Standardizing AI Annotations
The DICOM Way

David Clunie, MBBS, FRANZCR(Ret), FSIIM
PixelMed Publishing, LLC
Disclosures

- Editor of the DICOM Standard (NEMA Contract)
- Owner of PixelMed Publishing, LLC
- Author of book on DICOM Structured Reporting
- Consulting for GE, Carestream, MDDX (Bioclinica), Curemetrix, HCTS, Hologic
- Supported by NIH U24CA180918 QIICR, NCI Leidos BOA 29XS219 Task Order #05
Outline

- Background
  - radiologists and annotations
- Annotation use cases
  - in general and for AI
  - definition and different types of annotation
- DICOM Encoding
- Workflow and DICOM
- Beyond radiology
- Gaps and future improvements
I felt able to concurrently experience my proprioceptive awareness combined with other elements, such as a sense of narrative!
A Picture Is Worth A Thousand Words:

Needs Assessment for Multimedia Radiology Reports in a Large Tertiary Care Medical Center

Lina Nayak, MD, Christopher F. Beaulieu, MD, PhD, Daniel L. Rubin, MD, MS, Jafi A. Lipson, MD

Rationale and Objectives: Radiology reports are the major, and often only, means of communication between radiologists and their referring clinicians. The purposes of this study are to identify referring physicians’ preferences about radiology reports and to quantify their perceived value of multimedia reports (with embedded images) compared with narrative text reports.

Materials and Methods: We contacted 1800 attending physicians from a range of specialties at large tertiary care medical center via e-mail and a hospital newsletter linking to a 24-question electronic survey between July and November 2012. One hundred sixty physicians responded, yielding a response rate of 8.9%. Survey results were analyzed using Statistical Analysis Software (SAS Institute Inc, Cary, NC).

Results: Of the 160 referring physicians respondents, 142 (89%) indicated a general interest in reports with embedded images and completed the remainder of the survey questions. Of 142 respondents, 103 (73%) agreed or strongly agreed that reports with embedded images could improve the quality of interactions with radiologists; 129 respondents (91%) agreed or strongly agreed that having access to significant images enhances understanding of a text-based report; 110 respondents (77%) agreed or strongly agreed that multimedia reports would significantly improve referring physician satisfaction; and 85 respondents (60%) felt strongly or very strongly that multimedia reports would significantly improve patient care and outcomes.

Conclusions: Creating accessible, readable, and automatic multimedia reports should be a high priority to enhance the practice and satisfaction of referring physicians, improve patient care, and emphasize the critical role radiology plays in current medical care.

Key Words: Multimedia reports; radiology reporting; digital images; communication; radiology practice.
©AUR, 2013
Structured Radiology Reporting: Are We There Yet?\(^1\)

Given the prominent role that information technology will play in the future of health care delivery, the potential benefits of structured reporting systems now seem more relevant than ever. These systems may lead to rapid cohort design. The same 25 brain magnetic resonance (MR) imaging cases were reviewed in two distinct phases by two separate groups of residents: a control group and an intervention group. The MR imaging cases contained a representative
No, we are not f*cking there yet...

and other universal truths of parenting

Andrew Willis • Robin Swift
HOW DO WE GET THERE FROM HERE?
SUCCESS!

Idea, Plan, Activities

Analysis

Product

Consumers

Internet

Think outside the box

Work

E = % = $
Use Cases – Non-AI or AI, Use versus Re-Use

Clinical Use Cases
- Non-AI
  - communicate – rad-tech, rad-rad, radiologist-clinician, multidisciplinary team meetings
  - preserve state – for priors (where, what, what size, ...)
  - legal record
- AI
  - automated result output prior to human read (traditional mammo CADe, triage, priority)
  - human defined selection for targeted AI

Re-use Use Cases
- Non-AI
  - retrospective and prospective research
  - education
- AI
  - Non-AI annotations used for AI training/testing
  - AI annotations used for a different AI application training/testing
Use Cases – What is an “Annotation” anyway?

