RFC 9694
Guidelines for the Definition of New Top-Level Media Types

Abstract

This document defines best practices for defining new top-level media types. It also introduces a registry for top-level media types, and contains a short history of top-level media types. It updates RFC 6838.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

This document defines best practices for defining new top-level media types. Top-level media types ('top-level types' for short) appear to the left of the slash in a media type, examples being 'text/...', 'application/...', 'image/...', and so on. Please note that top-level types are different from trees (standards tree, vendor tree, personal tree), which (except for the standards tree) are indicated immediately to the right of the slash with a prefix of '.../vnd.' or '.../prs.'. Section 4.2.7 of RFC 6838 [RFC6838] only summarily gives criteria for defining additional top-level media types. This document provides more detailed criteria for defining additional top-level media types. It therefore updates RFC 6838 [RFC6838].

1.1. Background

New top-level types are rare enough and different enough from each other that each application needs to be evaluated separately. The main protocol extension point for media types are subtypes below each of the main types. For formats that do not fit below any other top-level type, the 'application' top-level type can always be used.

The main function of media types and subtypes is the dispatch of data formats to application code. In most cases, this requires and is done using the full type (i.e., including the subtype, and often some parameters). The top-level type can occasionally serve as a fallback for the tentative dispatch to applications handling a very wide range of related formats. Please note that assumptions about the correctness of a media type must be made carefully, as it could be under the control of an attacker.

In some older scenarios, it may also have been possible to identify a device (e.g., a phone for audio messages, a printer or fax device for images, a video recorder for videos, a computer for 'application' subtypes). However, the current hardware landscape, where computers and smartphones can handle a very wide variety of media, makes such a scenario look somewhat far-fetched.

The top-level type can be used for user-directed information. Besides direct inspection of the type string by the user, this includes using different types of default icons for different top-level types.

1.2. Requirements Language

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. Rules and Criteria for the Registration of New Top-Level Media Types

This section describes the rules and criteria for new top-level media types, including criteria already defined in RFC 6838 [RFC6838].

2.1. Required Criteria

The following is the list of required criteria for the definition of a new top-level type. Motivations for the requirements are also included.

- Every new top-level type **MUST** be defined in a Standards Track RFC (see Section 4.9 of RFC 8126 [RFC8126]). This will ensure there is sufficient community interest, review, and consensus appropriate for a new top-level type.
- The IANA Considerations section of an RFC defining a new top-level type **MUST** request that IANA add this new top-level type to the registry of top-level types.

- The criteria for what types do and do not fall under the new top-level type **MUST** be defined clearly. Clear criteria are expected to help expert reviewers evaluate whether or not a subtype belongs below the new type, and whether the registration template for a subtype contains the appropriate information. Criteria that cannot be defined clearly is a strong indication that whatever is being talked about is not suitable as a top-level type.
- Any RFC defining a new top-level type **MUST** clearly document the security considerations applying to all or a significant subset of subtypes.
- At a minimum, one subtype **MUST** be described. A top-level type without any subtypes serves no purpose. Please note that the 'example' top-level describes the subtype 'example'.

