Structured Email
Internet-Draft
Intended status: Standards Track
Expires: 3 September 2026

P. Tao
Apple Inc.

2 March 2026

Structured Quoted Content
draft-tao-sml-structured-quoted-content-00

Abstract

This document describes a machine-readable format for conveying quoted content in email messages. This can be used when replying to or forwarding an email message.

Structured quoted content is expected to be used in conjunction with conventional, human-readable quote formatting. They are based on the forthcoming "structured email" specification defined in [I-D.ietf-sml-structured-email-03] and related drafts.

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://ptao.github.io/id-structured-quoted-content/draft-tao-sml-structured-quoted-content.html. Status information for this document may be found at https://datatracker.ietf.org/doc/draft-tao-sml-structured-quoted-content/.

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

Source for this draft and an issue tracker can be found at https://github.com/ptao/id-structured-quoted-content.

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/.

This Internet-Draft will expire on 3 September 2026.

Copyright Notice

Table of Contents

1. Introduction

Email messages often content quoted sections of earlier email messages, typically as part of a reply or a forwarded message. The presentation is entirely determined by the sending MUA. Though various conventions have evolved around this quoting, the behavior is still largely unstandardized, and behavior differs widely between MUAs.

Some notable differences include:

* Where newly authored content may be inserted relative to the quoted content.
  - Some MUAs allow inserting new content only above the quoted message, while some allow inserting new content inline between blocks of quoted text.
* The formatting used, whether HTML or plaintext, to denote the quoted section.
* Whether some set of headers are also quoted, or whether an attribution line or preamble is inserted.
  - Whether the headers, attribution line, or preamble are localized into the sender’s device language.
* Whether attachments are included in the quoted content.

Due to the above, an email chain in a user’s inbox may involve a wide mix of differing quoting styles, which can be visually disorienting.

The goal of this draft is to introduce a standardized way to structure quoted content in message replies and forwarded messages. This allows the sending MUA to annotate the various components of the quoted content in such a way that the receiving MUA can then apply a consistent presentation across all messages.

2. Conventions and Definitions

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.

3. Structured Email Format

3.1. Structured Data Type

All quoted content will be transmitted as "Partial representation" as defined by https://datatracker.ietf.org/doc/html/draft-ietf-sml-structured-email-05#name-partial-representation.

3.2. Vocabularies

Structured Quoted Content uses the SchemaOrg vocabulary, as well as a custom vocabulary to be defined at *INSERT LINK*.

4. JSON-LD Objects

Structured Quoted Content uses the following object types.

EmailMessage

* vocabulary: schema.org and sml.draft.iana.org
* id format: "mid:<message-id header>"
* The following properties can be set: sender, toRecipient, dateSent, dateReceived, about (subject)
    * Each property should have its own id

QuotedContent

* vocabulary: sml.draft.iana.org
* id format: "quoted-content:<UUID>"
* Properties:
    * message
        * References EmailMessage which this item is quoting
    * headerBlock
        * References QuotedHeaderBlock
        * optional
    * attribution
        * References QuoteAttribution
        * optional
    * content
        * References one or more TextContent or QuotedContent

QuotedHeaderBlock

* vocabulary: sml.draft.iana.org
* id format: "quoted-header-block:<UUID>"
* Properties:
    * header
        * References one or more QuotedHeader

QuotedHeader

* vocabulary: sml.draft.iana.org
* id format: "quoted-header:<UUID>"
* Properties:
  * label
    * Text label for the header
  * value
    * References one of the properties of an EmailMessage

QuoteAttribution

* vocabulary: sml.draft.iana.org
* id format: "quote-attribution:<UUID>"
* Properties:
  * date
    * References either the dateSent or dateReceived property of an EmailMessage
  * from
    * References the sender property of an Emailmessage

TextContent

* vocabulary: sml.draft.iana.org
* id format: "text-content:<UUID>"
* Properties:

Note, TextContent has no properties.

TODO: There should be a type for a preamble for forwarding.

5. HTML Format

Each quoted block will be wrapped in an HTML element (typically div or span) with a data-source-id and a data-id property. The data-source-id is a cid: URI referencing the MIME part which contains the structured data definitions. The data-id is the id of the matching element.

6. Plaintext Format

TODO Plaintext

7. Backwards Compatibility

In a long reply chain, not all MUAs involved may be SML-aware. This results in two types of mixed content: 1. A MUA which is compliant with Structured Quoted Content quoting a message which does not have Structured Quoted Content. 1. A MUA which is not aware of Structured Quoted Content quoting a message which has Structured Quoted Content.

In the first case, the MUA SHOULD NOT try to retroactively determine the quoted content structure of the email it is quoting. In other words, it SHOULD NOT attempt to parse the HTML or plain-text formatting of that message in order to build a structured representation. Instead, any structured quoted content that it generates should treat the quoted message as one having no quoted content itself.

The MUA in the second situation is by definition not bound to any behavior specified in this draft. It may likely do one or both of the following: 1. Modify the HTML or plaintext in such a way that the existing structured content properties no longer behave as expected. 2. Modify the MIME structure in such a way that the structured quoted content is no longer valid. For example it may: * Change the MIME structure in such a way that the message is no longer classified as a "partial representation". * Remove or modify the application/ld+json part.

In general, it can be assumed that for a message which contains Structured Quoted Content, the Structured Quoted Content represents the most recent contiguous set of messages. This set of messages may be a subset of all the messages which actually comprise the quoted reply chain. Such a case occurs when the oldest message represented in the Structured Quoted Content has quoted a message sent by a MUA which was not aware of Structured Quoted Content.

When rendering such a message, only the first contiguous set of Structured Quoted Content should be treated as Structured Quoted Content. Any HTML or JSON-LD is found which hints at the presence of Structured Quoted Content further back in the reply chain, but which has not made it to the receiving MUA intact, should be ignored and that portion of HTML should be rendered as not containing Structured Quoted Content.

8. MUA Guidance

TODO: List potential usages/purposes/benefits

Some examples:

* More reliable quoted content parsing/recognition by all compliant MUAs

* Allow for consistent and localized quoting style (determined by receiver rather than sender)

* Better condensed display

* Clearer attribution of deeply nested content

* Clearer attribution for inline quoting

9. Security Considerations

TODO Security

10. IANA Considerations

This document has no IANA actions.

11. Normative References

[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/rfc/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/rfc/rfc8174>.

Acknowledgments

TODO acknowledge.

Author’s Address

Phillip Tao
Apple Inc.
Email: ptao@apple.com