- **Annotation**: “a note of explanation or comment added to a text or diagram”
- **Label**: “a classifying phrase or name applied to a person or thing [especially one that is inaccurate or restrictive 😞]”
- **NOT Markup**: “a set of tags assigned to elements of a text to indicate their relation to the rest of the text or dictate how they should be displayed”

Distinguish semantics from visual representation

“meaning” for machines not just humans
Use Cases – Level of Granularity

- Patient/Case
- Imaging Study
- Series/Acquisition
- Image
- Frame (pixel data array at one place in space/time/...)
- Region (“of interest” – ROI)
- Single point (label each/every voxel/pixel)
Use Cases – Qualitative, Quantitative

- **Qualitative**
  - categorical – shape = round, ...

- **Ordinal**
  - roundness – scale of 1 to 5

- **Quantitative**
  - morphology – size (diameter, volume, ...)
  - quantity – signal intensity, attenuation coefficient, ...
  - numerical features – entropy of GLCM, fractal dimension, ...
  - on transformed variants – registered, resampled, filtered, wavelet, ...
  - units – absolute and relative (to what – reference region, population)
  - derivation – mean, max, ...
  - method – model, fitting, sampling, binning, ...

#CMIMI18
Use Cases – Concept Representation

- Single concept
  - e.g., “round shape”; “42”

- Name-value pair
  - e.g., “shape” = “round (generic)”, “round shape”; “ultimate question” = “42”

- Coded versus text
  - SRT:M-02100 (SCT:42700002) v. languages: “qaab wareegsan”, “圆形”
  - synonyms (“round” v. “circular”), case, punctuation (“Shape, round”), …

- Same concepts in different classifications
  - SCT:42700002, NCIt:C48348, RadLex:RID5799
  - Metathesaurus (mapping): UMLS:C0332490
Use Cases – Localization Representation

- Patient/Case – Patient/Specimen Identifier
- Study, Series, Image – Unique Identifiers (UID, UUID)
- Frame – UID + frame number/offset
- ROI
  - contours - image (2D) or space (patient, 3D) relative
  - temporal coordinates (frame #, relative/absolute time)
  - segmentations (bitmap, partial occupancy, probability)
- Single point (each/every voxel/pixel)
  - parametric maps (bitmap, scaled integer, floating point)
- Transformations
  - rigid (affine), non-rigid (deformation field, spline, ...)

#CMIMI18
SO WHAT?

WHO CARES?

I've decided to stop caring about things.

If you care, you just get disappointed all the time. If you don't care, nothing matters, so you're never upset.

From now on, my rallying cry is, "So what?!"

That's a tough cry to rally around.

So what?!
Where DICOM fits in

- **DICOM “in”**
  - obviously all (radiology) images start out as DICOM
    - can be rearranged from multiple single files to single multi-frame DICOM file (happier algorithm developers)
  - “metadata” in DICOM “headers” are “annotations” too
    - e.g., Series Description = “T1 axial post-Gd”, B value = 1000
    - can be better structured/coded (retrospectively), e.g., Acquisition Contrast = T1
    - phase to cleanup/canonicalize multi-site data DICOM attributes
    - do not discard known-safe vendor private data elements during de-identification – may be useful for unanticipated re-use cases
Where DICOM fits in

- **DICOM “out”**
  - **All types of annotation** (including processed images) can be shared as DICOM
  - **Pros:**
    - allows re-use of clinical imaging/annotation systems
    - can be stored/shared/indexed in off-the-shelf DICOM archives (e.g., TCIA)
    - can be created/viewed/analyzed by (some) OTS DICOM tools, viewers, ...
    - not just a bunch of poorly labeled/organized files on somebody’s disk
    - self-describing/identifying – contains (pseudonymous) identifiers
  - **Cons:**
    - requires more attention to de-identification (if clinical origin)
    - requires use of DICOM toolkits/libraries to create/access
    - requires more attention to preservation/propagation of “composite context” (identities and UIDs) in processing pipeline (e.g., to restore identifiers from images to results)
    - more complex and arcane than making it up as you go along
Confessions of a Content Creator: I Don't Care About Data
DRINK?