2.2. Additional Considerations

- Existing wide use of an unregistered top-level type may be an indication of a need, and therefore may be an argument for formally defining this new top-level type.
- On the other hand, the use of unregistered top-level types is highly discouraged.
- Use of an IETF WG to define a new top-level type is not needed, but may be advisable in some cases. There are examples of new top-level type definitions without a WG (RFC 2077 [RFC2077]), with a short, dedicated WG (RFC 8081 [RFC8081]), and with a WG that included other related work (RFC 9695 [RFC9695]).
- The document defining the new top-level type should include initial registrations of actual subtypes. The exception may be a top-level type similar to 'example'. This will help show the need for the new top-level type, allow checking the appropriateness of the definition of the new top-level type, avoid separate work for registering an initial slate of subtypes, and provide examples of what is considered a valid subtype for future subtype registrations.
- The registration and actual use of a certain number of subtypes under the new top-level type should be expected. The existence of a single subtype should not be enough; it should be clear that new similar types may appear in the future. Otherwise, the creation of a new top-level type is most probably not justified.
- The proposers of the new top-level type and the wider community should be willing to commit to emitting and consuming the new top-level type in environments that they control.
- Desirability for common parameters: The fact that a group of (potential) types have (mostly) common parameters may be an indication that they belong under a common new top-level type.
- Top-level types can help humans with understanding and debugging. Therefore, evaluating how a new top-level type helps humans understand types may be crucial. But as often with humans, opinions may widely differ.
- Common restrictions may apply to all subtypes of a top-level type. Examples are the restriction to CRLF line endings for subtypes of type 'text' (at least in the context of electronic mail), or on subtypes of type 'multipart'.
- Top-level types are also used frequently in dispatching code. For example, "multipart/*" is frequently handled as multipart/mixed, without understanding of a specific subtype. The top-level types 'image', 'audio', and 'video' are also often handled generically. Documents with these top-level types can be passed to applications handling a wide variety of image, audio, or video formats. HTML-generating applications can select different HTML elements

(e.g., <img> or <audio>) for including data of different top-level types. Applications can select different icons to represent unknown types in different top-level types.

2.3. Negative Criteria

This subsection lists negative criteria for top-level types; it identifies criteria that are explicitly not reasons for a top-level type registration.

• A top-level type is not a pointer into another registration space that offers duplicate registrations for existing media types. Example: a top-level type of 'oid', leading to types of the form oid/nnnnn, where nnn is an OID (Object Identifier) designating a specific media format.

• A top-level type **MUST NOT** be defined for the mapping of other protocol elements to media types. For example, while there may be some merit to a mapping from media types to URIs, e.g., in the context of RDF (Resource Description Framework), there is very limited merit in a reverse mapping, and even less merit in creating a top-level type for such a mapping. The same applies to other protocol elements such as file extensions or URI schemes. If a mapping is needed, the recommended solution is to choose a single type/subtype and put the additional information in an appropriately named parameter. As an example, information on a file extension '.dcat' can be encoded as 'application/octet-string; filename=foo.dcat'.

• Media types are not a general type system. A top-level type **MUST NOT** be defined if its main or only purpose is to map other type systems, e.g., in programming languages or ontologies.

• A new top-level type **SHOULD NOT** generate aliases for existing widely used types or subtypes.

• Top-level types with an "X-" prefix cannot be registered, and **SHOULD NOT** be used. This is in line with RFC 6648 [RFC6648].

3.   Top-Level Media Type History

This section briefly describes the history of top-level types. The emphasis is on the aspects of the history that are relevant to the adoption of new top-level types.

RFC 1341 [RFC1341] first defined the structuring of content types into (top-level) type and subtype, and introduced the 'text', 'multipart', 'message', 'image', 'audio', 'video', and 'application' top-level types. That specification also allowed top-level types starting with 'X-'. With respect to new top-level types, it said the following:

> An initial set of seven Content-Types is defined by this document. This set of top-level names is intended to be substantially complete. It is expected that additions to the larger set of supported types can generally be accomplished by the creation of new subtypes of these initial types. In the future, more top-level types may be defined only by an extension to this standard. If another primary type is to be used for any reason, it must be given a name starting with "X-" to indicate its non-standard status and to avoid a potential conflict with a future official name.

The first time an additional top-level type was defined was in RFC 1437 [RFC1437], but this was an April Fools RFC, purely for entertainment purposes.

RFC 2046 [RFC2046] discouraged the use of "X-" for (new) top-level types, with the following words:

> In general, the use of "X-" top-level types is strongly discouraged. Implementors should invent subtypes of the existing types whenever possible. In many cases, a subtype of "application" will be more appropriate than a new top-level type.

RFC 2048 [RFC2048], published at the same time as RFC 2046 [RFC2046], defined requirements for the definition of new top-level types:

> In some cases a new media type may not "fit" under any currently defined top-level content type. Such cases are expected to be quite rare. However, if such a case arises a new top-level type can be defined to accommodate it. Such a definition must be done via standards-track RFC; no other mechanism can be used to define additional top-level content types.

