Digital Imaging and Communications in Medicine (DICOM)

Supplement 211:
DICOMweb Support for the application/zip Payload

Prepared by:

Bill Wallace, Brad Genereaux

DICOM Standards Committee, Working Group 27

1300 N. 17th Street

Rosslyn, Virginia 22209 USA

Developed in accordance with work item WI 2018-09-C

VERSION: 19

January 16, 2020

Table of Contents

Scope and Field of Application

Supplement 211 adds the “application/zip” payload to the RESTful Web Services in PS3.18. It enables retrieving all the instances in an entire DICOM study or series as a single zip payload.

A key use case is to enable machine learning researchers to use DICOMweb to retrieve studies, series, or other collections of images for training purposes. DICOMweb currently provides a means to retrieve an entire study or an entire series using the “multipart/related” media type, but this requires special non-browser implementations to retrieve the content because multipart/related is not currently supported by any of the major browsers.

Usage of this application/zip payload on interchange media is not addressed by this supplement.

The proposed ZIP payload does not alter the scope of the existing WADO-RS services to add multiple study content.

Open Questions

#	Question	Position
17	Is it appropriate to make normative file extensions for the representations within the ZIP payload.	
18	Is raw the appropriate file extension for octet/stream	
19	Should we use UUID to identify Bulkdata references.	
20	How should the Bulkdata URI/UUID choice be specified (it was in Part 19 previous, see yellow highlighting below)	
21	How should the group 2 meta information be present in the XML/JSON.	
22	Are there any additional security concerns about returning ZIP responses?	

Closed Questions

#	Question	Notes
1	Is “zip” a “DICOM media type”, a rendered media	A: ZIP is a response payload The section below, 8.7.3.5.2 DICOM ZIP Media Type, defines this as another DICOM payload. Including DICOMDIR and restricting the format to the older ZIP format doesn’t make sense for an XML/JSON response possibly including Bulkdata.

	type, or something else? e.g. Do we want to include support for DICOMDIR structures, or make it the default (or only) format? / Would we consider supporting the DICOM media type already defined in Part 12 section V?	
2	Do we want to have parameters to support anonymization?	No The existing parameter from Part 18 Section 9.4.1.2.1 applies only to the original URL based retrieve. This would require a separate CP which would apply to both ZIP and multipart/related retrieves.
5	Do we want to specify a different RESTful resource? e.g., /packaged	No We don't believe that's necessary; it's essentially a transformation on the same path as other WADO-RS services. As with any of the REST services DICOMweb offers, the design tended to follow the principles of REST Uniform Interface (https://en.wikipedia.org/wiki/Representational_state_transfer#Uniform_interface) and Clean URLs (https://en.wikipedia.org/wiki/Clean_URL); ZIP is just a transformed representation of a DICOM study and it is acceptable and encouraged by REST community to specify it in this way In line with how other APIs work and is more discoverable and what's expected
6	Do we want to allow support for tar or rar, be silent on it, or explicitly disallow it?	No Other packaging formats will not be specified as the intent is for easy access by other systems and minimizing how many other formats may be available diminishes the value of having one recommended format.
8	Can we support retrievals that are completely specified in a URL (and not require Accept header to be passed)?	Yes This is already supported in PS3.18 8.3.3.1 (query parameter "accept"). http://dicom.nema.org/medical/dicom/current/output/html/part18.html#sect_8.3.3.1
12	Are there security concerns about popping a URL into Chrome for example?	No There are no additional security concerns beyond the security concerns in DICOMweb itself because the data itself is already available via existing URLs.
13	What happens if a client accepts both multipart/related and application/zip and are	HTTP protocol specifies it is a server's decision to choose. We believe that mantra should still apply

	"weighted" equally?	
16	Should we consider ZIP at the rendered level as well?	No This would be added at the same time/way as multipart payloads are handled.
20	Should encrypted ZIP's be allowed.	No. The standard ISO ZIP's do not allow this. There is significant work required to describe a suitable mechanism for this. Future work to be done.
27	What is the directory structure?	There are no additional requirements on the directory structure. It was felt that allowing the providing application to specify the structure allows more flexibility in implementations.
28	Should this form of application/zip be supported on interchange media? e.g. CD, email etc	No Interchange media should be small and have a minimum number of options for parsing data. This payload can be in XML or ZIP with or without Bulkdata, all of which make media usage more problematic.

Insert Reference in PS3.18 Section 2.1

[ISO/IEC 21320-1] ISO/IEC. 2015/10. 1.0. Information Technology – Document Container File. Available at https://nfoworks.org/notes/2013/04/n130401-13R5EoW9HY4ZvQH/n130401c11-ISO_IEC_21320-1_2015.pdf

Update PS3.18 Section 8.6.1 as follows:

8.6.1 Payload Format

Payloads may be in ~~either~~ single part**, ZIP,** or multipart format depending on the media type.