NOPE

MARK EWBLE
Annotation (interoperability) matters now

- **Previously:**
  - little incentive to annotate
  - few tools to create or view annotations
  - annotation interoperability was a low priority for product managers
  - presentation rather than semantics were the priority for annotation tools

- **Now:**
  - semantic annotations have (real monetary) value beyond primary use case
  - recognition of existence of unanticipated re-use cases
  - annotations are expensive to create/recreate retrospectively
  - more expensive to process if proprietary rather than OTS standard
  - AI-generated annotations need to be interoperable for display
  - “interactive” AI requires interoperable annotation exchange
  - AI vendors unlikely to be the same as scanner/PACS vendors – mix and match
Bad ways in DICOM, historically:

- Burned in graphics and text
  - including screen shots
- Overlay graphics and text
  - in pixel data, header or separate object
- Presentation states
  - still only graphics and text (no semantics)
  - currently very popular in clinical PACS
Good ways in DICOM – standardized for ages, variable use

- Structured Reports (SR)
  - tree of codes, numbers, 2D, 3D & temporal coordinates, references, ...
  - basis for Ultrasound, Cardiovascular, Mammo CAD, radiation dose
  - Key Object Selection (KOS) – flags key images with text/coded label

- Segmentation (SEG or DSO)
  - rasterized – bitmap, probability, occupancy; coded property/anatomy; ROI, atlases (i.e., pixel level categorical annotation), ...
  - surface mesh (rarely used)

- Radiotherapy Structure Sets (RTSS)
  - 3D coordinates, some component semantics, few quantities (volume)
  - widely used in RT planning and re-used in workstations, e.g., for PET
Relatively new things in DICOM

- Real World Value Maps
  - coded way to describe voxel values (beyond Rescale Type)
  - retrofitted to all existing DICOM images
  - form of “annotation” that makes pixel values semantically meaningful

- Parametric Maps
  - RWVM combined with floating point or scaled integer pixels

- Second-generation Radiotherapy annotations
  - Conceptual Volumes – “grammar” for combining contours, segmentations
DICOM and Annotations – Then to Now

- Related DICOM IODs
  - Fiducials
    - markers with shape and location
  - Registration
    - rigid
    - deformable
    - well-known frames of reference (e.g., atlases)
Putting it all together ...

Current DICOM Images from Modality

Analysis Workstation

Previous DICOM SR etc.

DICOM SR

DICOM Segmentation

DICOM Registration

DICOM Real World Value

DICOM Parametric Map Images

Previous DICOM Images from PACS

PACS Store, Distribute and Review
Putting it all together ...

Current DICOM Images from Modality

Previous DICOM Images from PACS

Analysis Workstation

DICOM SR

DICOM Segmentation

DICOM Registration

DICOM Real World Value

DICOM Parametric Map Images

PACS Store, Distribute and Review
Are you too busy to improve?

No thanks!

We are too busy

Håkan Forss @hakanforss http://hakanforss.wordpress.com
This illustration is inspired by and in part derived from the work by Scott Simmerman, “The Square Wheels Guy” http://www.performance-management.com/
Too busy? – No problem, we supply tools

- **Store and regurgitate**
  - OTS commercial and open source DICOM archives as well as PACS
  - most modern systems will handle any new or old DICOM object
  - unlike ancient PACS, which rejected things they could not view
  - this increasing flexibility driven by “vendor neutral archive”, “universal viewer” and “deconstructed PACS” phenomena

- **DICOM annotation objects are just like images**
  - some actually are “images” (segmentations, parametric maps)
  - non-images all share common “composite context” encoding (e.g., patient/study/series/equipment identification/description)
Too busy? – No problem, we supply tools

**Creation**
- long history of ultrasound, mammo CAD, radiation dose, key object SR authoring
- increasing SEG and SR support in toolkits (esp. open source)
- extraction/propagation/merging of composite context
- merging descriptions of metadata created in XML or JSON into DICOM object content (included coded descriptions) – easier than hand-coding in programming language

**Transcoding**
- of other formats, proprietary, academic, and project-specific
- other structured annotation formats into DICOM SR (e.g., AIM via PS3.21 mapping)
- segmentations and label maps in other formats into DICOM SEG
- parametric maps in other formats into DICOM Parametric Maps
- single frame DICOM image sets into multi-frame single DICOM files
Too busy? – No problem, we supply tools