The 'model' top-level type was introduced by RFC 2077 [RFC2077] in 1997.

RFC 4735 [RFC4735] introduced the 'example' top-level type for use in documentation examples.

The 'font' top-level type was defined in RFC 8081 [RFC8081], a work of the 'justfont' IETF WG, in 2017. This was formalizing the widespread use of the unofficial 'font' top-level type that people were using in preference to official, registered types.

There is ongoing work to define a new 'haptics' top-level media type in RFC 9695 [RFC9695].

Wikipedia (at https://en.wikipedia.org/wiki/Chemical_file_format) reports the unofficial use of a 'chemical' top-level type. This top-level type was proposed by Peter Murray-Rust and Henry Rzepa at a workshop at the First WWW conference in May 1994 [CHEMIME]. It is in widespread use but remains unregistered.

Some Linux desktop logic uses what looks like a top-level type of 'x-scheme-handler' to map URI schemes to applications. In addition, the type 'inode/directory' is used. However, this is a purely local, system-specific use, and is not intended for exchange. If exchange or standardization are desired, a change from, for example, 'x-scheme-handler/http' to something like 'application/ scheme-handler; scheme=http' or 'inode/directory' to 'multipart/inode-directory' or 'application/ inode-directory (in all cases, properly registered) is strongly recommended.

The document currently defining the requirements for new top-level media types is RFC 6838 [RFC6838]. Of particular relevance to the work in this document are Sections 4.2.5 (Application Media Types) and 4.2.7 (Additional Top-Level Types) of [RFC6838]. These two sections are not

strictly aligned, because the first says that anything that doesn't go under a more specific type can go under the 'application' top-level type, while the later section allows for new top-level types.

4. IANA Considerations

4.1. Registration of Top-level Media Types

Registrations of new top-level types follow the "Standards Action" policy (see Section 4.9 of RFC 8126 [RFC8126]).

Registrations of new top-level types have to provide the name of the top-level type, the defining specification (RFC, or the respective draft during the approval process), and, if applicable, some comments. The defining specifications have to contain an "IANA Considerations" section requesting addition to the registry of top-level media types and document security considerations for the top-level types they register.

The comments field is empty or contains short comments about the usage of the type. Comments can be added or updated by the experts for subtype registrations under the respective top-level type, and by IANA itself.

There should be at least one subtype, except for registrations that are for demonstration purposes only (e.g. the example top-level type).

4.2. Initialization of the Registry of Top-Level Media Types

IANA has created the "Top-Level Media Types" registry and populated it with the values in Table 1. IANA also added a pointer to this registry from the "Media Types" registry group.

For each top-level media type, the registry contains the name of the type, a pointer to the RFC defining the type, a pointer to IANA's registry of subtypes for that type, and a comment field.

The initial state of the registry is as follows:

name	Defining RFC	Registry of Subtypes	Comments
application	[RFC2046]	[Application Media Types]	-
audio	[RFC2046]	[Audio Media Types]	-
example	[RFC4735]	[Example Media Types]	no registrations, for examples only
font	[RFC8081]	[Font Media Types]	-

name	Defining RFC	Registry of Subtypes	Comments
haptics	[RFC9695] [RFC9695]	[Haptics Media Types]	-
image	[RFC2046]	[Image Media Types]	-
message	[RFC2046]	[Message Media Types]	-
model	[RFC2077]	[Model Media Types]	-
multipart	[RFC2046]	[Multipart Media Types]	-
text	[RFC2046]	[Text Media Types]	requires CRLF for newlines
video	[RFC2046]	[Video Media Types]	-

Table 1: Initial Values for the Registry of Top-level Media Types

IANA has also added pointers to this document and to the "Top-Level Media Types" registry in the application for a media type at <https://www.iana.org/form/media-types>.

5. Security Considerations

This document as such is not expected to introduce any security issues. The security issues related to introducing a new top-level media type **MUST** be evaluated and documented carefully.

