Network Working Group
Request for Comments: 2912
Category: Standards Track

G. Klyne
Content Technologies

September 2000

Indicating Media Features for MIME Content

Status of this Memo

Copyright Notice

Abstract

In "A Syntax for Describing Media Feature Sets", an expression format is presented for describing media feature capabilities using simple media feature tags.

This memo defines a Multipurpose Internet Mail Extensions (MIME) 'Content-features:' header that can be used to annotate a MIME message part using this expression format, and indicates some ways it might be used.

Table of Contents

1. Introduction

In "A Syntax for Describing Media Feature Sets" [1], an expression format is presented for describing media feature capabilities as a combination of simple media feature tags, registered according to "Media Feature Tag Registration Procedure" [2]. This provides a format for message handling agents to describe the media feature content of messages that they can handle.

This memo defines a MIME 'Content-features:' header that can be used to annotate a MIME message part using these feature expressions. This header provides a means to indicate media-related features of message content that go beyond the MIME content type.

Consideration is also given to how it may be used to present message media content information that is problematic to express within the basic MIME framework.

1.1 Terminology and document conventions

This section defines a number of terms and other document conventions, which are used with specific meaning in this memo.

media feature
: information that indicates facilities assumed to be available for the message content to be properly rendered or otherwise presented. Media features are not intended to include information that affects message transmission.

feature set
: some set of media features described by a media feature assertion, as described in "A Syntax for Describing Media Feature Sets" [1]. (See that memo for a more formal definition of this term.)

feature set expression
: a string that describes some feature set, formulated according to the rules in "A Syntax for Describing Media Feature Sets" [1] (and possibly extended by other specifications).

This specification uses syntax notation and conventions described in RFC 2234 "Augmented BNF for Syntax Specifications: ABNF" [3].

> NOTE: Comments like this provide additional nonessential information about the rationale behind this document. Such information is not needed for building a conformant implementation, but may help those who wish to understand the design in greater depth.

2. Motivation and goals

It is envisaged that media feature labelling of message parts may be used in the following ways:

- to supply more detailed media feature information about a message content than can be provided by the 'Content-type:' header.
- to provide summary media feature information (possibly including MIME content types) about the content of a composite MIME message part (e.g. 'multipart' or 'message'), without having to open up the inner content of the message.
- to supply media feature information about external data referenced by a message part (e.g. 'message/external-body' MIME type). This information would not be available by examination of the message content.
- to describe the content of a message that is encrypted or encoded using some application-specific file structure that hides the content from a MIME processor. This information also would not be generally available by examination of the message content.

3. The 'Content-features:' MIME header

A new header field is defined that extends the allowable formats for 'optional-field' [4] with the following syntax:

```
optional-field =/ "Content-features" ":" Feature-expr
Feature-expr   =  filter      ; See [1], section 4.1
```

where 'filter' is the media feature expression format defined by "A Syntax for Describing Media Feature Sets" [1].

This header provides additional information about the message content directly contained or indirectly referenced in the corresponding MIME message part.

3.1 Whitespace and folding long headers

In some circumstances, media feature expressions can be very long.

According to "A Syntax for Describing Media Feature Sets" [1], whitespace is allowed between lexical elements of a media feature expression. Further, RFC822/MIME [4,5] allows folding of long headers at points where whitespace appears to avoid line length restrictions.

Therefore, it is recommended that whitespace is included as permitted, especially in long media feature expressions, to facilitate the folding of headers by agents that do not otherwise understand the syntax of this field.

3.2 Usage considerations

3.2.1 Simple message parts

When applied to a simple MIME message part, the header should appear just once and is used to convey additional information about the message part content that goes beyond that provided by the MIME 'Content-type:' header field. The 'Content-features:' header may indicate a content type that is different than that given by the MIME 'Content-type:' header. This is possible but not recommended when applied to a non-composite body part: in any case, MIME content type processing must be performed in accordance with the 'Content-type:' header.

> NOTE: Once the message content has been delivered to an application, it is possible that subsequent processing may be affected by content type information indicated by the media feature expression. See example 4.5 below.