**Access**
- annotation objects can be queried/retrieved using traditional DICOM network protocols widely supported by many toolkits
- DICOMweb (WADO-RS, QIDO-RS, …) protocols increasingly supported by archives and are just like any other HTTP request (can use curl, postman, as well as DICOM toolkit utilities)
- DICOMweb payloads (XML and JSON metadata) and separate bulk data access increase accessibility and simplify parsing

**Viewing and Analysis**
- toolkit support for extracting metadata and structured content into generic XML, JSON, CSV
- more viewers support display of DICOM SEG superimposed on underlying reference images
- some viewers can display content of DICOM SR objects rendered into hierarchical plain text, HTML or PDF – few (if any) can “tabulate” content into user friendly form
- voice dictation systems have long been able to ingest DICOM SRs, extract content and make available “merge fields” in dictation templates
Interoperable communication of quantitative image analysis results using DICOM standard

Andrey Fedorov1, Daniel Rubin2, Jayashree Kalpathy-Cramer3, Justin Kirby3, David Clunie4, Michael Orkmen5, David Flade5, Pattanasak Mongkolwat6, Rajesh Venkataramani7, Jan Bertling8, Steve Pieper1, Ron Kikinis9
1Brigham and Women’s Hospital, 2Stanford University, 3Massachusetts General Hospital, 4NCI Fredrick, 5PixelMed Publishing, 6OpenConnections GmbH, 7Brainlab, 8Mahidol University, 9Eisen Medical, 10Hemes Medical, 11Isomics Inc

Introduction
As quantitative imaging (QI) is gaining momentum in research and commercial platforms, it becomes important to support its usage scenarios:
• Clinical workflows: storage of the analysis results on PACS alongside the imaging data; longitudinal followup of the patient with quantitative imaging across workstations.
• Research workflows: validation of imaging biomarker analysis tools; community repositories of the analysis results; secondary analysis of data.
Various types of derived data important in quantitative imaging research include image annotations (points, distance measurements, contours, labeling of image voxels), parametric maps and numeric results of the quantitative measurements.
Image segmentation is a preprocessing task concerned with defining a region of interest for subsequent analysis and quantitation. It is therefore of critical importance to support interchange of the segmentation results.
Digital Imaging and Communication in Medicine (DICOM) is the standard used ubiquitously for communicating imaging data. Although DICOM provides the means to also describe derived image-related information, thus far it has found very limited acceptance in the quantitative imaging community.
As a step towards improving QI analysis results interoperability, we investigate the use of DICOM Image Segmentation Storage SOP Class (DICOM SEG) for communicating image segmentation results.

Workstation Support of DICOM SEG
Workstations evaluated (commercial products are in italics):
• 3D Slicer X, http://slicer.org
• ePARD v1.7, http://epard.stanford.edu
• Brainlab PDM 2.2
• Siemens syngo via VA36A_F06

DICOM for Image Segmentation Storage
DICOM SEG is the preferred way of communicating segmentations represented as labeled voxels. Some of the important features supported include:
• size efficiency with multi-frame storage and bit encoding
• structured terminology for encoding semantics
• binary and fractional segmentation (e.g., probability maps)
• encoding of the presentation (color)
• multiple voxel occupancy

Being part of the DICOM object “family”, integrates with other types of data:
• patient and study composite contexts, frame of reference maintained
• references source image data
• can be referenced from the measurement documents (DICOM SR)

Know of a workstation/toolkit supporting DICOM SEG but not listed?
Want to learn more or get help adopting DICOM SEG?
Have sample DICOM SEG datasets you would like to contribute?

Please email andrey.fedorov@gmail.com

Support in Developer Toolkits
DCMTK - DICOM Toolkit (C++, free open source)
http://dcmtk.org

PixelMed Toolkit (Java, free open source)
http://www.pixelmed.com/pixelmedsoftware/

classes providing abstractions to support interaction with SEG

Support in OpenSource:
• ePARD
• 3D Slicer
• Brainlab
• ClearVox
• AIM

Acknowledgments
This work was supported in part by the National Institutes of Health through the grants U24 CA158315 (QBRC), U01 CA142535, U01 CA189214 and U01 CA140208.
Crowds Cure Cancer project at RSNA 2017 (http://doi.org/10.7937/K9/TCIA.2018.OW73VLO2)

