

Minutes

<u>MEETING NAME</u>	09-WG33: WG-33 Data Archive and Management	
<u>MEETING PLACE/DIAL IN</u>		
<u>DATE & TIME</u>	Wednesday, July 8, 2020 11:30 am – 1:00 pm US ET	
<u>PRESIDING OFFICERS</u>	Matthew Bishop, UnityPoint Health Keith Eklund, Healthcare Tech Solutions	
<u>VOTING MEMBERS PRESENT</u>	Ambra Health Argentix Informatics Canon/Vital Images, Inc. Change Healthcare DesAcc EMEA Ltd. GE Healthcare Grafimedia Healthcare Tech Solutions Mach7 Technologies Laitek, Inc. Laitek, Inc. PixelMed Publishing SIIM SIIM	Ostrow, Daniel Silver, Elliot Dawson, Tim Ho, Kinson King, Graham Numan, Jouke Georgiadis, Pantelis Eklund, Keith Ulanov, Alexey Behlen, Fred Solomon, Harry Clunie, David Bishop, Matthew Henson, Kyle
<u>OTHERS</u>	Canon Medical Research USA, Inc. Canon/Vital Images, Inc. DeJarnette Research Systems GE Healthcare Hyland Software Laitek, Inc.	Kevin O'Donnell Whitby, Jonathan Wineke, Steve Nichols, Steven Ullrich, Mike Brown, Barry

Laitek, Inc.
London Health Science
Center
Mega Informatics Ltd.
Ochsner Health System
SIIM

Costea-Barluti, Razvan
Aizawa, Luiz
Fauquex, Jacques
Hayes, Matt
Carey, Cheryl

VOTING

AAPM

Bevins, Nicholas

MEMBERS

AAPM

Knazik, Shayna

ABSENT

European Society of
Radiology

Mildenberger, Peter

Laitek, Inc.

Sluis, Douglas

Mayo Clinic

Persons, Kenneth

Varian Medical Systems, Inc.

Schwere, Thomas

DICOM

Anna Zawacki, SIIM

SECRETARIAT

**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 one edit.

4 TOPIC ITEMS TO BE DISCUSSED (All)

- Review & discuss the New Work Item Proposal on Data Migration
In 7-8 meeting folder on FTP (please ask Anna for link)

Work item as posted to the server after a pretty heavy rework with Kevin O'Donnell. Clarifying a couple of things in the introduction: what PACS audit and QC might use as part of the production of this work item.

Discussing the limitations of the current standard, focused on what the limitations of implementation are:

- Optimization of the standard for network query for routine operations (doesn't really address the ability to get an inventory for the entire archive)
- Movement of data is constrained by the overhead of the C-MOVE service

Proposal includes an IOD for an inventory. Question in previous discussion how different it is from a manifest using key object selection – patient oriented manifest and this is an archive oriented manifest. Differences from the DICOMDIR are not conceptual but in terms of structure and scale.

IOD of the inventory

Mechanism to enable production of the inventory – asynchronously

Need to provide mechanisms to access that inventory – we already have DICOM non-patient object SOP classes such as color palettes and would be adding the inventory to the list of non-patient objects would make the access mechanisms a lot easier.

Direct access to the Part 10 compliant files.

Flagging whether there is metadata that has changed in the database and not been propagated to the stored files.

Because we are exchanging info via file exchange, we have to update the media storage service.

Limiting the scope to DICOM objects and not anything else, i.e. native CDA files.

Given that there is a bunch more things in the standard that will be affected. But we'll have to make updates to several Parts as outlined in the document.

Part 15 – should we update it for any of the research use cases, with any of the security requirements, need some security profile. Define access to the shared files.

Part 1 – will need updating because you provide a new access mechanism.

Some words in part 1 at least. The only difference from the current file service requirements – we'd have to discuss whether DICOMDIR is not required.

Using a shared file system. You can stuff that into part 12 possibly but need to mention in in Part 1.

We cannot gloss over the fact that it is network files access.

Add it as - Address overview of shared file access in Part 1.

Migrating a live archive – is it out of scope for this proposal?

No, it's one of the use cases.

It will be a technical detail that we'll be discussing later.

We need to mention that the intention is to be able to migrate a live archive.

Potential option is if you look at it as an asynchronous approach. You do your initial query and get back everything, and any changes after that come in as asynchronous notifications.

Delete the sentence about solving through UPS.

Also need to address a way to find all the instances/studies that were received after a particular point in time.

The goal is essentially to synchronize one archive with the other and achieve eventual consistency. Synchronization of part or the entire archive.

Patients are attributes of studies, so no need to add it.

