DICOM Secretariat/MITA
1300 North 17th Street, Suite 900
Arlington, VA 22209, USA
+1-703-475-9217
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, January 20, 2021 | 11:00 am – 12:30 pm US ET

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

VOTING MEMBERS PRESENT

Argentix Informatix	Silver, Elliot
Change Healthcare	Ho, Kinson
DesAcc EMEA	King, Graham
Grafimedia	Georgiadis, Pantelis
Healthcare Tech Solutions	Eklund, Keith
Laitek, Inc.	Brown, Barry
Laitek, Inc.	Costea-Barlutiu, Razvan
Laitek, Inc.	Solomon, Harry
Society for Imaging Informatics In Medicine	Bishop, Matthew
Society for Imaging Informatics In Medicine	Carey, Cheryl

OTHERS

Hyland Software	Ullrich, Mike
Laitek, Inc.	Behlen, Fred
London Health Science Center	Aizawa, Luiz

VOTING MEMBERS ABSENT

AAPM	Bevins, Nicholas
Ambra Health	Ostrow, Daniel
Canon Medical Research USA	O'Donnell, Kevin
Canon/Vital Images	Dawson, Tim
Canon/Vital Images	Whitby, Jonathan

European Society of Radiology	Mildenberger, Peter
GE Healthcare	Nichols, Steven
GE Healthcare	Numan, Jouke
Mach7 Technologies	Ulanov, Alexey
Mayo Clinic Rochester	Persons, Kenneth
PixelMed Publishing	Clunie, David
Society for Imaging Informatics In Medicine	Henson, Kyle
Varian Medical Systems, Inc.	Schwere, Thomas

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.

4 TOPIC ITEMS TO BE DISCUSSED (All)

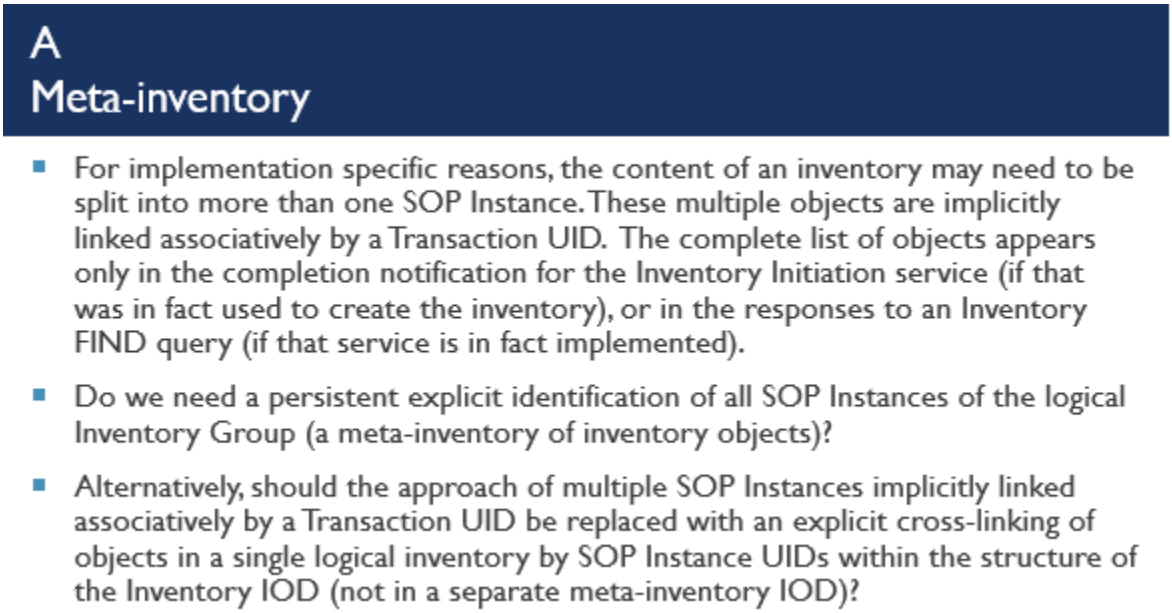
- Start addressing Open Issues, including those raised by WG-06 during their recent review:

ftp://medical.nema.org/MEDICAL/Private/Dicom/WORKGRPS/WG33/2021/2021-01-20/**issues_after_Jan_WG6.pptx**