- Booth and mobile app to find mass and draw longest diameter on liver, kidney, lung and ovarian tumors on TCIA DICOM images
- Sent a CSV file with all annotation coordinates and lengths, and metadata including patient, study, series, instance identifiers of images

Created two XSLT stylesheets
- extract cells from CSV table into XML files, one for each row (annotation) – driven by column headers
- convert extracted row cells into DICOM SR instance of DICOM TID 1500 Measurement Report encoded as PixelMed toolkit specific XML format

- Applied existing open source PixelMed toolkit XML to DICOM SR converter
- Validated DICOM SRs that had been created
  - ran automated DICOM SR validator in PixelMed toolkit (knows about TID 1500 since QIICR project)
  - visually inspected dumps of DICOM SR content with PixelMed and dicom3tools toolkit utilities and rendered SR coordinates and measurements on images using PixelMed tool

- One morning’s work (leveraging familiar toolkits and techniques and stylesheets used from other projects as a starting point)
- Plan is to to put back into TCIA along with the already publicly accessible images
| order | anatomy | seriesID | parentID | instanceID | length | start_x | start_y | end_x | end_y | annotator | id | srid | date_x | date_y | date_z | StudyID | StudyDate | StudyTime | SOPClassUID |
|-------|---------|----------|----------|------------|--------|---------|---------|-------|-------|-----------|-----|------|--------|--------|--------|----------|----------|-----------|-----------|-------------|
| 1     | Renal   | D11752   | 11752    | 0          | 0      | 0       | 0       | 0     | 0     | 11752     | 0  | 0    | 0      | 0      | 0      | 0        | 0        | 1         | 1         |

Courtesy of Jayashree Kalpathy-Cramer
CONTAINER: (126000,DCM,"Imaging Measurement Report") [SEPARATE] (DCMR,1500)

1.1: HAS CONCEPT MOD: CODE: (121049,DCM,"Language of Content Item and Descendants") = (eng,RFC5646,"English")
   >>1.1.1: HAS CONCEPT MOD: CODE: (121046,DCM,"Country of Language") = (US,ISO3166_1,"United States")

1.2: HAS OBS CONTEXT: PNAME: (121008,DCM,"Person Observer Name") = "accomplished_peafowl"

1.3: HAS CONCEPT MOD: CODE: (121858,DCM,"Procedure reported") = (41800-1-LN,"CT Abdomen")

1.4: CONTAINS: CONTAINER: (111028,DCM,"Image Library") [SEPARATE]
   >>1.4.1: CONTAINS: CONTAINER: (126200,DCM,"Image Library Group") [SEPARATE]
      >>1.4.1.1: CONTAINS: IMAGE: = (1.2.840.10008.5.1.4.1.1.2.1.3.6.1.4.1.14519.5.2.1.9203.4004.26881842288818573226516023762)
         >>>1.4.1.1.1: HAS ACQ CONTEXT: CODE: (121139,DCM,"Modality") = (CT,DCM,"Computed Tomography")
            >>>1.4.1.1.2: HAS ACQ CONTEXT: DATE: (111000,DCM,"Study Date") = "19870620"
            >>>1.4.1.1.3: HAS ACQ CONTEXT: TIME: (111061,DCM,"Study Time") = "135823"

1.5: CONTAINS: CONTAINER: (126010,DCM,"Imaging Measurements") [SEPARATE]
   >>1.5.1: CONTAINS: CONTAINER: (125007,DCM,"Measurement Group") [SEPARATE]
      >>1.5.1.1: HAS OBS CONTEXT: TEXT: (112839,DCM,"Tracking Identifier") = "5b6eb473013275542d29985a3d0fbb00"
            >>1.5.1.2: HAS OBS CONTEXT: UUIDREF: (112840,DCM,"Tracking Unique Identifier") = "1.3.6.1.4.1.9621.1.0.0.0.1535644357.22655.1"
            >>1.5.1.3: HAS CONCEPT MOD: CODE: (G-C8E3,SRT,"Finding Site") = (T-71000,SRT,"Kidney")
            >>1.5.1.4: CONTAINS: NUM: (G-D7E,SRT,"Length") = 66.43856134 (mm,UCUM,"mm")
                >>>1.5.1.4.1: INFERRED FROM: SCOORD: = POLYLINE {172.835357666016,270.06406914062,133.798889160156,343.045318603516}
                    >>>1.5.1.4.1.1: SELECTED FROM: IMAGE: = (1.2.840.10008.5.1.4.1.1.2.1.3.6.1.4.1.14519.5.2.1.9203.4004.26881842288818573226516023762)
Image size: 1224 x 1584
View size: 1246 x 751
WL: 127 WW: 255
X: -623 px Y: 320 px