Do you see an ability to migrate non patient objects? I would not put that in scope of this proposal.

We have hanging protocols, color pallets, templates of various sorts.

Is there a way of discovering a full set of inventory non patient object?

Do you need to include it? Let's put it out of scope for now.

Specifically say – non patient object migration is out of scope for now.

Orders – absolutely not, that's HL7.

Move approval of this work item for forwarding to the DICOM Standards Committee.

None opposed.

No abstentions.

Passed unanimously.

- Start reviewing open issues to be addressed in the Supplement Presentation in 7-8-2020 working group folder

| Consensus 5/27 – inventory / directory / metadata index

- Important use cases for inventory / directory / metadata index export, independent of storage format
 - Supporting data migration from current PACS
 - PACS audit/quality control
 - Research
- Functionally equivalent (+/-) to Query (C-FIND/QIDO) response for entire PACS database
 - But Query typically has restrictions on number of responses
 - Query requires synchronous network protocol response
- **Requirement: inventory / directory / metadata index of entire PACS database (or described subset), with results written to persistent objects (SOP Instance(s) / files)**

Don't have to close all the issues before we get to public comment but need to understand what the issues are and possible solutions.

Need to confirm what we will be calling it – inventory vs directory vs metadata index.

2

Consensus 5/27 – access to stored instances

- Use cases for direct filesystem access to SOP Instances
 - PACS replacement data migration
 - DICOMweb server decoupled from PACS
 - On principle, whatever the user organization wants to do – it's their damn data!
- Functionally equivalent (+/-) to network Retrieve (C-MOVE/WADO)
 - But Retrieve has performance issues for non-conventional access patterns
- **Requirement: inventory includes optional pointer for each archived object to files directly accessible through filesystem (or cloud object system)**
 - Support for target Part10 files, and Part10 files in ZIP or TAR

One variation – is this assuming they are in a standard transfer syntax or proprietary? If proprietary it will be identified and treat them as if they weren't there.

Does the inventory need to include the transfer syntax? Yes, and need to state it. Media profiles may include constraints of transfer syntax. Some basic list.

Why is the pointer optional? Because if the PACS doesn't make them available, then you don't have a pointer then go back to using C-MOVE or C-GET.

Is there any way to strongly encourage its presence? Does inventory have a value if you end up doing C-GET? Yes. We should strongly encourage it.

By default you always have C-GET or C-MOVE.

Interesting question whether the inventory should specify what access services are available? DICOMWEB, C-GET, C-MOVE. Inventory must include at least one means of access.

How persistent are the URLs? How long will they last, and can they change? Or is it unspecified? Leave it unspecified. It may go stale.

Could we set an expiry date for the inventory? There is no such thing as a permanent URL. Would it be a stretch to consider an additional interface that would consider everything that was changed? Give me everything that was changed or added after a certain date?

3

Consensus 5/27 – updated metadata

- Use case - updates to metadata attributes may not (yet) have been propagated into stored SOP Instance files
 - Recognition of common PACS implementation (descriptive, not prescriptive) - important technical mechanism supporting efficient archive operations when update of substantial numbers of objects is invoked, or objects have long access time
 - Metadata updates required for second party (non-PACS) direct access to SOP Instance files - requires applications directly accessing files to obtain current metadata separately
 - Acknowledged potential patient safety concern*
- **Requirement: inventory includes current metadata attributes that have not yet been propagated into stored SOP Instance files, and flag for this condition**
 - Possible service to request flushing updates to stored objects (see issue 9)

one extreme – so, if nothing has changed, the metadata may be completely empty.

9

Open Issue – metadata update invocation / flush

- Use case – in preparation for migration or other direct file access, external apps may want to request stored object files be updated to reflect current metadata
 - Asynchronous mechanism required – possible extended period to process
- **? Requirement: propagation of current metadata to stored instances may be invoked by standard service with asynchronous completion**
 - Mechanism specifies list of studies to process
 - Possible use of Unified Procedure Step Service (UPS), or variant of Storage Commitment
 - Requests may not complete successfully (cf. Storage Commitment)

So if that's the case maybe we should have a service that allows a user force metadata commitment to identify a study and say to the archive – I want you to update all the metadata in the stored instances to these studies. It's just like storage commitment.

That's a privileged command.

It's a cool idea, but may be impractical, don't know if a lot of people would implement it.

Don't dismiss the idea just yet. Keep it for now. But needs work.

