

Notes on XML and Media Types

Chris Lilley

Changes from RFC 3023 to 3023-bis

- <http://www.w3.org/2006/02/son-of-3023/latest.html#anchor45>
- text/xml and text/xml-external-parsed-entity deprecated
- XPointer ([XPointerFramework] and [XPointerElement] added (fragment identifiers)
- XPointer Registry ([XPtrReg]) mentioned [XBase] added
- charset language now conforms to TAGMIME
- many references updated

Outstanding issues:

- text/xml or text/*+xml with missing charset parameter
- applications all rely on the xml encoding declaration
- MIME says to assume US-ASCII (good choice for text email, but...)
- RFC 3023 says to assume US-ASCII - universally ignored requirement
- RFC 3023-bis proposes to deprecate text/xml
- existing non-ASCII content (like RSS feeds, xslt) served as text/xml not going away
- deprecation rejected by implementors: not a solution. FIX text/xml instead.

text/* vs application/*

- email and HTTP differ in emphasis and goals here
- is the original distinction still relevant/important
- <http://www.mail-archive.com/whatwg@lists.whatwg.org/msg20855.html>

HTTPbis ISSUE 20 may help here

- <http://trac.tools.ietf.org/wg/httpbis/trac/ticket/20>

Closed issues:

- Inconsistency with W3C TAG finding Internet Media Type registration
 - consistency of use [TAGMIME]
<http://www.w3.org/2001/tag/2004/0430-mime>
 - Don't add a charset if its unknown, or differs from the encoding declaration
 - Unambiguous when charset omitted
- Generic processing of fragment IDs in RFC 3023bis
 - extensive TAG discussion
 - XML vs RDF/XML handling
- 3023bis and Xpointer
 - element() scheme required, or not
- XML 5th edition vs. 4th edition