Imaging Measurement Report

Concept Modifier: Language of Content Item and Descendants [Annex 1]
Observation Context: Person Observer Name = accomplished_peafowl
Concept Modifier: Procedure reported = CT Abdomen (41806-1, LN)

Image Library

Image Library Group

CT image

Acquisition Context: Modality = Computed Tomography (CT, DCM)
Acquisition Context: Study Date = 1987-06-20
Acquisition Context: Study Time = 13:58:23

Imaging Measurements

Measurement Group

Observation Context: Tracking Identifier = "5b6eb43b1d5175942d2990fe3d0b600"
Observation Context: Tracking Unique Identifier = 1.3.6.1.4.1.5962.1.6.0.0.0.1535644535.22655.1
Concept Modifier: Finding Site = Kidney (T-71000, SRT)

Length:
66.43856134 mm
Inferred from: Spatial Coordinates [Annex 2]

Annex

Annex 1

Language of Content Item and Descendants:
Annotation workflow and use cases

- Traditional “box under the table/in back room” (“headless”)
  - acquire images -> process -> send both to PACS for reader
  - “push workflow” (“unmanaged”) familiar from Mammography CADe
  - cloud variants and large data volumes add security/timing challenges
  - data to algorithm or algorithm to data

- Interactive (virtually “embedded”)
  - user views images -> selects task/frames/ROI -> requests CAD -> see result
  - user expects immediate response (or likely will not use)
  - responsiveness may depend on prearranged proximity of data & algorithm
  - standard payload not only for results but request as well
  - standard protocol for command and control of request/progress/complete
Annotation workflow standards

- Standard payload for images in/out & results (DICOM)
  - for visualization – “pretty picture” screen shots and presentation states
  - semantically meaningful – segmentations, parametric maps, SR

- Standard protocols to choose from:
  - DICOM – Unified Procedure Step (UPS) and DICOMweb UPS-RS variant
  - IHE – Invoke Image Display (IID) – HTTP request to view specified study(ies)
  - HL7 – synchronized applications – CCOW and now FHIRcast
  - DICOM Application Hosting – SOAP-based, never popular; revisit RESTfully?

- “Push” unmanaged request – image timeout, UPS N-CREATE, IHE IID
- “Pull” from worklist – UPS C-FIND or UPS-RS, do work and update status
- Fetch payload – DICOM C-MOVE/C-GET or WADO-RS (metadata)
- Return result – DICOM C-STORE or STOW-RS
- Security – transport (TLS), authentication (Kerberos, SAML, JWT, OAUTH)
An illustration of how digital slides are stored in a pyramid structure.

http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0030783
Extending Segmentation and Parametric Maps to tiled images for Whole Slide Imaging (CP 1830)
- Plane Position (Slide) macro instead of Plane Position (Patient)

DICOM query or DICOMweb RESTful API for annotation access?
- “spatial” queries, e.g., all annotations that intersect defined frame/region

More compact representation of very large numbers of contours?
- e.g., all nuclei, all membranes in WSI (versus SEG bit-plane representation)

Explicit Label Map rather than Segmentation bit planes?
- in SEG object have one bit-plane per segment (label), each as a frame
- in traditional label maps, one multi-valued voxel where each voxel is an index whose value represents a segment (label)
- same semantics but “gratuitously” different representation – convenience?

Always need more coded concepts
- e.g., more texture features – all of IBSI (CP 1764)
I may not be there yet,
but I am closer than I was yesterday.
How long until you are assimilated?

“we will add your biological and technological distinctiveness to our own”

“your culture will adapt to service us”

“resistance is futile”