

Network Working Group  
Internet-Draft  
Updates: [2822](#) (if approved)  
Obsoletes: [1036](#) (if approved)  
Expires: September 25, 2003

D. Kohn  
Skymoon Ventures  
March 27, 2003

**News Article Format**  
**draft-kohn-news-article-03.txt**

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#).

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 September 25, 2003.

Copyright Notice

Copyright (C) The Internet Society (2003). All Rights Reserved.

Abstract

This document defines the format of network news articles.

Network news articles resemble mail messages but are broadcast to potentially large audiences, using a flooding algorithm that propagates one copy to each interested host (or group thereof), typically stores only one copy per host, and does not require any central administration or systematic registration of interested users. Network news originated as the medium of communication for Usenet, circa 1980. Since then Usenet has grown explosively, and many Internet sites participate in it. In addition, the news technology is now in widespread use for other purposes, on the Internet and



elsewhere.

This document defines the format of network news articles in the context of the Internet Message Format, and adds Multipurpose Internet Mail Extensions (MIME) support for multimedia and internationalized message bodies.

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">3</a>
<a href="#">1.1</a>	Scope . . . . .	<a href="#">3</a>
<a href="#">1.2</a>	Requirements Notation . . . . .	<a href="#">4</a>
<a href="#">1.3</a>	Errata . . . . .	<a href="#">4</a>
<a href="#">1.4</a>	Syntax Notation . . . . .	<a href="#">4</a>
<a href="#">1.5</a>	Definitions . . . . .	<a href="#">4</a>
<a href="#">1.6</a>	Structure of This Document . . . . .	<a href="#">4</a>
<a href="#">2.</a>	Format . . . . .	<a href="#">5</a>
<a href="#">2.1</a>	Base . . . . .	<a href="#">5</a>
<a href="#">2.2</a>	MIME Conformance . . . . .	<a href="#">5</a>
<a href="#">2.3</a>	Other MIME Support . . . . .	<a href="#">5</a>
<a href="#">3.</a>	Headers . . . . .	<a href="#">6</a>
<a href="#">3.1</a>	New Internet Message Format Headers . . . . .	<a href="#">6</a>
<a href="#">3.2</a>	Mandatory Headers . . . . .	<a href="#">6</a>
<a href="#">3.3</a>	Date . . . . .	<a href="#">6</a>
<a href="#">3.4</a>	Subject . . . . .	<a href="#">7</a>
<a href="#">3.5</a>	Message-ID . . . . .	<a href="#">7</a>
<a href="#">3.6</a>	News Headers . . . . .	<a href="#">8</a>
<a href="#">3.6.1</a>	Newsgroups . . . . .	<a href="#">8</a>
<a href="#">3.6.2</a>	Path . . . . .	<a href="#">9</a>
<a href="#">3.6.3</a>	Followup-To . . . . .	<a href="#">9</a>
<a href="#">3.6.4</a>	Expires . . . . .	<a href="#">10</a>
<a href="#">3.6.5</a>	Control . . . . .	<a href="#">10</a>
<a href="#">3.6.6</a>	Distribution . . . . .	<a href="#">10</a>
<a href="#">3.6.7</a>	Summary . . . . .	<a href="#">10</a>
<a href="#">3.6.8</a>	Approved . . . . .	<a href="#">11</a>
<a href="#">3.6.9</a>	Organization . . . . .	<a href="#">11</a>
<a href="#">3.6.10</a>	Xref . . . . .	<a href="#">11</a>
<a href="#">3.6.11</a>	Supersedes . . . . .	<a href="#">11</a>
<a href="#">3.7</a>	Other Message Headers . . . . .	<a href="#">11</a>
<a href="#">4.</a>	Internationalization Considerations . . . . .	<a href="#">12</a>
<a href="#">5.</a>	Security Considerations . . . . .	<a href="#">13</a>
	Normative References . . . . .	<a href="#">14</a>
	Informative References . . . . .	<a href="#">15</a>
	Author's Address . . . . .	<a href="#">16</a>
<a href="#">A.</a>	Acknowledgements . . . . .	<a href="#">17</a>
	Intellectual Property and Copyright Statements . . . . .	<a href="#">18</a>



## **1. Introduction**

### **1.1 Scope**

"Netnews" is a set of protocols for generating, storing and retrieving news "articles" (which use the Internet Message Format) and for exchanging them among a readership which is potentially widely distributed. It is organized around "newsgroups", with the expectation that each reader will be able to see all articles posted to each newsgroup in which she participates. These protocols most commonly use a flooding algorithm which propagates copies throughout a network of participating servers. Typically, only one copy is stored per server, and each server makes it available on demand to readers able to access that server.