3.2.2 Multipart and other composites

'Content-features:' headers may be applied to a MIME multipart indicating information about the inner content of the multipart.

Implementations must not assume a one-to-one relationship between 'Content-features' headers and contained body parts. Headers may appear on a containing multipart wrapper in a different order than the body parts to which they refer; a single header may refer to more than one contained body part; several headers may refer to the same contained body part.

If it is important to relate specific media features to specific contained MIME body parts, then the 'Content-features:' header should be applied directly to the body part concerned, rather than the surrounding composite.

> NOTE: The intent here is to allow summary media feature information to be provided without having to open up and examine the inner content of the MIME message.

Similar usage may apply when the message format is a non-MIME or opaque composite; e.g. 'application/zip', or an encrypted message. In these cases, the option of examining the message content to discover media feature information is not available.

3.2.3 Reference to external data

Media feature information about data indirectly referenced by a MIME body part rather than contained within a message can be conveyed using one or more 'Content-features:' headers.

For example, media information --including contained MIME content type(s)-- about the data referenced by a MIME 'Message/external-body' may be conveyed.

4. Examples

4.1 Simple message

```
Mime-Version: 1.0
Content-type: text/plain;charset=US-ASCII
Content-features: (& (paper-size=A4) (ua-media=stationery) )

 :
(data)
 :
```

4.2 Fax message

```
Mime-Version: 1.0
Content-Type: multipart/mixed; boundary="break"
Content-features:
  (& (Type="image/tiff")
     (color=Binary)
     (image-file-structure=TIFF-S)
     (dpi=200)
     (dpi-xyratio=200/100)
     (paper-size=A4)
     (image-coding=MH) (MRC-mode=0)
     (ua-media=stationery) )

--break
Content-Type: image/tiff; name="coverpage.tiff"
Content-Transfer-Encoding: base64
Content-Description: This part is a coverpage
Content-Disposition: attachment; filename="coverpage.tiff"

0M8R4KGxGuEAAAAAAAAAAAAAAAAAAAAPgADAP7/CQAGAAAAAAAAAAAAAA
AAAAAAAAEAAAZAAAAAEAAAD+////AAAAAAAAAAD/////////////////
 :
(more data)
 :
--break

Content-Type: image/tiff; name="document.tiff"
Content-Transfer-Encoding: base64
Content-Disposition: attachment; filename="document.tiff"

AAAADgAAAA8AAAAQAAAAEQAAABIAAAATAAAAFAAAABUAAAAWAAAAFwAAABg
GgAAABsAAAAcAAAAHQAAAB4AAAAfAAAAIAAAACEAAAAiAAAAIwAAACQAAAA
 :
(more data)
 :
--break--
```

4.3 Multipart/alternative data

This example illustrates three points:

- Information about the various parts in a multipart/alternative can be made available before the alternative body parts are processed. This may facilitiate optimum one-pass processing of multipart/alternative data.

- There may be alternatives having the same basic MIME content-type, but differing in the content features that they use.

- There is NO defined correspondence between 'Content-features' headers and contained body parts.

```
Mime-Version: 1.0
Content-Type: multipart/alternative; boundary="break"
Content-features: (& (Type="text/plain") (charset=US-ASCII) )
Content-features:
  (& (Type="text/html") (charset=ISO-8859-1) (color=limited) )
Content-features:
  (& (Type="text/html") (charset=ISO-8859-1) (color=binary) )

--break
Content-type: "text/plain";charset=US-ASCII
Content-features: (color=binary)

 :
(data)
 :
--break
Content-type: "text/plain";charset=US-ASCII
Content-features: (color=limited)

 :
(data)
 :
--break


Content-type: text/html;charset=iso-8859-1
Content-features: (color=binary)

 :
(data)
 :
--break
Content-type: text/html;charset=iso-8859-1
Content-features: (color=limited)

 :
(data)
 :
--break--
```

4.4 Reference to external message data

```
Mime-Version: 1.0
Content-type: message/external-body; access-type=URL;
              URL="http://www.foo.com/file1.html"

Content-type: Multipart/mixed
Content-features: (& (Type="text/plain") (charset=US-ASCII) )
Content-features: (& (Type="image/tiff") (color=limited) )

<end>
```

