



Describing Media Content of Binary Data in XML

W3C Working Group Note 4 May 2005

This version:

<http://www.w3.org/TR/2005/NOTE-xml-media-types-20050504>

Latest version:

<http://www.w3.org/TR/xml-media-types>

Previous version:

<http://www.w3.org/TR/2005/NOTE-xml-media-types-20050502>

Editors:

Anish Karmarkar, Oracle

Ümit Yalçınalp, SAP

Copyright © 2005 W3C® (MIT, ERCIM, Keio), All Rights Reserved. W3C liability, trademark and document use rules apply.

>

>Abstract

This document addresses the need to indicate the content-type associated with binary element content in an XML document and the need to specify, in XML Schema, the expected content-type(s) associated with binary element content. It is expected that the additional information about the content-type will be used for optimizing the handling of binary data that is part of a Web services message.

Status of this Document

This section describes the status of this document at the time of its publication. Other documents may supersede this document. A list of current W3C publications and the latest revision of this technical report can be found in the W3C technical reports index at <http://www.w3.org/TR/>.

This document is a W3C Working Group Note. This document includes the resolution of the comments received on the Last Call Working Draft previously published. The comments on this document and their resolution can be found in the Web Services Description Working Group's issues list.

There is no technical difference between this document and the 2 May 2005 version; the acknowledgement section has been updated to thank external contributors. A diff-marked version against the previous version of this document is available.

It has been produced jointly by the XML Protocol Working Group, and the Web Services Description Working Group, which are part of the Web Services Activity.

No further work on this topic is planned at this point. Errors in this document can be reported to the public public-ws-media-types@w3.org mailing list (public archive).

Publication as a Working Group Note does not imply endorsement by the W3C Membership. This is a draft document and may be updated, replaced or obsoleted by other documents at any time. It is inappropriate to cite this document as other than work in progress.

Table of Contents

1 Introduction [p.2]

- > 1.1 Notational Conventions<a>
- > 1.2 Requirements<a>
- > 2 Attributes for Declaring Content-Type<a>
- > 2.1 xmime:contentType Attribute<a>
- > 2.2 xmime:expectedContentTypes Attribute<a>
- > 3 Declaring Content-Type for Binary Data<a>
- > 3.1 Role of xmime:expectedContentTypes Schema Annotation Attribute<a>
- > 4 Examples<a>
- > 4.1 Binary Data with Known Media Type<a>
- > 4.2 Binary Data with Preferred Media Type<a>
- > 5 Normative References<a>
- > 6 Informative References<a>
- ><p>

Appendices

A Acknowledgements [p.3]

- > B Schema<a>
- ><p>

>

>1 Introduction

Data sent and received over the Web typically uses the MIME media type defined by [IETF RFC 2046] [p.4] , as the type system. For example, "image/jpeg", "application/pdf". There is a need to indicate the content-type of the XML element content, for example, in messages sent and received by Web services. There is also a need to express the content-type information using [XML Schema: Datatypes] [p.4] and [XML Schema: Datatypes] [p.4] , which is the type system used by [WSDL 2.0 Part 1] [p.5] . This would allow XML-based applications, such as Web services, to utilize the widely deployed and supported MIME media type infrastructure.

[XOP] [p.5] and [MTOM] [p.6] enables one to serialize binary content (element content that is in a canonical lexical representation of the *xs:base64Binary* type) in an optimized way using MIME packaging. There is a desire to specify the content-type information of such binary element content in a standard way in the [XML Information Set] [p.8] and not just in the optimized serialization of that Infoset.

This document specifies:

- An attribute (*xmime:contentType*) to indicate the content-type of an XML element content whose type is *xs:base64Binary* or *xs:hexBinary*. The value of the attribute is a valid content-type string (e.g., "text/xml; charset=utf-16"). This attribute specifies the content-type of the element content on which it occurs.
- A XML Schema annotation attribute (*xmime:expectedContentTypes*) to indicate, in XML Schema, the expected content-type(s) for an element content whose type is *xs:base64Binary* or *xs:hexBinary*.

The XML Schema annotation, *xmime:expectedContentTypes*, specifies the expected range of values for the *xmime:contentType* attribute and the expected range of content-type for the binary element content.