This is the first of four documents that obsolete [RFC 1036](#). This document focuses on the syntax and semantics of network news articles. [\[useprot\]](#) is also a standards-track document, and describes the protocol issues of network news articles, independent of transmission protocols such as NNTP [\[RFC0977\]](#) and IMAP [\[RFC3501\]](#). An informational document, [\[useimpl\]](#), describes implementation recommendations to improve interoperability and usability. The fourth document, [\[useint\]](#), an experimental standard, specifies internationalization of message headers.

The predecessor to this document [\[RFC1036\]](#) said that: "In any situation where this standard conflicts with the Internet [email standard, the latter] should be considered correct and this standard in error." The basic philosophy of this document follows that previous convention, so as to standardize news article syntax firmly in the context of Internet Message Format syntax. In the context of the Internet messaging architecture, different protocols (such as IMAP, POP3 [\[RFC1939\]](#), NNTP and SMTP [\[RFC2821\]](#)) are seen as alternative ways of moving around the same content. That content is the Internet Message Format as specified by [\[RFC2822\]](#), including optional enhancements such as MIME [\[RFC2049\]](#). A user should be able to ingest an article via NNTP, read it via IMAP, forward it off to someone else via SMTP and have them read it via POP3 all without having to alter the content.

This document uses a cite by reference methodology, rather than trying to repeat the contents of other standards, which could otherwise result in subtle differences and interoperability challenges. Although this document is as a result rather short, it requires complete understanding and implementation of the normative references to be compliant.



## **1.2 Requirements Notation**

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 [[RFC2119](#)].

## **1.3 Errata**

The RFC Editor makes available errata for RFCs at [[errata](#)]. Implementers should review that page for normative references, noting in particular that errata currently exist for [[RFC2046](#)].

## **1.4 Syntax Notation**

Headers defined in this specification use the Augmented Backus-Naur Form (ABNF) notation (including the Core Rules) specified in [[RFC2234](#)] and many constructs (specifically <date-time>, <mailbox-list>, <obs-zone>, and <unstructured>) defined in [[RFC2822](#)]. [Section 3.5](#) updates the [[RFC2822](#)] definition of <msg-id>.

## **1.5 Definitions**

Add definitions here for at least article, user agent, injector, moderator, server, and gateway. Agent refers generically to all roles.

## **1.6 Structure of This Document**

[Section 2](#) defines the format of news articles. [Section 3](#) defines some additional headers necessary for the netnews environment.





## **2. Format**

### **2.1 Base**

News articles MUST conform to the "legal to generate syntax" specified in [Section 3 of \[RFC2822\]](#). News agents SHOULD also support the obsolete syntax specified in [Section 4 of \[RFC2822\]](#), particularly to support old news messages and gatewayed obsolete mail messages, but they MUST NOT generate such syntax.

### **2.2 MIME Conformance**

User agents MUST meet the definition of MIME-conformance in [\[RFC2049\]](#). This level of MIME Conformance provides support for internationalization and multimedia in message bodies, and for receipt (but not generation) of internationalized headers. Generation of internationalized message headers is specified by [\[useint\]](#).

### **2.3 Other MIME Support**

User agents conformant with this document SHOULD support receipt (and automatic reassembly) of message/partial MIME messages, as specified in [Section 5.2.2 of \[RFC2046\]](#) and MAY support generation of message/partial articles for excessively large articles.

User agents SHOULD send regular paragraph text as "text/plain; format=flowed" as specified in [\[RFC2646\]](#) and SHOULD preserve flowed text (including quoting) when replying or forwarding, as described in that specification.

User agents SHOULD support on receipt and MAY generate MIME extension header fields, including but not limited to Content-Disposition [\[RFC2183\]](#) and Content-Language [\[RFC3282\]](#).



