Workgroup: openpgp
Internet-Draft:
draft-dkg-openpgp-userid-conventions-00
Published: 25 August 2023
Intended Status: Informational
Expires: 26 February 2024
Authors: D. K. Gillmor
ACLU

OpenPGP User ID Conventions

Abstract

OpenPGP User IDs are UTF-8 strings. Existing documents claim that by conventione, they contain "an RFC 2822 name-addr object", but that's not the case. This document attempts to better describe the actual conventions about User IDs in the deployed OpenPGP ecosystem.

About This Document

This note is to be removed before publishing as an RFC.

The latest revision of this draft can be found at https://dkg.gitlab.io/openpgp-userid-conventions/. Status information for this document may be found at https://datatracker.ietf.org/doc/draft-dkg-openpgp-userid-conventions/.

Discussion of this document takes place on the OpenPGP Working Group mailing list (mailto:openpgp@ietf.org), which is archived at https://mailarchive.ietf.org/arch/browse/openpgp/. Subscribe at https://www.ietf.org/mailman/listinfo/openpgp/.

Source for this draft and an issue tracker can be found at https://gitlab.com/dkg/openpgp-userid-conventions.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

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

This Internet-Draft will expire on 26 February 2024.

Table of Contents

1. Introduction

OpenPGP certificates contain User IDs. An OpenPGP User ID packet contains a simple UTF-8 string. According to [RFC4880] and its successor [I-D.ietf-openpgp-crypto-refresh]:

> By convention, it includes an [RFC2822] mail name-addr

But in practice, this is not what most OpenPGP implementations generate or expect. This document tries to better describe the actual convention used.

1.1. Terminology

The term "OpenPGP Certificate" is used in this document interchangeably with "OpenPGP Transferable Public Key", as defined in Section 10.1 of [I-D.ietf-openpgp-crypto-refresh].

The key words "**MUST**", "**MUST NOT**", "**REQUIRED**", "**SHALL**", "**SHALL NOT**", "**SHOULD**", "**SHOULD NOT**", "**RECOMMENDED**", "**NOT RECOMMENDED**", "**MAY**", and "**OPTIONAL**" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. OpenPGP User ID Conventions

An OpenPGP User ID has no formal constraints other than being a UTF-8 string, but common conventions govern its use in specific contexts.

In the context of sending and receiving signed and encrypted e-mail messages, a User ID typically contains an e-mail address. While [RFC5322] and [RFC6531] describe an addr-spec as it is used e-mail message headers (us-ascii in the former, and Unicode in the latter), the common OpenPGP User ID convention is somewhat simpler, while still permitting extraction of a valid addr-spec. An e-mail-oriented OpenPGP implementation that follows this simpler convention is more likely to be interoperable with Transferable Public Keys found in the wild.

In particular, the common convention for an OpenPGP User ID related to e-mail can be described with the following ABNF (see [RFC5234]), which uses the Unicode-augmented definitions of atext and dot-atom-text found in [RFC6532]:

```
openpgp-addr-spec            = dot-atom-text "@" dot-atom-text

openpgp-email-prefix-char    = atext / specials / SPACE

openpgp-email-wrapped-addr   = *openpgp-uid-prefix-char
                               "<" openpgp-addr-spec ">"

openpgp-email-uid-convention = openpgp-addr-spec /
                               openpgp-email-wrapped-addr
```

Note that any openpgp-addr-spec described in the above sequence is also a valid Unicode addr-spec. The only addr-specs not matched are obsolete forms, or those with CWFS or quoted-string in the local

part, or those with domain literals for the domain part. Using such a non-matching addr-spec in an OpenPGP User ID is likely to lead to interoperability problems.

3. Internationalized Domain Names

FIXME: if a domain name in the openpgp-addr-spec contains non-ASCII characters, will existing implementations accept A-labels? Or should we encourage standardization on U-labels (see [RFC5980])?

4. Example User IDs

4.1. Conventional User IDs

Most tools will work fine with the following User IDs, even though most of them are not technically [RFC5322] name-addr objects:

*Alice Jones <alice@example.org>

*Alice T. Jones <alice@example.org>

*Sean O'Brian <sean@example.org>

*Jörg Schmidt <js@example.org>

*Mr. Ed, the Talking Horse <ed@example.org>

*alice@example.org

4.2. Examples of Atypical User IDs

The following examples are UTF-8 strings that are valid [RFC5322] name-addr objects, but would most likely cause interoperability problems if they were used as an OpenPGP User ID:

4.2.1. RFC 2047 Encoding of non-ASCII Characters

Do not use [RFC2047] encoding:

=?utf-8?Q?J=C3=B6rg?= Schmidt <js@example.org>

4.2.2. quoted-string Parts

Do not use [RFC5322] quoted-string parts:

"Sean O'Brian" <sean@example.org>

4.2.3. FWS

Do not use Folding Whitespace (FWS) (Section 3.2.2 of [RFC5322]).

Alice Jones <alice@example.org>

It's probably not a good idea to include any control character or whitespace character at all, other than " " (SPACE, U+0020) in an OpenPGP User ID. Do not include newline, carriage returns, tab characters, non-folding characters, byte order marks, etc.

Also, leading or trailing whitespace is likely to cause interoperability failures in any context where the User ID must be cleanly parsed.