Note that the use of this mechanism, in particular the *xmime:contentType* attribute, does not require the implementation, in whole or in part, of XML Schema. In the absence of XML Schema the type information (*xs:base64Binary* or *xs:hexBinary*) may have to be provided via other mechanisms; for example, using *xsi:type*.

1.1 Notational Conventions

The keywords "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [IETF RFC 2119] [p.8] .

This specification uses properties from the XML Information Set (see [XML Information Set] [p.9]). Such properties are denoted by square brackets, e.g. [namespace name].

This specification uses namespace prefixes that are listed in Table 1. Note that the choice of any namespace prefix is arbitrary and not semantically significant (see [XML Information Set] [p.9]).

Table 1. Prefixes and Namespaces used in this specification

Prefix	Namespace	Definition
xmime	http://www.w3.org/2005/05/xmlmime	Defined by this specification
xs	"http://www.w3.org/2001/XMLSchema"	Defined in the W3C XML Schema specifications [XML Schema: Structures] [p.10] , [XML Schema: Datatypes] [p.10] .
xsi	"http://www.w3.org/2001/XMLSchema-instance"	Defined in the W3C XML Schema specification [XML Schema: Structures] [p.10] .

Namespace names of the general form "http://example.org/..." and "http://example.com/..." represent application or context-dependent URIs (see [IETF RFC 3986] [p.9]).

All parts of this specification are normative, with the exception of examples and sections explicitly marked as "Non-Normative".

1.2 Requirements

This section describes the set of requirements that this document addresses.

1. Define how to indicate the content-type of an XML element content whose type is *xs:base64Binary* or *xs:hexBinary*. This is meta-data that may be, but not required to, used by tools to infer the specific content-type of binary data.
2. Define how to indicate the expected content-type(s) of XML element content whose type is *xs:base64Binary* or *xs:hexBinary* in XML Schema. This information is needed to define the set of content-type that a binary data may have. For example, a Web services application may be willing to indicate that the binary data represents an image, but leaves it to a document instance to further specify whether it is "jpeg", or "gif". This meta-data is not required to be present.
3. Define the acceptable format of content-type values.
4. Define the relationship between the expected and the actual value of the content-type declared for binary data in XML documents.

2 Attributes for Declaring Content-Type

This section defines two global *attribute information items* for declaring the content-type of binary data and expected content-type(s) of binary data in XML Schema to address requirements (1) and (2) above. Their usage is addressed in Section 3 **Declaring Content-Type for Binary Data** [p.9] .

2.1 `mime:contentType` Attribute

The `mime:contentType` *attribute information item* has the following Infoset properties:

- A [local name] of `contentType`.
- A [namespace name] of "http://www.w3.org/2005/05/xmlmime".

The type of the `mime:contentType` *attribute information item* is *xs:string* with a minimum length of three and all leading and trailing white space characters are ignored.

The [normalized value] of the `mime:contentType` *attribute information item* MUST be a valid Content-Type string, e.g., "image/png", "text/xml; charset=utf-16" as defined by [IETF RFC 2045] [p.9] and indicates the content-type of the [owner element]. Note that [normalized value] consists of normalized attribute value as defined by [XML Information Set] [p.10] and does not mean that two equivalent values of `mime:contentType` will necessarily be equal.

The `mime:contentType` *attribute information item* allows Web services applications to optimize the handling of the binary data defined by a *binary element information item* and should be considered as meta-data. The presence of the `mime:contentType` attribute does not changes the value of the element content.

2.2 mime:expectedContentTypes Attribute

The `mime:expectedContentTypes` *attribute information item* has the following Infoset properties:

- A [local name] of `expectedContentTypes`.
- A [namespace name] of `http://www.w3.org/2005/05/xmlmime`.

The type of the `mime:expectedContentTypes` *attribute information item* is `xs:string`.

The value and the meaning of the `mime:expectedContentTypes` attribute is similar to the value allowed for the 'Accept' HTTP header defined by HTTP 1.1 specification, Section 14.1 (see [IETF RFC 2616] [p.10]) and MUST follow the production rules defined in that section except for the following:

1. The prefix "Accept:" MUST NOT be used.
2. The rule `qdttext` is changed from: `qdttext = <any TEXT except">>` to: `qdttext = <any CHAR except">>` This change is made to disallow non-US-ASCII OCTETs.