### **3. Headers**

#### **3.1 New Internet Message Format Headers**

The following header fields extend the fields defined in [section 3.6 of \[RFC2822\]](#) as follows:

```
fields          =/ *( newsgroups /
                        path /
                        followup-to /
                        expires /
                        control /
                        distribution /
                        summary /
                        approved /
                        organization /
                        xref /
                        supersedes )
```

Each of these headers may occur at most once in a news article.

#### **3.2 Mandatory Headers**

Each news article conformant with this specification MUST have exactly one of each of the following headers: From, Subject, Message-ID, Date, Newsgroups, and Path.

From is exactly as specified in [Section 3.6.2 of \[RFC2822\]](#). Date and Subject are fully conformant with [\[RFC2822\]](#), though with extra detail in [Section 3.3](#) and [Section 3.4](#), respectively.

In [Section 3.5](#), this document updates the <msg-id> construct from [\[RFC2822\]](#) so as to ensure that Internet Message Format Message-IDs are usable in widely deployed news software.

Following [\[RFC2822\]](#) syntax, the headers defined in this document do not require a space between the ":" and the field's contents. (E.g., "Subject:Hello World" is acceptable, as opposed to requiring "Subject: Hello World".) To be compliant with this specification, news agents MUST support 0 or more spaces between the colon and the field's contents. However, to maximize compatibility with the installed base of news agents, implementers SHOULD use exactly one space.

#### **3.3 Date**

The Date header is the same as that specified in Sections [3.3](#) and 3.6.1 of [\[RFC2822\]](#). However, the use of "GMT" as a time zone, which



is part of <obs-zone>, is widespread in news articles today. Therefore, agents MUST accept, but MUST NOT generate, <date-time> constructs where <obs-zone>="GMT". (As stated in [Section 2.1](#), support for <obs-zone> would otherwise have been SHOULD accept, MUST NOT generate.) Note that these requirements apply wherever <date-time> is used, including Expires in [Section 3.6.4](#).

### [3.4](#) Subject

The Subject header contains a short string identifying the topic of the message. [Section 3.6.5 of \[RFC2822\]](#) says:

When used in a reply, the field body MAY start with the string "Re: " (from the Latin "res", in the matter of) followed by the contents of the "Subject:" field body of the original message. If this is done, only one instance of the literal string "Re: " ought to be used since use of other strings or more than one instance can lead to undesirable consequences.

Because of the importance of threading in news, that MAY is amplified to a SHOULD: Follow-ups to an article SHOULD begin with the subject "Re: " followed by the original subject of the referenced article (with that original subject stripped of any starting "Re: ").

User agents MAY remove strings that are known to be used erroneously as back-reference (such as "Re(2): ", "Re:", "RE: ", or "Sv: ") from the beginning of the Subject field body when composing the subject of a followup, and add a correct back-reference in front of the result.

User agents replying to a message MUST NOT use any other string except "Re: " as a back reference. Specifically, a translation of "Re: " into a local language or usage MUST NOT be used.

User agents MAY present to the user a translation of "Re: ", but this MUST only be an artifact of the user interface and MUST NOT be part of the actual news article.

### [3.5](#) Message-ID

The "Message-ID:" field contains a single unique message identifier. This is the only header field definition that updates [\[RFC2822\]](#). The ABNF should be used as below, but the requirements and descriptive text from [Section 3.6.4 of \[RFC2822\]](#) still apply.



```
message-id      = "Message-ID:" msg-id CRLF
msg-id          = [CFWS] msg-id-core [CFWS]
msg-id-core     = "<" id-left "@" id-right ">"
                  ; maximum length is 250 octets
id-left         = dot-atom-text / no-fold-quote / obs-id-left
id-right        = dot-atom-text / no-fold-literal / obs-id-right
no-fold-quote   = DQUOTE *( qtext / no-space-qp ) DQUOTE
no-fold-literal = "[" *( htext / no-space-qp ) "]"
no-space-qp     = ( "\" ptext ) / obs-qp
ptext           = %d33-61 / ; Printable characters excluding ">"
                  %d63-126 /
                  obs-text
htext           = HEXDIG / ; hexadecimal digits, case-insensitive
                  "." / ; IPv4 separator
                  ":" / ; IPv6 separator
```

Although compliant agents MUST support [CFWS] between the "Message-ID:" and the <msg-id-core>, implementers SHOULD generate exactly one space there, to maximize compatibility with the installed base.