4.2.4. Comments

Avoid Comments (Section 3.2.2 of [RFC5322]) with the possible exception of a single comment just before the angle-bracket that delimits the openpgp-addr-spec

FIXME: should we discourage comments entirely? See, for example, the litany of complaints at https://dkg.fifthhorseman.net/blog/openpgp-user-id-comments-considered-harmful.html

Alice (the Great) Jones <alice@example.org>

or

(The Great) Alice Jones <alice@example.org>

or

Alice Jones <alice@example.org> (The Great)

5. IANA Considerations

This draft asks IANA to make one change to the OpenPGP protocol group.

In the "Packet Types/Tags registry", update row 13 ("User ID Packet") by adding this document to the "Reference" field.

6. Security Considerations

This document describes widespread conventions about User IDs in the OpenPGP ecosystem.

Generating an OpenPGP certificate with a User ID that does not match these conventions may result in security failures when a peer tries to find a certificate but cannot

7. References

7.1. Normative References

[RFC4880] Callas, J., Donnerhacke, L., Finney, H., Shaw, D., and R. Thayer, "OpenPGP Message Format", RFC 4880, DOI 10.17487/RFC4880, November 2007, <https://www.rfc-editor.org/info/rfc4880>.

[I-D.ietf-openpgp-crypto-refresh] Wouters, P., Huigens, D., Winter, J., and N. Yutaka, "OpenPGP", Work in Progress, Internet-Draft, draft-ietf-openpgp-crypto-refresh-10, 21 June 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-openpgp-crypto-refresh-10>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC5234] Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, DOI 10.17487/RFC5234, January 2008, <https://www.rfc-editor.org/info/rfc5234>.

[RFC6532] Yang, A., Steele, S., and N. Freed, "Internationalized Email Headers", RFC 6532, DOI 10.17487/RFC6532, February 2012, <https://www.rfc-editor.org/info/rfc6532>.

7.2. Informative References

[RFC2822] Resnick, P., Ed., "Internet Message Format", RFC 2822, DOI 10.17487/RFC2822, April 2001, <https://www.rfc-editor.org/info/rfc2822>.

[RFC5322] Resnick, P., Ed., "Internet Message Format", RFC 5322, DOI 10.17487/RFC5322, October 2008, <https://www.rfc-editor.org/info/rfc5322>.

[RFC6531] Yao, J. and W. Mao, "SMTP Extension for Internationalized Email", RFC 6531, DOI 10.17487/RFC6531, February 2012, <https://www.rfc-editor.org/info/rfc6531>.

[RFC5980] Sanda, T., Ed., Fu, X., Jeong, S., Manner, J., and H. Tschofenig, "NSIS Protocol Operation in Mobile

Environments", RFC 5980, DOI 10.17487/RFC5980, March 2011, <https://www.rfc-editor.org/info/rfc5980>.

[RFC2047] Moore, K., "MIME (Multipurpose Internet Mail Extensions) Part Three: Message Header Extensions for Non-ASCII Text", RFC 2047, DOI 10.17487/RFC2047, November 1996, <https://www.rfc-editor.org/info/rfc2047>.

Appendix A. Python Example

The following Python example can be used to parse a conventional OpenPGP User ID:

```
#!/usr/bin/python3

from typing import Optional, Tuple
import re

def openpgp_userid(test: str) -> Optional[Tuple[str, str]]:
    '''Returns a None if `test` is not a conventional User ID.

    if `test` is a conventional User ID, returns a Tuple containing
    the User ID and the embedded e-mail address.'''

    specials = r'[()<>\[\]:;@\\,."]'
    atext = "[-A-Za-z0-9!#$%&'*+/=?^_`{|}~\x80-\U0010ffff]"
    dot_atom_text = atext + r"+(?:\." + atext + "+)*"
    pgp_addr_spec = dot_atom_text + "@" + dot_atom_text
    pgp_uid_prefix_char = "(?:" + atext + "|" + specials + "| )"
    addr_spec_raw = "(?P<addr_spec_raw>" + pgp_addr_spec + ")"
    addr_spec_wrapped = pgp_uid_prefix_char + \
        "*<(?P<addr_spec_wrapped>" + pgp_addr_spec + ")>"
    pgp_uid_convention = "^(?:" + addr_spec_raw + "|" + \
        addr_spec_wrapped + ")$"

    pgp_uid_convention_re = re.compile(pgp_uid_convention,
                                       re.UNICODE)

    m = pgp_uid_convention_re.search(test)
    if m:
        return (m[0], m['addr_spec_wrapped'] or m['addr_spec_raw'])
    else:
        return None
```

Appendix B. Document History

RFC Editor Note: Please delete this section before publication.

B.1. Substantive Changes from Origin to draft-ietf-openpgp-userid-conventions-00

*added positive and negative examples

*added Python implementation

*added FIXME about internationalized domain names

B.2. Origin

This was originally discussed on the mailing list at https://mailarchive.ietf.org/arch/msg/openpgp/wNo27-0STfGR9JZSlC7s6OYOJkI and was formulated as a patch to the OpenPGP specification at https://mailarchive.ietf.org/arch/msg/openpgp/wNo27-0STfGR9JZSlC7s6OYOJkI .

Author's Address

Daniel Kahn Gillmor
ACLU

Email: dkg@fifthhorseman.net