The `mime:expectedContentTypes` *attribute information item* is intended to be used as part of XML Schema annotation for a *binary element information item* declaration (see **3 Declaring Content-Type for Binary Data** [p.10]). This *attribute information item* is meant to allow XML Schema authors to indicate the range of media types and/or associated parameters that are acceptable for the binary data. It serves as a static constrain on the `mime:contentType`. Users of this *attribute information item* are urged to avoid using wild cards (for example, "image/*") as it may lead to interoperability problems. If the set of expected media types is not known, the use of `mime:expectedContentTypes` is NOT RECOMMENDED.

3 Declaring Content-Type for Binary Data

Documents that want to specify additional content-type information for binary data SHOULD denote this by using a *binary element information item*. A *binary element information item* is an *element information item* defined with the following additional constraints.

- An OPTIONAL `mime:contentType` *attribute information item* as described above in **2.1 mime:contentType Attribute** [p.10] .
- The *character information items* comprising the [children] of the *element information item* MUST conform to the lexical constraints of `xs:base64Binary` or `xs:hexBinary`.

If the media type identified by the value of an `xmime:contentType` attribute information item is a text based media type then the value of the `xmime:contentType` attribute information item SHOULD include a `charset` parameter.

For authoring convenience, two types `xmime:base64Binary` and `xmime:hexBinary` are defined in **B Schema** [p.10]

Example 1: Element with binary content and `xmime:contentType` attribute

```
<?xml version="1.0" ?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns:tns="http://example.com/ct-required"
  xmlns:xmime="http://www.w3.org/2005/05/xmlmime"
  targetNamespace="http://example.com/ct-required">

  <xs:import namespace="http://www.w3.org/2005/05/xmlmime"
    schemaLocation="http://www.w3.org/2005/05/xmlmime"/>

  <!-- This element has binary content and requires the xmime:contentType
    attribute that indicates the content-type of the binary element -->
  <xs:element name="MyBinaryData">
    <xs:complexType>
      <xs:simpleContent>
        <xs:extension base="xs:base64Binary" >
          <xs:attribute ref="xmime:contentType" use="required"/>
        </xs:extension>
      </xs:simpleContent>
    </xs:complexType>
  </xs:element>

</xs:schema>
```

3.1 Role of `xmime:expectedContentTypes` Schema Annotation Attribute

The `xmime:expectedContentTypes` attribute is used for annotating XML Schema to indicate the expected range of content-type of the binary element content and the expected range of values for `xmime:contentType` attribute.

The value of the `xmime:contentType` attribute, if present, SHOULD be within the range specified by the `xmime:expectedContentTypes` annotation attribute, if specified in the schema. See Section 14.1 of [IETF RFC 2616] [p.10] on how to interpret content-type ranges that may be defined with respect to actual content. When the `xmime:expectedContentTypes` annotation attribute contains a wild card ("*") or a list of acceptable content-type separated by commas (","), the schema SHOULD require the `xmime:contentType` attribute to be present.

Applications that need to specify expected content-type SHOULD use the schema annotation to declare the range of expected values. `xmime:expectedContentTypes` annotation attribute MAY be used in conjunction with the declaration of *binary element information items* or with complex type definitions that are derived from `xs:base64Binary` or `xs:hexBinary` in XML Schema. If the `xmime:expectedContentTypes` annotation attribute is used in both the *binary element information item* declaration as well as definition of the complex type which the *binary element information item*

belongs to, then the expected range of values defined for the *binary element information item* MUST be a subset of the expected range of values defined for the complex type.

The `xmime:expectedContentTypes` annotation can be used in conjunction with either type or element declarations. Certain data-binding frameworks which use static type mappings can be more specific if the `xmime:expectedContentTypes` annotation is applied to the complexType declarations instead of the element declarations using those types. For this reason, the use of `expectedContentTypes` on element declarations using named complex types is not recommended. An example is provided in Example 6.

The example below consists of a type definition, `PictureType`, and an element declaration, `Picture`. The `xmime:contentType` attribute is required to be present and specifies the content-type of the binary content. The schema annotation attribute `xmime:expectedContentTypes` specifies that the media type of the binary content is 'image', and the subtype name is either 'jpeg' or 'png'.

