DICOM Secretariat/MITA
1300 North 17th Street, Suite 900
Arlington, VA 22209, USA
http://www.dicomstandard.org/
dicom@dicomstandard.org

MINUTES

MEETING NAME 09-WG33: WG-33 Data Archive and Management

MEETING PLACE/DIAL IN

DATE & TIME Wednesday, September 16, 2020 | 11:00 am – 12:30 pm US ET

PRESIDING OFFICERS Matthew Bishop, UnityPoint Health
Keith Eklund, Healthcare Tech Solutions

VOTING MEMBERS PRESENT

Ambra Health	Ostrow, Daniel
DesAcc EMEA	King, Graham
Grafimedia	Georgiadis, Pantelis
Healthcare Tech Solutions	Eklund, Keith
Laitek, Inc.	Costea-Barlutiu, Razvan
Laitek, Inc.	Brown, Barry
Laitek, Inc.	Solomon, Harry
PixelMed Publishing	Clunie, David
Society for Imaging Informatics In Medicine	Bishop, Matthew
Society for Imaging Informatics In Medicine	Carey, Cheryl
Sygehus Lillebaelt	Hansen, Peer Møller
Varian Medical Systems, Inc.	Schwere, Thomas

OTHERS

Citius Tech Healthcare Technology	Mahalle, Prashant
DeJarnette Research Systems	Wineke, Steve
Experies Consulting	McCloskey, Thomas
Hyland Software	Ullrich, Mike
Mega Informatica Ltd	Fauquex, Jacques
	Bronson, Hokuf

		Liu, Daniel
		Minner, Steve
		Wong, Lori
VOTING MEMBERS	AAPM	Bevins, Nicholas
	AAPM	Knazik, Shayna
ABSENT	Argentix Informatix	Silver, Elliot
	Canon Medical Research USA	O'Donnell, Kevin
	Canon/Vital Images	Dawson, Tim
	Canon/Vital Images	Whitby, Jonathan
	Change Healthcare	Ho, Kinson
	GE Healthcare	Nichols, Steven
	GE Healthcare	Numan, Jouke
	Mach7 Technologies	Ulanov, Alexey
	Mayo Clinic Rochester	Persons, Kenneth
	Society for Imaging Informatics In Medicine	Henson, Kyle

DICOM SECRETARIAT	Anna Zawacki, SIIM

1 CALL TO ORDER AND REVIEW OF ANTI-TRUST RULES AND DICOM PATENT POLICY (Co-Chairs, Secretariat)

The meeting was called to order. Guidelines for Conducting NEMA Meetings were read and attendance was recorded.

2 REVIEW AND APPROVE AGENDA (Co-Chairs)

The agenda was reviewed and approved.

3 REVIEW MINUTES (Co-Chairs)

The minutes of the previous meeting were reviewed and approved with a few edits.

4 TOPIC ITEMS TO BE DISCUSSED (All)

- Discuss how in the DICOM Standard to specify filesystem access to archive files

  ftp://d9-workgrps:Private15@medical.nema.org/MEDICAL/Private/Dicom/WORKGRPS/WG33/2020/2020-09-16/File%20access%20discussion.pptx

Open Issues / To Do

1	Referencing Files, File Sets, etc. – how to specify in Parts 10/11/12?	HS
2	How to specify matching keys for inventory?	RC
3	Count of inventory objects in set (n of m) – do we need to record 'm' in each object of set? How is it known before entire set is complete?	KH DO
4	PACS that manages multiple distributed archives (e.g., VNAs), but holds metadata updates itself (it proxies DIMSE/DICOMweb retrieves and applies updates) – how does it identify to migration client that metadata updates in inventory need to be applied when retrieving directly from VNA?	DO
5	Need required behavior for applying metadata updates to SOP Instances retrieved by non-DICOM protocol. (in Part 4 Annex I? – depends on approach to Pts 10/11/12)	HS
6	Do we need variable richness of metadata in inventory (e.g., additional Instance level attributes)? Can one size fit all, or do we need a few sizes, or totally variable?	
7	For inventory divided into multiple SOP Instances (for creator performance with enormous data sets), do we need a meta-inventory of inventory objects?	
8	What topics should be included in an informative Pt17 Annex?	