Note that this updated ABNF applies wherever <msg-id> is used, including the In-Reply-To and References headers mentioned in [Section 3.7](#).

## [3.6](#) News Headers

### [3.6.1](#) Newsgroups

The Newsgroups header specifies to which newsgroup(s) the article is posted.





```
newsgroups      = "Newsgroups:" newsgroup-list CRLF
newsgroup-list  = [FWS] newsgroup-name
                  *( "," [FWS] newsgroup-name ) [FWS]
newsgroup-name  = component *( "." component ) ; 71 character max
component       = plain-component
plain-component = component-start *29component-rest
component-start = ALPHA / DIGIT
component-rest  = ALPHA / DIGIT / "+" / "-" / "_"
```

A newsgroup name consists of one or more components separated by periods, with no more than 71 characters total. Each component consists of less than 30 or less letters and digits.

### [3.6.2](#) Path

The Path header's content indicates which relayers the article has already visited, so that unnecessary redundant transmission can be avoided.

```
path            = "Path:" [FWS]
                  *( path-host [FWS] path-delimiter [FWS] )
                  path-host [FWS] CRLF
path-host       = ( ALPHA / DIGIT )
                  *( ALPHA / DIGIT / "-" / "." / ":" / "_" )
path-delimiter  = "!"
```

### [3.6.3](#) Followup-To

The Followup-To header specifies to which newsgroup(s) followups should be posted.

```
followup-to     = "Followup-To:" ( newsgroup-list / poster-text )
                  CRLF
poster-text     = [FWS] %d112.111.115.116.101.114 [FWS]
                  ; "poster" in lower-case
```

The syntax is the same as that of the Newsgroups content, with the exception that the magic word "poster" (which is always lowercase) means that followups should be mailed to the article's reply address



rather than posted.

#### **3.6.4 Expires**

The Expires header specifies a date and time when the article is deemed to be no longer useful and could usefully be removed ("expired").

expires = "Expires:" date-time CRLF

#### **3.6.5 Control**

The Control header marks the article as a control message, and specifies the desired actions (additional to the usual ones of storing and/or relaying the article). The verb indicates what action should be taken, and the argument(s) (if any) supply details. In some cases, the body of the article may also contain details. Control messages are further specified in the companion document, [[useprot](#)].

control = "Control:" verb \*( FWS argument ) CRLF

An article with a Control header MUST NOT have a Supersedes header.

#### **3.6.6 Distribution**

The Distribution header specifies geographic or organizational limits on an article's propagation.

distribution = "Distribution:" dist-name \*( "," dist-name ) CRLF

dist-name = [FWS] ALPHA / DIGIT  
\*( ALPHA / DIGIT / "+" / "-" / "\_" ) [FWS]

"All" MUST NOT be used as a distribution-name. Distribution-names SHOULD contain at least three characters, except when they are two-letter country names as in [[ISO.3166.1988](#)]. Distribution-names are case-insensitive (i.e. "US", "Us", "uS", and "us" all specify the same distribution).

#### **3.6.7 Summary**

The Summary header is a short phrase summarizing the article's content.

summary = "Summary:" unstructured CRLF



### **3.6.8 Approved**

The Approved header indicates the mailing addresses (and possibly the full names) of the persons or entities approving the article for posting.

approved = "Approved:" mailbox-list CRLF

### **3.6.9 Organization**

The Organization header is a short phrase identifying the poster's organization.

organization = "Organization:" unstructured CRLF

There is no "s" in Organization.

### **3.6.10 Xref**

The Xref header indicates where an article was filed by the last relay to process it.

xref = "Xref:" [CFWS] path-host  
1\*( CFWS location ) [CFWS]

location = newsgroup-name ":" 1\*16DIGIT

### **3.6.11 Supersedes**

The Supersedes header specifies articles to be cancelled.

supersedes = "Supersedes:" 1\*( [FWS] msg-id-core ) CRLF

There is no "c" in Supersedes.

## **3.7 Other Message Headers**

The headers Reply-To, Sender, Comments, and Keywords are often used in news articles and have the identical meaning as that specified in [\[RFC2822\]](#). References and In-Reply-To are also regularly used in news articles and have the same meaning as that specified in [\[RFC2822\]](#), except that they use the updated <msg-id> construct defined in [Section 3.5](#).