Example 2: Schema declaring an element with binary content and expected media type of "image/jpeg" or "image/png"

```
<?xml version="1.0" ?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns:tns="http://example.com/wildcard"
  xmlns:xmime="http://www.w3.org/2005/05/xmlmime"
  targetNamespace="http://example.com/wildcard">

  <xs:import namespace="http://www.w3.org/2005/05/xmlmime"
    schemaLocation="http://www.w3.org/2005/05/xmlmime"/>

  <xs:complexType name="PictureType">
    <xs:simpleContent>
      <xs:restriction base="xmime:base64Binary" >
        <xs:attribute ref="xmime:contentType" use="required" />
      </xs:restriction>
    </xs:simpleContent>
  </xs:complexType>

  <!-- This element designates the range of values
    that the element definition will accept -->
  <xs:element name="Picture" type="tns:PictureType"
    xmime:expectedContentTypes="image/jpeg, image/png"/>

</xs:schema>
```

The example document instance below conforms to the element declaration of `Picture` and specifies that the binary content is of type "image/png".

Example 3: Document instance containing element with binary content-type "image/png"

```
<?xml version="1.0" ?>
<Picture xmlns="http://example.com/wildcard"
  xmlns:xmime="http://www.w3.org/2005/05/xmlmime"
  xmime:contentType="image/png">aWKKapGGyQ</Picture>
```

4 Examples

4.1 Binary Data with Known Media Type

The examples in this section consists of a binary elements whose media type is known in advance to be "image/jpeg".

In Example [p.10] 4, a fixed media type is specified by declaring it with an annotation in conjunction with the complex type definition. The attribute `xmime:contentType` is not used as the media type of the binary data is know in advance.

Example 4: Element with binary content, known media type and no `xmime:contentType` attribute

```
<?xml version="1.0" ?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns:tns="http://example.com/know-type"
  xmlns:xmime="http://www.w3.org/2005/05/xmlmime"
  targetNamespace="http://example.com/know-type">

  <xs:import namespace="http://www.w3.org/2005/05/xmlmime"
    schemaLocation="http://www.w3.org/2005/05/xmlmime"/>

  <xs:simpleType name="JPEGPictureType"
    xmime:expectedContentTypes="image/jpeg">
    <xs:restriction base="xs:base64Binary"/>
  </xs:simpleType>

  <xs:element name="JPEGPicture" type="tns:JPEGPictureType"/>

</xs:schema>
```

In Example [p.9] 5, a fixed media type is specified by declaring it with an annotation in conjunction with the element declaration. The attribute `xmime:contentType` is optionally used in document instances to indicate the media type of the binary data.

Example 5: Element with binary content, known media type and optional `xmime:contentType` attribute

```
<?xml version="1.0" ?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns:tns="http://example.com/know-type"
  xmlns:xmime="http://www.w3.org/2005/05/xmlmime"
  targetNamespace="http://example.com/know-type">

  <xs:import namespace="http://www.w3.org/2005/05/xmlmime"
    schemaLocation="http://www.w3.org/2005/05/xmlmime"/>

  <xs:element name="JPEGPicture" type="xmime:base64Binary"
    xmime:expectedContentTypes="image/jpeg" />

</xs:schema>
```

4.2 Binary Data with Preferred Media Type

This example illustrates that binary data with media type 'image/jpeg' is preferred but binary data with media type of 'image/tiff' is also allowed (with a lower preference).

Example 6: Element with binary content and preferred media type

```
<?xml version="1.0" ?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns:tns="http://example.com/preferred-type"
  xmlns:xmime="http://www.w3.org/2005/05/xmlmime"
  targetNamespace="http://example.com/preferred-type">

  <xs:import namespace="http://www.w3.org/2005/05/xmlmime"
    schemaLocation="http://www.w3.org/2005/05/xmlmime"/>

  <xs:complexType name="JPEGPreferredPictureType"
    xmime:expectedContentTypes="image/jpeg;q=1.0, image/tiff;q=0.8">
    <xs:simpleContent>
      <xs:restriction base="xmime:base64Binary" >
        <xs:attribute ref="xmime:contentType" use="required" />
      </xs:restriction>
    </xs:simpleContent>
  </xs:complexType>

  <xs:element name="JPEGPeferredPicture"
    type="tns:JPEGPreferredPictureType"/>

</xs:schema>
```

5 Normative References

XML Schema: Structures

XML Schema Part 1: Structures Second Edition [p.9] , H. Thompson, D. Beech, M. Maloney, and N. Mendelsohn, Editors. World Wide Web Consortium Recommendation, 28 October 2004. (See <http://www.w3.org/TR/2004/REC-xmlschema-1-20041028/>.)

