker (or "phisher") to spoof the domain or URLs of
   businesses. The same kind of attack is also possible on the local
   part of internationalized email addresses. It should be noted that
   the proposed fix involving forcing all displayed elements into
   normalized lower-case works for domain names in URLs, but not email
   local parts since those are case sensitive.

   Since email addresses are often transcribed from business cards and
   notes on paper, they are subject to problems arising from confusable
   characters (see [RFC4690]). These problems are somewhat reduced if
   the domain associated with the mailbox is unambiguous and supports a
   relatively small number of mailboxes whose names follow local system
   conventions. They are increased with very large mail systems in
   which users can freely select their own addresses.

   The internationalization of email addresses and headers must not
   leave the Internet less secure than it is without the required
   extensions. The requirements and mechanisms documented in this set
   of specifications do not, in general, raise any new security issues.
   They do require a review of issues associated with confusable
   characters -- a topic that is being explored thoroughly elsewhere
   (see, e.g., [RFC4690]) -- and, potentially, some issues with UTF-8
   normalization, discussed in [RFC3629], and other transformations.
   Normalization and other issues associated with transformations and
   standard forms are also part of the subject of ongoing work discussed
   in [Net-Unicode], in [IDNAbis-BIDI] and elsewhere. Some issues
   specifically related to internationalized addresses and headers are
   discussed in more detail in the other documents in this set.
   However, in particular, caution should be taken that any
   "downgrading" mechanism, or use of downgraded addresses, does not
   inappropriately assume authenticated bindings between the
   internationalized and ASCII addresses.

   The new UTF-8 header and message formats might also raise, or
   aggravate, another known issue. If the model creates new forms of an
   'invalid' or 'malformed' message, then a new email attack is created:
   in an effort to be robust, some or most agents will accept such
   message and interpret them as if they were well-formed. If a filter
   interprets such a message differently than the final MUA, then it may
   be possible to create a message that appears acceptable under the
   filter's interpretation but should be rejected under the
   interpretation given to it by the final MUA. Such attacks already
   exist for existing messages and encoding layers, e.g., invalid MIME

syntax, invalid HTML markup, and invalid coding of particular image types.

Models for the "downgrading" of messages or addresses from UTF-8 form to some ASCII form, including those described in [EAI-downgrade], pose another special problem and risk. Any system that transforms one address or set of mail header fields into another becomes a point at which spoofing attacks can occur and those who wish to spoof messages might be able to do so by imitating a message downgraded from one with a legitimate original address.

In addition, email addresses are used in many contexts other than sending mail, such as for identifiers under various circumstances (see Section 6.3). Each of those contexts will need to be evaluated, in turn, to determine whether the use of non-ASCII forms is appropriate and what particular issues they raise.

This work will clearly impact any systems or mechanisms that are dependent on digital signatures or similar integrity protection for mail headers (see also the discussion in Section 6.4). Many conventional uses of PGP and S/MIME are not affected since they are used to sign body parts but not headers. On the other hand, the developing work on domain keys identified mail (DKIM [DKIM-Charter]) will eventually need to consider this work and vice versa: while this experiment does not propose to address or solve the issues raised by DKIM and other signed header mechanisms, the issues will have to be coordinated and resolved eventually if the two sets of protocols are to co-exist. In addition, to the degree to which email addresses appear in PKI (Public Key Infrastructure) certificates, standards addressing such certificates will need to be upgraded to address these internationalized addresses. Those upgrades will need to address questions of spoofing by look-alikes of the addresses themselves.

10. Acknowledgements

This document, and the related ones, were originally derived from documents by John Klensin and the JET group [Klensin-emailaddr], [JET-IMA]. The work drew inspiration from discussions on the "IMAA" mailing list, sponsored by the Internet Mail Consortium and especially from an early document by Paul Hoffman and Adam Costello [Hoffman-IMAA] that attempted to define an MUA-only solution to the address internationalization problem.

More recent documents have benefited from considerable discussion within the IETF EAI Working Group and especially from suggestions and text provided by Martin Duerst, Frank Ellermann, Philip Guenther, Kari Hurtta, and Alexey Melnikov, and from extended discussions among

the editors and authors of the core documents cited in Section 3:
Harald Alvestrand, Kazunori Fujiwara, Chris Newman, Pete Resnick,
Jiankang Yao, Jeff Yeh, and Yoshiro Yoneya.

Additional comments received during IETF Last Call, including those
from Paul Hoffman and Robert Sparks, were helpful in making the
document more clear and comprehensive.

11. References

11.1. Normative References

[ASCII]   American National Standards Institute (formerly
                        United States of America Standards Institute),
                        "USA Code for Information Interchange",
                        ANSI X3.4-1968, 1968.

                        ANSI X3.4-1968 has been replaced by newer
                        versions with slight modifications, but the 1968
                        version remains definitive for the Internet.

[RFC1652]   Klensin, J., Freed, N., Rose, M., Stefferud, E.,
                        and D. Crocker, "SMTP Service Extension for
                        8bit-MIMEtransport", RFC 1652, July 1994.