#1 – Harry Solomon, major topic for today's discussion
#3 – Kinson Ho and Daniel Ostrow will work on
#4 – Daniel Ostrow will work on

#7 ties back a bit to #3 – if there is no meta inventory of the objects it will be very hard to get a count, unless we include a special SOP systems that just has the count of inventory objects in the given inventory

Changes for DICOM part 10, Agreed to be able to reference part 10 compliant files and provide a non DICOM protocol access

WG-33 Consensus on access to stored instances

- PACS provides inventory of studies (new Inventory SOP Class)
- PACS must provide DIMSE and/or DICOMweb access to studies, series, and SOP Instances
- PACS may also provide direct "filesystem" access to SOP Instances
  - Specified by URI with protocol-based scheme (smb://, nfs://, http:// …)
  - SOP Instances are in Part 10 conformant files, or Part 10 files within an archive file format (ZIP, TAR, or TAR+GZIP)
- PACS may provide study- or series-level 'filesystem get' access if all instances are in single ZIP/TAR file or single folder/directory

Initial conceptual approach to DICOM standard

- Part 10 already specifies File Format, File Sets, and abstract File Service (minimum functionality)
- Add Part 12 media formats for archive File Sets
  - Already includes ZIP
  - TAR, TAR+GZIP
  - URI-defined access mapped to Part 10 abstract File Service
- Part 11 Profile for migration use case invoking Part 12 media formats

BUT …

Issues with initial concept

- Part 10 definition of File Sets intimately connected with presence of DICOMDIR file and constrained file IDs (8 characters)
- Current WG-33 approach uses Inventory SOP Instances, which are somewhat similar to Media Storage Directory SOP Instances used in DICOMDIR, but still different
  - May be more than one per archive File Set
- WG-33 approach of URI-based access should not be limited to Part 10 definition of constrained File IDs within File Sets

Option 1

- Refactor Parts 10, 11, 12 to separate DICOMDIR from File Set definition
- Allow File Sets without DICOMDIR, or with alternate means of identifying content of File Set (e.g., through Inventory SOP Instances)
- Allow unconstrained File IDs in URIs
- Retain all restrictions for existing Media Formats

Option 2

- Amend Part 10 to allow either Basic or Enhanced Directory
- Define Inventory as "Enhanced Directory" and place in file named DICOMDIR
- Allow unconstrained File IDs in URIs in Enhanced Directory

Option 3

- Require PACS to create DICOMDIR, but with no Directory records
  - Allowed in current DICOM specifications
- No substantive interoperability purpose
  - Just conveys a UID for the File Set (which no one cares about)
- But requires fewer changes to Parts 10/11/12
  - Allow unconstrained File IDs in URIs

DICOM specifications allowing empty directory

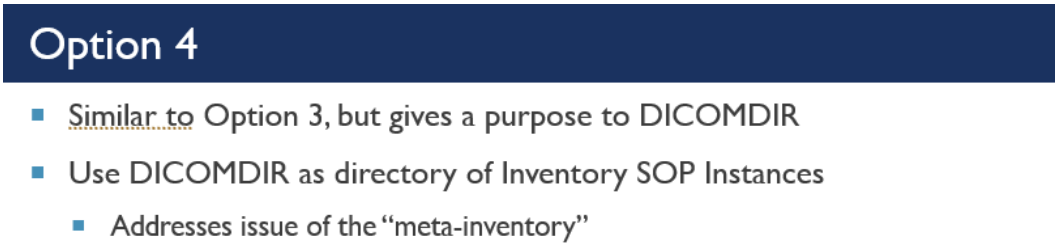
PS3.10 Section 8.6 Reserved DICOMDIR File ID

A single File with a File ID, DICOMDIR, shall exist as a member of every File-set. This File ID is made of a single Component (see Section 8.2 for the File ID structure). It contains the DICOM Media Storage Directory (see PS3.3 for detailed specification of the Basic Directory IOD), which includes general information about the whole File-set. This general information is always present, **but optionally the directory content may be left empty in environments where it would not be needed**. If the DICOMDIR File does not exist in a File-set, the File-set does not conform to PS3.10. The DICOMDIR shall not reference Files outside of the File-set to which it belongs.

PS3.3 Section F.3.1 Module Table

Module	Reference	Usage	Module Description
File-set Identification	F.3.2.1	M	File-set identification information
Directory Information	F.3.2.2	**U**	Directory Information followed by a Sequence of Directory Records. Note **The Directory Information Module is optional.** This Directory Information Module should be present in all but primitive environments where a directory is not needed. In this case, only the File-set Identification Information is present.

After seeing that, Option 4 came to mind, as follows:



One thing you must not do – post this as a re-documentation effort.

The most important thing is to define what we want to do, and not how we want to document it.
Need to have a way of identifying file access mechanisms not currently outlined in the standard. Need to avoid physical media constraints. Don't reuse the DICOMDIR mechanism, extremely difficult to build and doesn't scale well at all. Need a new way to encode this info. And then there is a granularity question.

Make up a new option - Option 5
An inventory stratified, if necessary, into multiple sub-files, with UIDs of whatever persistent entities it needs to be at the appropriate level, coded in DICOM format with a part 10 meta information header.
One stop shopping – migration engineer would only have to read one part of DICOM.

- Continue going through the Draft Supplement

  ftp://d9-workgrps@medical.nema.org/MEDICAL/Private/Dicom/WORKGRPS/WG33/2020/2020-09-16/SupXXX_00_ArchiveInventoryIODandServices.docx

In the Inventory Module done further refinement, basically clarifying when different things are required.