- Here is the latest version of the Supplement:

ftp://medical.nema.org/MEDICAL/Private/Dicom/WORKGRPS/WG33/2021/2021-01-20/**Sup223_18_InventoryIODandServices.docx**

Discussed Open Issues



Issue is: wanting to have a way to deliver meta inventory on media. Hand over a bunch of files and meta-inventory to a new system which then imports the totality of the inventory. May not be essential. Comes down to knowing whether what you have is complete. What's the whole set? This is what will get implemented – those who have the most to gain from this standard are systems wanting to import bulk data. Another aspect of this is – it's possible there will be stable subsets of the archive that may already have inventories, if you need inventory of the whole archive system then we take it and reference it instead of re-processing it – presuming patient info gets updated along the way, not adding new studies.

B
Containers

- C.YY.1.2.7 and C.YY.1.3.1 specify that a container file (e.g., ZIP) at the Study or Series level must contain all the SOP Instances of that Study or Series, i.e., only a single container is allowed to be identified. (This is an optimization since the file location of each SOP Instance is recorded.) Is this an acceptable limitation?
- Specification says container files at Study or Series level may not include objects from other Studies/Series. Is this an acceptable limitation? Does this affect actual PACS implementations? How would this affect use of Study/Series containers vs. (unpacked) individual SOP Instance files in migrations?

Ultimately, you have a file location for each SOP instance, if you have an identification at the study and series level.

C
Non-DICOM protocol filepaths

- Only a single non-DICOM protocol is allowed for accessing the Instances in a Study (although different Studies can use different protocols). This is a consequence of splitting the URI into a root and a filename, under the assumption that all instances of a study will be on the same server, and we can avoid repeating the protocol and server name for each instance. Is this an acceptable limitation? Is that an issue for implementations that, e.g., may store different series of a study on different storage devices?
- Only a single non-DICOM protocol filepath is allowed for accessing an Instance. Is this an acceptable limitation? Are there implementations where an instance may be stored redundantly in two or more locations, and is it necessary for the inventory to record such multiple locations?

Question: is splitting the URI into a root and into specific file path on per instance basis appropriate? It's an optimization. We could specify a full URI for every instance, but when you get a 1000 slice CT study, you will be repeating the same 30 bytes of each URI – with a lot of studies it adds up to a lot of data. Or optimization by zipping inventory file later. Optimizations imply limitations on what you have. Can we effectively have root or default value and then per instance populate it or not populate it?
Do we just say one location is sufficient and if that fails through direct protocol you go to you DICOM CMOVE and let your PACS figure out where the redundant one is? Single instance could be replicated any number of times depending how customer configures it, magnitude of effort and time … prefer to list all redundant storage locations, rather not have it as a limitation. Plenty of reasons why an instance can occur in multiple

locations. Hashing, hierarchical storage, redundancy, different representations (compressed and not compressed version)? Question is whether those need to be listed? Always one is authoritative – the one you get when you do a C-MOVE. Multiple authoritative copies are a possibility too.

D
De-Identifiation

- The current (and only) de-identification profile, the Basic Application Level Confidentiality Profile, is oriented toward classic image SOP Instances, with PHI attributes in the top-level data set. It is not clear that that profile can be effectively applied to the Inventory IOD, where all PHI is within Sequence attributes. Further, it is not clear that the use cases for the Inventory are even appropriate for applying that profile. Is there is a need for a de-identification profile for Inventory instances? Which new data elements should be added to the tables of PS3.15?

Issue– once introduce new attributes, they tend to get re-used, if don't say: here is what the anonymization is for these attributes is now, and they get re-used in another situation, somebody may forget to put them in an applicable table. List of patient IDs not tied to patient data. Will have to go through the data elements carefully considering each one.

E
List of studies or patients

- The Scope of Inventory allows specification of a Study Instance UID List to be included in the inventory (e.g., to support research uses where applicable studies have been identified elsewhere). The Study Instance UID List has VR UI, which limits the list to 65534 bytes (~2000 UIDs). Is this an acceptable limit, considering that a client application that needs more can simply request multiple inventories? Should we define a new VR to allow 2^{32} bytes in the list?
- Similar limitation on Patient ID List

Falling into the same idea of implementation limitations. Seems fine.