Acknowledgements

Continuous encouragement for writing this document came from Harald Alvestrand. Further encouragement was provided by Murray S. Kucherawy. Both Harald and Murray also provided ideas for actual text. Without them, this memo would never have reached even the first draft stage. Alexey Melnikov provided the difficult to find pointer to RFC 2077 [RFC2077] and examples for applications dispatching on top-level types. Additional information and comments were received from Chris Lilley, Graham Kline, Henry S. Rzepa, Francesca Palombini, Zaheduzzaman Sarker, Amanda Baber, Paul Wouters, Roman Danyliw, John Scudder, Radia Perlman, Lars Eggert, and Antoine Fressancourt. Inspiration for negative criteria or examples were provided by Phillip Hallam-Baker, Donald E. Eastlake 3rd, Petter Reinholdtsen, and Christian Heller.

References

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/info/rfc2119>.

[RFC6838] Freed, N., Klensin, J., and T. Hansen, "Media Type Specifications and Registration Procedures", BCP 13, RFC 6838, DOI 10.17487/RFC6838, January 2013, <https://www.rfc-editor.org/info/rfc6838>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[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>.

Informative References

[CHEMIME] Rzepa, H.S., Murray-Rust, P., and B. Whitaker, "The Application of Chemical Multipurpose Internet Mail Extensions (Chemical MIME) Internet Standards to Electronic Mail and World Wide Web Information Exchange", Journal of Chemical Information Computer Science, vol. 38, no. 6, pp. 976-982, DOI 10.1021/ci9803233, 14 August 1998, <https://pubs.acs.org/doi/10.1021/ci9803233>.

[RFC1341] Borenstein, N. and N. Freed, "MIME (Multipurpose Internet Mail Extensions): Mechanisms for Specifying and Describing the Format of Internet Message Bodies", RFC 1341, DOI 10.17487/RFC1341, June 1992, <https://www.rfc-editor.org/info/rfc1341>.

[RFC1437] Borenstein, N. and M. Linimon, "The Extension of MIME Content-Types to a New Medium", RFC 1437, DOI 10.17487/RFC1437, April 1993, <https://www.rfc-editor.org/info/rfc1437>.

[RFC2046] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types", RFC 2046, DOI 10.17487/RFC2046, November 1996, <https://www.rfc-editor.org/info/rfc2046>.

[RFC2048] Freed, N., Klensin, J., and J. Postel, "Multipurpose Internet Mail Extensions (MIME) Part Four: Registration Procedures", RFC 2048, DOI 10.17487/RFC2048, November 1996, <https://www.rfc-editor.org/info/rfc2048>.

[RFC2077] Nelson, S., Parks, C., and Mitra, "The Model Primary Content Type for Multipurpose Internet Mail Extensions", RFC 2077, DOI 10.17487/RFC2077, January 1997, <https://www.rfc-editor.org/info/rfc2077>.

[RFC4735] Taylor, T., "Example Media Types for Use in Documentation", RFC 4735, DOI 10.17487/RFC4735, October 2006, <https://www.rfc-editor.org/info/rfc4735>.

[RFC6648]   Saint-Andre, P., Crocker, D., and M. Nottingham, "Deprecating the "X-" Prefix and Similar Constructs in Application Protocols", BCP 178, RFC 6648, DOI 10.17487/RFC6648, June 2012, <https://www.rfc-editor.org/info/rfc6648>.

[RFC8081]   Lilley, C., "The "font" Top-Level Media Type", RFC 8081, DOI 10.17487/RFC8081, February 2017, <https://www.rfc-editor.org/info/rfc8081>.

[RFC9695]   Muthusamy, Y. K. and C. Ullrich, "The 'haptics' Top-level Media Type", RFC 9695, DOI 10.17487/RFC9695, December 2024, <https://www.rfc-editor.org/info/rfc9695>.

Author's Address

Martin J. Dürst
Aoyama Gakuin University
Fuchinobe 5-10-1, Chuo-ku, Sagamihara, Kanagawa
252-5258
Japan
Phone: +81 42 759 6329
Email: duerst@it.aoyama.ac.jp
URI: https://www.sw.it.aoyama.ac.jp/Dürst/