>Include Table C.12-x "Original Attributes Macro Attributes"		Recording and provenance of metadata changes. See C.YY.1.2.4
(0008,0054)	1C	AE Title from which referenced study may be retrieved Required if Retrieve URL (0008,1190) is not present. May be present otherwise.
(0008,1190)	1C	DICOMweb RS Origin Server from which referenced study may be retrieved Required if Retrieve AE Title (0008,0054) is not present. May be present otherwise.
(0008,0FxA)	1C	First part of URI for accessing SOP Instances within the study through a non-DICOM protocol. See C.YY.1.2.5.

		Required if Folder Pathname (0008,0FxB) or File Pathname (0008,0FxC) are present within this Sequence Item (including in subsidiary Sequence attributes)
(0008,0FxB)	3	Second part of URI, which when combined with Stored Instance Root URI (0008,0FxA) forms an access URI for a folder containing all SOP Instances for the Study. See C.YY.1.2.6.
(0008,0FxC)	3	Second part of URI, which when combined with Stored Instance Root URI (0008,0FxA) forms an access URI for an archive format file containing all SOP Instances for the Study. See C.YY.1.2.6.

C.YY.1.2.6 Folder Pathname and File Pathname

If all of the stored SOP Instances of the Study are in the DICOM File Format accessible through a non-DICOM protocol that provides read or move operations on a set of objects, e.g., all the files in a directory folder, Folder Pathname (0008,0FxB) concatenated with Stored Instance Root URI (0008,0FxA) provides the URI for such operations.

If all of the stored SOP Instances of the Study are in a single file archive format (ZIP or TAR) as specified in PS3.12, File Pathname (0008,0FxC) concatenated with Stored Instance Root URI (0008,0FxA) provides the URI for accessing that file.

Folder Pathname and File Pathname shall not be present if SOP Instances of other Studies are present in the folder or archive format file.

This references to this new part. It's not going to be defined in part 12. We will define it when necessary.

2 useful cases to mention:

- Supplement 211 - working on the zip
- 3D printing group – encapsulation of OBJ, the OBJ format references a separate material file and also potential texture map images

Separate the transactions related to this from the static content.

Patient			
>Patient's Name	(0010,0010)	2	Patient's full name
>Patient ID	(0010,0020)	2	Primary identifier for the Patient
>Include Table 10-18 "Issuer of Patient ID Macro Attributes"			Identifier of the Assigning Authority that issued the Patient ID
>Other Patient IDs Sequence	(0010,1002)	3	A Sequence of identification numbers or codes used to identify the Patient. One or more Items are permitted in this Sequence
>>Patient ID	(0010,0020)	2	An identifier for the Patient
>>Include Table 10-18 "Issuer of Patient ID Macro Attributes"			Identifier of the Assigning Authority that issued the Patient ID
>Patient's Birth Date	(0010,0030)	2	Birth date of the Patient
>Patient's Sex	(0010,0040)	2	Sex of the named Patient
>Other Patient Names	(0010,1001)	3	Other names used to identify the Patient
>Referenced Series Sequence	(0008,1115)	1C	Required if Inventory SOP Instance is produced in response to an Initiate Inventory Creation N-ACTION with Inventory Content Level (0008,0Fx9) value of SERIES or INSTANCE (see PS3.4 Section ZZ.2). May be present otherwise.

HS **Harry Solomon**
Does it matter that patient attributes will be encoded after the referenced series seq?

HS **Harry Solomon**
Specify relative to attribute in Scope of Inventory Sequence

Don't offer the options of studies and series – everything is at the instance level.
Had the discussion, agreed on hierarchical structure.
Can have a hierarchical structure, but when data is not actually hierarchical, you push the data down to the level that you need to.
You are migrating what you have not what you want it to be.

When you are doing a migration you need to be able to migrate all non-conformant data?
No, need to be able to supply all non-conformant data to the migrator to be able to make a decision on what to do with the data
Another way – flag those studies that have non-conformant data. But if there is a consistent pattern of deviance – you will be flagging a lot of studies.

What we want to do here is – functionality that handles bulk of the data, it's up to migration vendors to have any extra value-added services they want. We are trying to facilitate the process not to solve every problem in the process. We are trying though to make this the only interface to the PACS database, you shouldn't have to go back to the PACS database for other reasons, you may have to consult other sources of information like RIS or EHR.

5 NEW BUSINESS

6 Adjourned: 12:30PM ET

7 DATE AND TIME OF NEXT MEETINGS (Secretariat)

- Continue T-con meetings bi-weekly for the time being
- Next call is September 30, 2020 between 11:00 am and 12:30 pm ET

Reviewed by Counsel Peter Tolsdorf on 12/1/2020

NEMALINK CODE	09-WG33
SUBMITTED BY	Hull, Carolyn
SUBMITTED ON	11/16/20
LEGAL APPROVAL	Enter approval date.
UPLOAD LOCATION	Enter upload location.