#### **4. Internationalization Considerations**

Internationalization of news article bodies is provided using MIME mechanisms in [Section 2.2](#). Generation of internationalized message headers is not specified in this document, and is instead specified in the experimental standard, [[useint](#)].

## 5. Security Considerations

The news article format specified in this document does not provide any security services, such as confidentiality, authentication of sender, or non-forgery. Instead, such services need to be layered above, using such protocols as S/MIME [[RFC2633](#)] or PGP/MIME [[RFC3156](#)], or below, using secure versions of news transport protocols. Additionally, several currently non-standardized protocols [[PGPVERIFY](#)] will hopefully be standardized in the near future.

Message-IDs (see [Section 3.5](#)) in news are required to be unique; articles are refused (in server-to-server transfer) if the ID has already been seen. So if you can predict the ID of a message, you can preempt it by posting a message (possibly to a quite different group) with the same ID, stopping your target message from propagating. Agents that generate message-ids for news articles SHOULD ensure that they are unpredictable.





## Normative References

- [RFC2046] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types", [RFC 2046](#), November 1996.
- [RFC2049] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part Five: Conformance Criteria and Examples", [RFC 2049](#), November 1996.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC2183] Troost, R., Dorner, S. and K. Moore, "Communicating Presentation Information in Internet Messages: The Content-Disposition Header Field", [RFC 2183](#), August 1997.
- [RFC2234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", [RFC 2234](#), November 1997.
- [RFC2646] Gellens, R., "The Text/Plain Format Parameter", [RFC 2646](#), August 1999.
- [RFC2822] Resnick, P., "Internet Message Format", [RFC 2822](#), April 2001.
- [RFC3282] Alvestrand, H., "Content Language Headers", [RFC 3282](#), May 2002.



## Informative References

- [ISO.3166.1988]  
International Organization for Standardization, "Codes for the representation of names of countries, 3rd edition", ISO Standard 3166, August 1988.
- [PGPVERIFY]  
Lawrence, D., "PGPverify <<ftp://ftp.isc.org/pub/pgpcontrol/README.html>>", June 1999.
- [RFC0977] Kantor, B. and P. Lapsley, "Network News Transfer Protocol", [RFC 977](#), February 1986.
- [RFC1036] Horton, M. and R. Adams, "Standard for interchange of USENET messages", [RFC 1036](#), December 1987.
- [RFC1939] Myers, J. and M. Rose, "Post Office Protocol - Version 3", STD 53, [RFC 1939](#), May 1996.
- [RFC2633] Ramsdell, B., "S/MIME Version 3 Message Specification", [RFC 2633](#), June 1999.
- [RFC2821] Klensin, J., "Simple Mail Transfer Protocol", [RFC 2821](#), April 2001.
- [RFC3156] Elkins, M., Del Torto, D., Levien, R. and T. Roessler, "MIME Security with OpenPGP", [RFC 3156](#), August 2001.
- [RFC3501] Crispin, M., "INTERNET MESSAGE ACCESS PROTOCOL - VERSION 4rev1", [RFC 3501](#), March 2003.
- [errata] "RFC Editor Errata <<http://www.rfc-editor.org/errata.html>>".
- [useimpl] "Usenet Implementation Recommendations (work in progress)".
- [useint] "Usenet Internationalization (work in progress)".
- [useprot] "Usenet Protocol (work in progress)".



Author's Address

Dan Kohn  
Skymoon Ventures  
3045 Park Boulevard  
Palo Alto, California 94306  
USA

Phone: +1-650-327-2600

EMail: dan@dankohn.com

URI: <http://www.dankohn.com/>

**[Appendix A](#). Acknowledgements**

Comments and/or text were provided by Mark Crispin, Claus Faerber, Ned Freed, Andrew Gierth, Tony Hansen, Paul Hoffman, Simon Josefsson, Bruce Lilly, Charles Lindsey, Ken Murchison, Pete Resnick, and Henry Spencer.

## Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Information on the IETF's procedures with respect to rights in standards-track and standards-related documentation can be found in [BCP-11](#). Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementors or users of this specification can be obtained from the IETF Secretariat.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this standard. Please address the information to the IETF Executive Director.

## Full Copyright Statement

Copyright (C) The Internet Society (2003). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assignees.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION





HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF  
MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

#### Acknowledgement

Funding for the RFC Editor function is currently provided by the  
Internet Society.