Insert PS3.18 Section 8.6.1.3 as follows:

8.6.1.3 ZIP Payload

A ZIP payload contains one or more representations. The media type of the payload shall be 'application/zip'. It is used as a container for distributing a set of related DICOM SOP Instances and/or BulkData. See PS3.18 Section 8.7.3.5.4 for details of the application/zip payload.

8.6.1.3.1 File Extensions

The media type of the representation determines the file extension that shall be used. Table 8.6.1.3.1-1 shows the DICOM Media Type and their associated file extensions.

Table J.1.2-1 DICOM Media Types and File Extensions

Media Type	File Extension
application/dicom	"dcm"
application/dicom+json	"json"
application/dicom+xml	"xml"
octet/stream	"raw"

8.6.1.3.2 BulkData URI

BulkData URIs within the DICOM Metadata representations (see Section 10.4.1.1.2) for any included BulkData Files, shall be relative references according to [RFC3986] section 4.

Relative URIs:

- the path shall not begin with a slash or two slashes
- the path shall not contain a backslash
- the path shall not refer to a higher level in the file system hierarchy (i.e., use of ".." is not permitted)
- file name extensions corresponding to executable file types (exe, dll, etc.) are not permitted
- embedded white space is not permitted

Note:

The following are examples of valid relative URI references:

"image1.json"

"series1/image3.dcm "

"./study1/series2/bulkdata1.raw"

The following are examples of invalid relative URI references:

"file:///matlist.mtl" (not a relative reference)

"/matlist.mtl" (no leading slashes)

"c:/matlist.mtl" (not a relative reference)

"../matlist.mtl" (no "..")

“setup.exe” (executable extension not permitted)

"mat list.mtl" (no embedded spaces)

8.6.1.3.3 Logical Format

The origin server shall format the ZIP according to [ISO/IEC 21320-1].

8.6.1.3.4 Metadata Representations

When a Metadata representation includes any Bulkdata references, the Metadata shall include File Meta Information Attributes within the Infoset.

Update PS3.18 Section 8.7 as follows:

8.7 Media Types

Media types are the basis for both content negotiation and data typing of message payloads. Each PS3.18 service, and/or transaction defines the media types and associated representations that are default, required and optional.

The media type also specifies whether the payload contains a single representation (single part), or multiple representations (multipart **or ZIP**). ~~**Multipart**~~ **Multiple representation** payloads are only defined for the RESTful APIs. See Section 8.6.1.2**, Section 8.6.1.3** and Section 10.4.3.

…

Update PS3.18 Section 8.7.1 as follows:

8.7.1 Multipart Media Types

Some of the services defined in this Part of the Standard support the multipart media types [RFC2387]. The syntax is:

```
multipart-media-type = "multipart" "/" subtype *(OWS ";" OWS parameter)
```

The application/multipart-related media type is used by the RESTful services. Its syntax is:

```
multipart-related = "multipart/related"
                    OWS ";" OWS "type" "=" DQ media-type DQ
                    OWS ";" OWS "boundary" "=" boundary
                    [related-parameters]
```

Where

```
boundary           ; See Section 8.6.1.2.1
```

```
bchar              = bchar-nospace / SP
bchar-nospace      = DIGIT / ALPHA / "'" / "(" / ")" / "+" / "_" / "," / "-"
                     / "." / "/" / ":" / "=" / "?" "/" / ":" / "=" / "?"
related-parameters = [";" "start" "=" cid]
                     [";" "start-info" "=" cid-list]
cid-list           = cid cid-list
cid                = token / quoted-string
```

The "type" parameter is required. It contains the media type of the "root" body part. It always contains the special character "/" and thus requires quote marks.

The cid is a content identifier. It should be unique for each part of the multipart message.

105 Typically, the "start" and "start-info" parameters are not specified, and the "root" is the first body part.

Bulkdata references to Bulkdata contained within the Multipart payload shall be UUID references.

Update PS3.18 Section 8.7.3.5 as follows:

8.7.3.5 DICOM Media Type Syntax

110 The syntax of DICOM Media Types is:

```
dicom-media-type = (dcm-singlepart / dcm-multipart / dcm-zip) [dcm-parameters]
```

Where

```
dcm-singlepart = dcm-mt-name

dcm-multipart  ;see Section 8.7.3.5.1
```

115 **`dcm-zip        ;see Section 8.7.3.5.4`**

```
dcm-parameters = transfer-syntax-mtp ;see Section 8.7.3.5.2

               / charset-mtp;see Section 8.7.3.5.3

dcm-mt-name    = dicom / dicom-xml / dicom-json ;DICOM Media Type name

dicom          = "application/dicom"

dicom-xml      = "application/dicom+xml"

dicom-json     = "application/dicom+json"

octet-stream   = "application/octet-stream"
```

All DICOM Media Types may have a Transfer Syntax parameter, but its usage may be constrained by the service for which they are used.

Note

The application/dicom+xml and application/dicom+json Media Types may have a Transfer Syntax parameter in order to specify the encoding of base64 data.

All DICOM Media Types may have a character set parameter, but its usage may be constrained by the service for which they are used.