XML Schema: Datatypes

XML Schema Part 2: Datatypes Second Edition [p.9] , P. Byron and A. Malhotra, Editors. World Wide Web Consortium Recommendation, 28 October 2004. (See <http://www.w3.org/TR/2004/REC-xmlschema-2-20041028/>.)

IETF RFC 3986

Uniform Resource Identifiers (URI): Generic Syntax [p.9] , T. Berners-Lee, R. Fielding, L. Masinter, January 2005. (See <http://www.ietf.org/rfc/rfc3986.txt>.)

IETF RFC 2045

RFC 2045 - Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies [p.5] , N. Freed, N. Borenstein, November 1996. (See <http://www.ietf.org/rfc/rfc2045.txt>.)

IETF RFC 2046

RFC 2046 - Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types [p.9] , N. Freed, N. Borenstein, November 1996. (See <http://www.ietf.org/rfc/rfc2046.txt>.)

IETF RFC 2616

Hypertext Transfer Protocol--HTTP 1.1 [p.10] , R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, T. Berners-Lee, June 1999. (See <http://www.w3.org/Protocols/rfc2616/rfc2616.html>.)

IETF RFC 2119

Key words for use in RFCs to Indicate Requirement Levels [p.10] , S. Bradner, Author. Internet Engineering Task Force, June 1999. (See <http://www.ietf.org/rfc/rfc2119.txt>.)

XML Information Set

XML Information Set (Second Edition) [p.5] , J. Cowan and R. Tobin, World Wide Web Consortium Recommendation, 4 February 2004. (See <http://www.w3.org/TR/2001/REC-xml-infoset-20011024/>.)

6 Informative References

WSDL 2.0 Part 1

Web Services Description Language (WSDL) Version 2.0 Part 1: Core Language [p.4] , Roberto Chinnici, Martin Gudgin, Jean-Jacques Moreau, Jeffrey Schlimmer, Sanjiva Weerawarana, World Wide Web Consortium Working Draft 3 August 2004 (See <http://www.w3.org/TR/2004/WD-wsdl20-20040803/>.)

XOP

XML-binary Optimized Packaging [p.10] , Martin Gudgin, Noah Mendelsohn, Mark Nottingham, Herve Ruellan, W3C Recommendation, 25 January 2005 (See <http://www.w3.org/TR/2005/REC-xop10-20050125/>.)

MTOM

SOAP Message Transmission Optimization Mechanism [p.10] , Martin Gudgin, Noah Mendelsohn, Mark Nottingham, Herve Ruellan, W3C Recommendation, 25 January 2005 (See <http://www.w3.org/TR/2005/REC-soap12-mtom-20050125/>.)

A Acknowledgements

This document is developed by the participants of the joint media types task force formed by Web Services Description [p.8] and XML Protocol [p.8] Working Groups. Participants of the taskforce, specifically Martin Gudgin, and Mark Nottingham are gratefully acknowledged.

The authors would also like to thank the commentators, in particular Joe Fialli, Larry Massinter, and Henry Thompson.

B Schema

```
<?xml version="1.0" ?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
            xmlns:xmime="http://www.w3.org/2005/05/xmimime"
            targetNamespace="http://www.w3.org/2005/05/xmimime" >

  <xs:attribute name="contentType">
    <xs:simpleType>
      <xs:restriction base="xs:string" >
        <xs:minLength value="3" />
      </xs:restriction>
    </xs:simpleType>
  </xs:attribute>
</xs:schema>
```

B Schema

```
        </xs:restriction>
      </xs:simpleType>
    </xs:attribute>

    <xs:attribute name="expectedContentTypes" type="xs:string" />

    <xs:complexType name="base64Binary" >
      <xs:simpleContent>
        <xs:extension base="xs:base64Binary" >
          <xs:attribute ref="xmime:contentType" />
        </xs:extension>
      </xs:simpleContent>
    </xs:complexType>

    <xs:complexType name="hexBinary" >
      <xs:simpleContent>
        <xs:extension base="xs:hexBinary" >
          <xs:attribute ref="xmime:contentType" />
        </xs:extension>
      </xs:simpleContent>
    </xs:complexType>
  </xs:schema>
```