F
Patient matching demographics

- The Patient IE in the Inventory IOD contains a basic set of patient demographics (name, sex, DoB). Would additional attributes be useful for patient matching during migration/consolidation of multiple repositories?

Issues of patient ID, assigning authority, mapping, qualify sequence – do we need to consider? Do we want to be able to have the production of the inventory do any simplified processing of the data while producing the inventory? I.e. telling it to add in this issue where a patient ID data where all of the studies you are inventorying? That could be possibly done as a post-inventory process. If we can define a couple of simple transforms, it may be useful.

G
Study inventoried once

- Section C.YY.1.1.5 requires that a given study shall appear only once among all SOP Instances of a group, and it shall be completely described in that one Item. Is that an issue for implementations that, e.g., may store different series of a study on different storage devices?
- What are the implications for the inventory consumer if a study is listed more than once in an inventory? What if the contents of those two listings are different, either due to the study being split to two storage subsystems (each inventorying its own data), or due to the study changing during the production of the inventory (with the two listings at different timepoints)?
- In other words, do we specify relaxed rules to simplify the producer of the inventory, or constraints to simplify the consumer?

If you impose constraints, will they be followed?

H
Scope of Inventory encoding

- The Scope of Inventory Sequence is specified with almost all attributes Type 2. Should these be specified with conditional inclusion (Type 1C), such that the data elements would be absent (rather than zero-length) if not used?

It could make sense if the purpose of the inventory is clinical studies then certain additional attributes could be filled in there.

Doesn't change which attributes we are allowing, it's just whether they are present and empty or not present at all.
Would be interesting to see a paper that looks at type 1 C implementations are there and if there are any rules out there.
If we say is not present, unless we have a particular use, it does provide a little more flexibility down the road if we decide to change conditions: extend, expand a bit.

J
Physician roles

- Inventory explicitly identifies data elements for the physician roles: Referring, Reading, Consulting, and Physician(s) of Record. Which of these are actually managed in PACS implementations? What is the use case for any of the Physician roles in migration? Which are required?

What do PACS actually manage? How many physician roles they actually manage? If they are doing reporting in the PACS, you may have the reading physician. Only seen referring and reading physician.

K
Protocols

- Non-DICOM file access protocols are left open. Should specific protocols be described normatively? If so, which ones?

Suggested wording: if you can make it work, use it.
They want something more normative, then for example, you can use this and that.

Are there any of the Open Issues that we want to dispose of right now?
Dispose of Scope of Inventory Encoding.

The group has moved on to discussing the latest Supplement #223 draft:
Group 6 objected to - Not for clinical use code – that section has been renamed and the attributes have been renamed. Now called – removed from operational use, and there is a reason for removal code sequence.

Did we ever consider the nearline case for migration? Calling it out separately somehow? If you need to switch tapes for particular studies – is there something around that that needs to be considered in your migration? Tapes is an interesting situation – is it at the series level or instance level? All of the above. PACS may have an identifier for a tape. Study could be split among multiple tapes. Open issue to come back to.

Some questions about original attributes macro, recording the prior values and provenance of updates, further discussion around 600.
What about attributes that an instance has that it's private or extended SOP classes?
Line 602 – attributes defined in an IOD, the values are authoritative.
The question is whether the receiver of the inventory needs to apply that tag to the SOP instances? If it doesn't know about it, it doesn't know it's used, and it doesn't apply it.
Do we say that we can add it anything else in the original attribute sequence?
We're using original attribute sequence within the item in each study.
If you have an update to a patient demographics that changes the demographics, but it only changes it in the inventory, perhaps an illustrative example would be good for Part 17.

Action: Harry to draft a new subsection here in the Operational Considerations for handling updates.

Next meeting: Discuss first open issue, everyone come prepared with options.

5 OLD BUSINESS

6 NEW BUSINESS

7 DATE AND TIME OF NEXT MEETINGS (Secretariat)

- Continue T-con meetings bi-weekly
- Next call is February 3, 2021 between 11:00 am and 12:30 pm ET

NEMALINK CODE	09-WG33
SUBMITTED BY	Hull, Carolyn
SUBMITTED ON	1/28/21
LEGAL APPROVAL	2/3/2021
UPLOAD LOCATION	Enter upload location.