Insert PS3.18 Section 8.7.3.5.4

8.7.3.5.4 DICOM ZIP Media Types

The syntax of ZIP media types is:

```
dcm-zip = "application/zip"
                [ OWS ";" OWS "type" "=" dcm-mp-mt-name ]
                [dcm-parameters]
```

Each ZIP media type may include a "type" parameter that defines the media types included with the ZIP file. The ZIP response shall be encoded according to Section 8.6.1.3. The type shall default to “application/dicom” if not specified.

Update PS3.18 Section 10.4.1.1.2 as follows:

10.4.1.1.2 Metadata Resources

Table 10.4.1-2 defines the resources used to retrieve the metadata contained in Instances.

Table 10.4.1-2. Retrieve Transaction Metadata Resources

Resource	URI Template
Study Metadata	`/studies/{study}/metadata`
Series Metadata	`/studies/{study}/series/{series}/metadata`
Instance Metadata	`/studies/{study}/series/{series}/instances/{instance}/metadata`

The Metadata Resources are used to retrieve the DICOM instances without retrieving Bulkdata. The Metadata returned for a study, series, or instance resource includes all Attributes in the resource. For Data Elements having a Value Representation (VR) of DS, FL, FD, IS, LT, OB, OD, OF, OL, OW, SL, SS, ST, UC, UL, UN, US, and UT, the origin server is permitted to replace the Value Field of the Data Element with a Bulkdata URI. The user agent can use the Bulkdata URI to retrieve the Bulkdata. **The Bulkdata URI shall be an HTTP(S) URI. A Bulkdata UUID shall not be used.**

Update PS3.18 Section 10.4.4 as follows:

10.4.4 Media Types

The origin server shall support the media types specified as default or required in Table 10.4.4-1.

Table 10.4.4-1. Default, Required, and Optional Media Types

Media Type	Usage	Section
application/dicom	Required	Section 8.7.3.1
application/dicom+json	Default	Section 8.7.3.2
multipart/related; type="application/dicom+xml"	Required	Section 8.7.3.2
Rendered Media Types **application/zip**	Optional **Optional**	Section 8.7.4 **Section 8.6.1.3**

The origin server shall support the Transfer Syntax and Character Set media type parameters. See Section 8.7.3.5.2 and Section 8.7.3.5.3.

Update PS3.19 Table A.1.5-2 as follows:

Table A.1.5-2. DICOM Data Set Macro

Name	Optionality	Cardinality	Description
…			
>BulkData	C	1	A reference to a blob of data that the recipient may retrieve through use of the GetData() method, a WADO-RS call or a STOW-RS call. Required if the DICOM Data Element represented is not zero length and an XML Infoset Value, Item, InlineBinary or PersonName element is not present. The provider of the data may use a BulkData reference at its discretion to avoid encoding a large DICOM Value Field as text by value in the Infoset. For example, pixel data or look up tables.

Name	Optionality	Cardinality	Description
			There is a single BulkData Infoset element representing the entire Value Field, and not one per Value in the case where the Value Multiplicity is greater than one. Note E.g., a LUT with 4096 16 bit entries that may be encoded in DICOM with a Value Representation of OW, with a VL of 8192 and a VM of 1, or a US VR with a VL of 8192 and a VM of 4096 would both be represented as a single BulkData element. All rules (e.g., byte ordering and swapping) in PS3.5 apply. Note Implementers should in particular pay attention the PS3.5 rules regarding the value representations of OD, OF, OL and OW. If the BulkData has a string or text Value Representation, the value(s) of the DICOM Specific Character Set Data Element, if present, might be necessary to determine its encoding.
>>uuid	C	A	An identifier of this bulk data reference formatted as a UUID using the hexadecimal representation defined in ITU-T Recommendation X.667. Required if BulkData URI is not present. Shall not be present otherwise.
>>uri	C	A	The ~~**HTTP(S)**~~ URI for this bulk data reference. Required if **the BulkData UUID is not present**. ~~**the NativeDicomModel was:**~~ • ~~**returned in response to a WADO-RS Retrieve Metadata request**~~ Shall not be present otherwise.
>InlineBinary	C	1	The Value Field of the enclosing Attribute encoded as base64. Required if the DICOM Data Element represented is: • not zero length • the VR if the enclosing Attribute is either OB, OD, OF, OW, or UN • an XML Infoset Value or BulkData XML element is not present Shall not be present otherwise.

Name	Optionality	Cardinality	Description
			There is a single InlineBinary Infoset element representing the entire Value Field, and not one per Value in the case where the Value Multiplicity is greater than one. Note E.g., a LUT with 4096 16 bit entries that may be encoded in DICOM with a Value Representation of OW with a VL of 8192 and a VM of 1 would be represented as a single InlineBinary element. All rules (e.g., byte ordering and swapping) in PS3.5 apply. Note Implementers should in particular pay attention to the PS3.5 rules regarding the value representations of OD, OF, OL and OW.