4.5 Compressed data

This example shows how the 'Content-features' header can be used to overcome the problem noted in the MIME registration for 'Application/zip' regarding information about the data content.

```
Mime-Version: 1.0
Content-type: application/zip
Content-features: (& (Type="text/plain") (charset=US-ASCII) )
Content-features: (& (Type="image/tiff") (color=limited) )
Content-transfer-encoding: base64

 :
(data)
 :
<end>
```

4.6 Multipart/related data

(See also: RFC 2387, "The MIME Multipart/Related Content-type" [8])

```
Mime-Version: 1.0
Content-Type: multipart/related; boundary="boundary-example";
              type="text/html"; start="<foo3@foo1@bar.net>"
Content-features: (& (type="text/html") (charset=US-ASCII) )
Content-features: (type="image/gif")

--boundary-example
Content-Type: text/html;charset=US-ASCII
Content-ID: <foo3@foo1@bar.net>

referencing a resource in another body part, for example
through a statement such as:
<IMG SRC="http://www.ietf.cnri.reston.va.us/images/ietflogo.gif"
 ALT="IETF logo">
```

```
--boundary-example
Content-Location:
   http://www.ietf.cnri.reston.va.us/images/ietflogo.gif
Content-Type: IMAGE/GIF
Content-Transfer-Encoding: BASE64

R0lGODlhGAGgAPEAAP/////ZRaCgoAAAACH+PUNvcHlyaWdodCAoQykgMTk5
NSBJRVRGLiBVbmF1dGhvcml6ZWQgZHVwbGljYXRpb24gcHJvaGliaXRlZC4A
etc...

--boundary-example--
```

5. Security Considerations

When applied to simple or multipart MIME formatted data, a media feature expression provides summary information about the message data, which in many cases can be determined by examination of the message content. Under these circumstances, no additional security considerations appear to be raised.

When applied to other message composites, especially encrypted message content, feature expressions may disclose information that is otherwise unavailable. In these cases, some security considerations associated with media content negotiation [1,2] may have greater relevance.

It is suggested here that media feature descriptions may be usefully employed with encrypted message content. In doing this, take care to ensure that the purpose of encryption is not compromised (e.g. encryption might be intended to conceal the fact that a particular application data format is being used, which fact might be disclosed by an injudiciously applied Content-features header).

If a 'Content-features' header is applied to a multipart/signed object (or indeed outside any other form of signed data) the media feature information is not protected. This unprotected information could be tampered with, possibly fooling implementations into doing inappropriate things with the contained material. (Putting the media feature information inside the signed information would overcome this, at the cost of requiring implementations to parse the inner structure to find it.)

6. Acknowledgements

This proposal draws from discussions with Dan Wing. The fax message example was taken from a proposal by Mike Ruhl. The multipart/related example is developed from RFC 2557 [7].

The author would like to thank the following people who offered comments that led to significant improvements: Mr Hiroshi Tamura, Ted Hardie, Maurizio Codogno, Jacob Palme, Ned Freed.

7. References

[1] Klyne, G., "A Syntax for Describing Media Feature Sets", RFC 2533, March 1999.

[2] Holtman, K., Mutz, A. and T. Hardie, "Media Feature Tag Registration Procedure", RFC 2506, March 1999.

[3] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", RFC 2234, November 1997.

[4] Crocker, D., "Standard for the format of ARPA Internet text messages", STD 11, RFC 822, August 1982.

[5] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part 1: Format of Internet message bodies", RFC 2045, November 1996.

[6] Levinson, E., "The MIME Multipart/Related Content-type", RFC 2387, August 1998.

[7] Palme, J., Hopmann, A. and N. Shelness, "MIME Encapsulation of Aggregate Documents, such as HTML (MHTML)", RFC 2557, March 1999.

8. Author's Address

Graham Klyne
Content Technologies Ltd.
1220 Parkview,
Arlington Business Park
Theale
Reading, RG7 4SA
United Kingdom

Phone: +44 118 930 1300
Fax: +44 118 930 1301
EMail: GK@ACM.ORG

9. Full Copyright Statement

Acknowledgement

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