[RFC2119]   Bradner, S., "Key words for use in RFCs to
                        Indicate Requirement Levels'", RFC 2119, BCP 14,
                        March 1997.

[RFC2821]   Klensin, J., "Simple Mail Transfer Protocol",
                        RFC 2821, April 2001.

[RFC3490]   Faltstrom, P., Hoffman, P., and A. Costello,
                        "Internationalizing Domain Names in Applications
                        (IDNA)", RFC 3490, March 2003.

[RFC3629]   Yergeau, F., "UTF-8, a transformation format of
                        ISO 10646", STD 63, RFC 3629, November 2003.

11.2. Informative References

[DKIM-Charter]   IETF, "Domain Keys Identified Mail (dkim)",
                        October 2006, <http://www.ietf.org/
                        html.charters/dkim-charter.html>.

[EAI-DSN]   Newman, C., "UTF-8 Delivery and Disposition
                        Notification", Work in Progress, January 2007.

   [EAI-SMTPext]   Yao, J., Ed. and W. Mao, Ed., "SMTP extension
                          for internationalized email address", Work
                          in Progress, June 2007.

   [EAI-UTF8]   Yeh, J., "Internationalized Email Headers", Work
                          in Progress, April 2007.

   [EAI-downgrade]   Yoneya, Y., Ed. and K. Fujiwara, Ed.,
                          "Downgrading mechanism for Internationalized
                          eMail Address (IMA)", Work in Progress,
                          March 2007.

   [EAI-imap]   Resnick, P. and C. Newman, "IMAP Support for
                          UTF-8", Work in Progress, March 2007.

   [EAI-pop]   Newman, C., "POP3 Support for UTF-8", Work
                          in Progress, January 2007.

   [EAI-scenarios]   Alvestrand, H., "UTF-8 Mail: Scenarios", Work
                          in Progress, February 2007.

   [Hoffman-IMAA]   Hoffman, P. and A. Costello, "Internationalizing
                          Mail Addresses in Applications (IMAA)", Work
                          in Progress, October 2003.

   [IDNAbis-BIDI]   Alvestrand, H. and C. Karp, "An IDNA problem in
                          right-to-left scripts", Work in Progress,
                          October 2006.

   [JET-IMA]   Yao, J. and J. Yeh, "Internationalized eMail
                          Address (IMA)", Work in Progress, June 2005.

   [Klensin-emailaddr]   Klensin, J., "Internationalization of Email
                          Addresses", Work in Progress, July 2005.

   [Net-Unicode]   Klensin, J. and M. Padlipsky, "Unicode Format
                          for Network Interchange", Work in Progress,
                          March 2007.

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

   [RFC2046]   Freed, N. and N. Borenstein, "Multipurpose
                          Internet Mail Extensions (MIME) Part Two: Media
                          Types", RFC 2046, November 1996.

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

   [RFC2231]   Freed, N. and K. Moore, "MIME Parameter Value
                          and Encoded Word Extensions:
                          Character Sets, Languages, and Continuations",
                          RFC 2231, November 1997.

   [RFC2368]   Hoffman, P., Masinter, L., and J. Zawinski, "The
                          mailto URL scheme", RFC 2368, July 1998.

   [RFC2822]   Resnick, P., "Internet Message Format",
                          RFC 2822, April 2001.

   [RFC3028]   Showalter, T., "Sieve: A Mail Filtering
                          Language", RFC 3028, January 2001.

   [RFC3156]   Elkins, M., Del Torto, D., Levien, R., and T.
                          Roessler, "MIME Security with OpenPGP",
                          RFC 3156, August 2001.

   [RFC3461]   Moore, K., "Simple Mail Transfer Protocol (SMTP)
                          Service Extension for Delivery Status
                          Notifications (DSNs)", RFC 3461, January 2003.

   [RFC3464]   Moore, K. and G. Vaudreuil, "An Extensible
                          Message Format for Delivery Status
                          Notifications", RFC 3464, January 2003.

   [RFC3851]   Ramsdell, B., "Secure/Multipurpose Internet Mail
                          Extensions (S/MIME) Version 3.1 Message
                          Specification", RFC 3851, July 2004.

   [RFC3987]   Duerst, M. and M. Suignard, "Internationalized
                          Resource Identifiers (IRIs)", RFC 3987,
                          January 2005.

   [RFC4155]   Hall, E., "The application/mbox Media Type",
                          RFC 4155, September 2005.

   [RFC4409]   Gellens, R. and J. Klensin, "Message Submission
                          for Mail", RFC 4409, April 2006.

   [RFC4690]   Klensin, J., Faltstrom, P., Karp, C., and IAB,
                          "Review and Recommendations for
                          Internationalized Domain Names (IDNs)",
                          RFC 4690, September 2006.

Authors' Addresses

   John C Klensin
   1770 Massachusetts Ave, #322
   Cambridge, MA 02140
   USA

   Phone: +1 617 491 5735
   EMail: john-ietf@jck.com

   YangWoo Ko
   ICU
   119 Munjiro
   Yuseong-gu, Daejeon 305-732
   Republic of Korea

   EMail: yw@mrko.pe.kr