4

Open Issue – inventory subset scope

- Use case – inventory may be needed for subset of archive
 - Studies in datetime range (e.g., since last inventory creation)
 - Studies with specific modalities, procedure codes, or procedure descriptions (e.g., for research)
 - Studies from particular locations (e.g., for business divestments)
 - Studies with metadata updated flag (e.g., to check for updates to specific study sets)
- **? Requirement: inventory includes description of scope of subset**

Can we leverage list for IHE schedule workflow as a list of mandatory matching keys?

We could make it burdensome or simple or leave it to people's discretion.

For these use cases, can you use the required matching keys we have.

Do we need to address modalities and SOP classes, or can we assume if you want to look at the series level attributes you have to do a series level specification? Patient would be very useful, cohort of patients.

Patient access – definitely. Study access to. Is series access something we have to think about?

Table this for later.

What about studies with no date?

Are you planning on using this service to find the stuff that you'll go back to clean up? Yes, QA is very important here.

What about duplicate UIDs? You can use inventory to find those.

Are we talking about the date and time of study or study's last update?

We need to include both.

Is there a requirement to be able to query on a storage location?

5

Open Issue – inventory content

- User application needs to select archive studies/series/instances relevant to its use case without retrieving/reading archived objects
 - Minimally for data migration use, potentially for other envisioned uses
 - Identifiers (UIDs, IDs, direct access pointers)
 - Clinical context (patient, order/accession, procedure)
 - Technical (modality, protocol, equipment, image/document type)
 - Possible: Cross-references to other studies/series/instances
- Inventory needs to record metadata attributes that were updated
- **? Requirement: specify levels of conformance on creator for minimal (Type 1) and rich (Types 2 & 3) inventory content**
 - Compare Storage Service Levels of Conformance for receivers ([Part 4 Section B.4.1.1](#))

Need identifiers

Clinical context to some level

Some amount of clinical details

Possibly any cross references from that study to other studies

Should we treat it as a query – if you don't ask for them, you don't get them

Don't include all –

Essentially this is kind of like your return keys – ask for what to include, ask for what you want, and we will hopefully include all that you want

Is this also a list of updated metadata attributes?

Do we want to be able for the inventory requestor to specify the degree of detail of the inventory – YES

As a practical choice – you'd want to have it, but you need to test it

Could you also specify as an implementor that you don't implement some level of all those profiles? Here are my supported profiles you may ask for them in this.

The baseline will allow a very min subset + anything else you ask for and the profiles will define the subset. Need to agree on a small set of profiles to help explore the issue.

You want to be able to indicate you want updates, but not sure you need to be able to indicate which updates you want.

The modified attributes record what the attributes used to be.

Migration takes a long time, you may want to catch up with the latest metadata, you may want to do a refresh. Having an option to have the updated metadata

that is separable from the inventory is desirable. Should be able to say: now re-run it to include what has changed since the last time.

Will it also record an attribute by attribute? No, but it will record study by study. You won't necessarily know what's been updated.

Definitive solution would be to serialize inventories with the studies. But that leads to inconsistencies between front end and store data. So, you crawl the files system to learn the metadata in there. Maybe that's a sweet spot, coupled with the query mechanism.

Remember that database is the source of truth.

II

Open Issue – distributed inventory structure

- Inventory instance may apply to described scope of subset
- Use case: inventory instances might be created and persisted for partitions of the archive
 - Allows incremental or parallel production of inventory
- **? Requirement: inventory supports inclusion by reference of other inventory instances for subsets of archive**
- See note in Part 3 Section F.1: *The directory information has been defined so that a future version may be extended to support the distribution of the directory information among a logical tree of several files (with the DICOMDIR file at its root). However in this version, the entire directory information is specified to be stored in a single File with a DICOMDIR File ID.*

What actually is this list? It would be a hierarchical list of studies including series and instances.

If we include series and instance level info, it may take us to it being as slow as C-FIND.

The model we need to encode is clear, but the representation of this model is not so clear.

How is this going to be chunkable?

Also need to consider – putting this in a database is nice, but you'd have to pick one database, there isn't a standardized database.

For the next call – pick up the discussion at Issue 5/6.

OLD BUSINESS

5 NEW BUSINESS

6 DATE AND TIME OF NEXT MEETINGS (Secretariat)

- Continue T-con meetings bi-weekly for the time being (at least until the New Work Item Proposal is submitted to DSC)
- Next call is July 22, 2020 between 11:00 am and 12:30 pm ET

Reviewed by Counsel Peter Tolsdorf on July 20, 2020

<u>NEMALINK CODE</u>	09-WG33
<u>SUBMITTED BY</u>	Hull, Carolyn
<u>SUBMITTED ON</u>	7/15/2020
<u>LEGAL APPROVAL</u>	Enter approval date.
<u>UPLOAD LOCATION</u>	Enter upload location.