

# Message format

---

The Internet e-mail message format is defined in [RFC 5322](#) and a series of [RFCs](#), [RFC 2045](#) through [RFC 2049](#), collectively called, [Multipurpose Internet Mail Extensions](#), or *MIME*. Although as of July 13, 2005, [RFC 2822](#) is technically a proposed [IETF](#) standard and the MIME RFCs are draft IETF standards,<sup>[24]</sup> these documents are the standards for the format of Internet e-mail. Prior to the introduction of [RFC 2822](#) in 2001, the format described by [RFC 822](#) was the standard for Internet e-mail for nearly 20 years; it is still the official IETF standard. The IETF reserved the numbers 5321 and 5322 for the updated versions of [RFC 2821](#) (SMTP) and [RFC 2822](#), as it previously did with [RFC 821](#) and [RFC 822](#), honoring the extreme importance of these two RFCs. [RFC 822](#) was published in 1982 and based on the earlier [RFC 733](#) (see <sup>[25]</sup>).

Internet e-mail messages consist of two major sections:

- *Header* — Structured into [fields](#) such as summary, sender, receiver, and other information about the e-mail.
- *Body* — The message itself as unstructured text; sometimes containing a [signature block](#) at the end. This is exactly the same as the body of a regular letter.

The header is separated from the body by a blank line.

## [\[edit\]](#) Message header

Each message has exactly one [header](#), which is structured into [fields](#). Each field has a name and a value. [RFC 5322](#) specifies the precise syntax.

Informally, each line of text in the header that begins with a [printable character](#) begins a separate field. The field name starts in the first character of the line and ends before the separator character ":". The separator is then followed by the field value (the "body" of the field). The value is continued onto subsequent lines if those lines have a space or tab as their first character. Field names and values are restricted to 7-bit [ASCII](#) characters. Non-ASCII values may be represented using MIME [encoded words](#).

## [\[edit\]](#) Header fields



This section **needs additional citations for verification**.

Please help [improve this article](#) by adding [reliable references](#). Unsourced material may be [challenged](#) and [removed](#). *(August 2010)*

The message header should include at least the following fields:

- *From*: The [e-mail address](#), and optionally the name of the author(s). In many e-mail clients not changeable except through changing account settings.
- *To*: The e-mail address(es), and optionally name(s) of the message's recipient(s). Indicates primary recipients (multiple allowed), for secondary recipients see Cc: and Bcc: below.
- *Subject*: A brief summary of the topic of the message. [Certain abbreviations](#) are commonly used in the subject, including ["RE:"](#) and ["FW:"](#).
- *Date*: The local time and date when the message was written. Like the *From*: field, many email clients fill this in automatically when sending. The recipient's client may then display the time in the format and time zone local to him/her.
- *Message-ID*: Also an automatically generated field; used to prevent multiple delivery and for reference in In-Reply-To: (see below).

Note that the *To*: field is not necessarily related to the addresses to which the message is delivered. The actual delivery list is supplied separately to the transport protocol, [SMTP](#), which may or may not originally have been extracted from the header content. The "To:" field is similar to the addressing at the top of a conventional letter which is delivered according to the address on the outer envelope. Also note that the "From:" field does not have to be the real sender of the e-mail message. One reason is that it is very easy to fake the "From:" field and let a message seem to be from any mail address. It is possible to [digitally sign](#) e-mail, which is much harder to fake, but such signatures require extra programming and often external programs to verify. Some ISPs do not relay e-mail claiming to come from a domain not hosted by them, but very few (if any) check to make sure that the person or even e-mail address named in the "From:" field is the one associated with the connection. Some ISPs apply [e-mail authentication](#) systems to e-mail being sent through their MTA to allow other MTAs to detect forged spam that might appear to come from them.

[RFC 3864](#) describes registration procedures for message header fields at the [IANA](#); it provides for [permanent](#) and [provisional](#) message header field names, including also fields defined for MIME, netnews, and http, and referencing relevant RFCs. Common header fields for email include:

- *Bcc*: [Blind Carbon Copy](#); addresses added to the SMTP delivery list but not (usually) listed in the message data, remaining invisible to other recipients.

- Cc: [Carbon copy](#); Many e-mail clients will mark e-mail in your inbox differently depending on whether you are in the To: or Cc: list.
- [Content-Type](#): Information about how the message is to be displayed, usually a [MIME](#) type.
- In-Reply-To: [Message-ID](#) of the message that this is a reply to. Used to link related messages together.
- Precedence: commonly with values "bulk", "junk", or "list"; used to indicate that automated "vacation" or "out of office" responses should not be returned for this mail, e.g. to prevent vacation notices from being sent to all other subscribers of a mailinglist.
- Received: Tracking information generated by mail servers that have previously handled a message, in reverse order (last handler first).
- References: [Message-ID](#) of the message that this is a reply to, and the message-id of the message the previous was reply a reply to, etc.
- Reply-To: Address that should be used to reply to the message.
- Sender: Address of the actual sender acting on behalf of the author listed in the From: field (secretary, list manager, etc.).

## [\[edit\]](#)Message body



This section **needs additional [citations for verification](#)**.

Please help [improve this article](#) by adding [reliable references](#). Unsourced material may be [challenged](#) and [removed](#). *(November 2007)*

## [\[edit\]](#)Content encoding

E-mail was originally designed for 7-bit [ASCII](#).<sup>[26]</sup> Much e-mail software is [8-bit clean](#) but must assume it will communicate with 7-bit servers and mail readers.

The [MIME](#) standard introduced character set specifiers and two content transfer encodings to enable transmission of non-ASCII data: [quoted printable](#) for mostly 7 bit content with a few characters outside that range and [base64](#) for arbitrary binary data.

The [8BITMIME](#) extension was introduced to allow transmission of mail without the need for these encodings but many [mail transport agents](#) still do not support it fully. In some countries, several encoding schemes coexist; as the result, by default, the message in a non-Latin alphabet language appears in non-readable form (the only exception is coincidence, when the sender and receiver use the same encoding scheme). Therefore, for international [character sets](#), [Unicode](#) is growing in popularity.

## [\[edit\]](#)Plain text and HTML

Most modern graphic [e-mail clients](#) allow the use of either [plain text](#) or [HTML](#) for the message body at the option of the user. [HTML e-mail](#) messages often include an automatically-generated plain text copy as well, for compatibility reasons.

Advantages of HTML include the ability to include in-line links and images, set apart previous messages in [block quotes](#), wrap naturally on any display, use emphasis such as [underlines](#) and [italics](#), and change [font](#) styles. Disadvantages include the increased size of the email, privacy concerns about [web bugs](#), abuse of HTML email as a vector for [phishing](#) attacks and the spread of [malicious software](#).<sup>[27]</sup>

Some web based [Mailing lists](#) recommend that all posts be made in plain-text<sup>[28][29]</sup> for all the above reasons, but also because they have a significant number of readers using [text-based e-mail clients](#) such as [Mutt](#).

Some [Microsoft e-mail clients](#) allow rich formatting using [RTF](#), but unless the recipient is guaranteed to have a compatible [e-mail client](#) this should be avoided.<sup>[30]</sup>

In order to ensure that HTML sent in an email is rendered properly by the recipient's client software, an additional header must be specified when sending: "Content-type: text/html". Most email programs send this header automatically.