A DICOM Prototype in XML with Relationships
Robert C. Leif, Ph.D. & Stephanie H. Leif M.S.
Newport Instruments
San Diego, CA
www.cytometryml.org

Pathology Informatics 2012
Standards Paucity

• The most cost effective way to design a standard is to reuse an existing one.
  — Change the syntax not the semantics.

• Benefit from the existing design, experience, and documentation.

• In the case of software, design can cost more than coding. Maintenance can cost more than both.

• Evolution not revolution.
DICOM & RESTful Web Services (for a server)
Tim Dawson & Jim Philbin (HL7, Baltimore, ‘12)

• Retrieve DICOM (studies, series, or instances by UID)
• Retrieve all metadata in one XML set
• Retrieve bulk data (including pixels) in one multi-part MIME message
Organization & Hierarchy

- All of the metadata from the series together with all of the instances’ metadata are retrieved together.
  - The metadata of specific objects is retrieved on the client

- Binary or mixed data are subsequently retrieved
  - based on the content of the metadata of specific objects
Pathology-Cytometry Data Separation

- Specimen preparation should be separated into:
  1. That which is common to all instances, which is stored with the series description.
  2. That which is specific for each instance, which is stored with the instance.

- The instrument description should be separated into:
  1. Those items that are unchanged for all of the instances, which are stored with the series description.
  2. The settings and configuration that change between instances, which are stored with their specific instance.
Series & Instances Organization

Instance

Relations

- Header
- Protocol

Metadata

Metadata that describes the binary data files in the instance

Binary Data

Series

EPUB ToC

- Header
- Protocol

Metadata

Common metadata that describes all of the binary data files in the Instances

Specimen Prep.

Instrument
Proposed Metadata Organization

Series Relation List

Instance 1 Ref.  Instance 2 Ref.  Instance 3 Ref.  Instance n Ref.

Instance Relation List

n=1

Most Relevant Data File Ref.

n=1..25

Data File with Relation Ref.

Choice

Metadata File with Relation Ref.
DICOM in XML Dilemma

• XML Schema Definition Language (XSDL) is required for description of objects, such as in databases.
• Resource Description Framework (RDF) required for description of relations between objects.
  --Jules Berman was absolutely correct, a standard or similar document should show relationships.
• XSDL and RDF have different schema languages!
• XHTML5 (W3C) cannot be usefully interfaced with XML — Polyglot Markup
Solution to XML Dilemma

• Create the equivalent of RDF in XSDL.
• Requirement: optimize readability and flexibility.
• RDF Triple becomes: Subject, Verb, and Object
  – Old Idea
• Creation of data structure required replacement of inflexible attributes with elements.
  – This increased verbosity, but provided the full power of data structures.
Image3 is Diagnostic, since it is the Most_Relevant_Image_Reference which is the Data_Of_Greatest_Interest and is an Instance which is Processed_Data.

Image3 is descendant of Image2 and Image1.

Image3 is classification-results of gate_and_linear_descriminant metadata.

Fig 1. which is based upon the Image3 file, shows a dividing cell population.
Metadata Relation

The significance of gate_and_linear_discriminant metadata is Informational.

The role of the Metadata_Reference is a classifier.

The gate_and_linear_discriminant metadata is a classification description of Image3.

This is an imagined algorithm.

The data is hypothetical
Relation XML Code Example

<relations_image:Subject>
  <relations:Self>Image3</relations:Self>
</relations_image:Subject>

<relations_image:Predicate_Phrase/>
  <!--Each Predicate Phrase consists of a verb and one or more objects-->
<relations_image:Verb_Phrase_Std>
  is descendant of
</relations_image:Verb_Phrase_Std>
Code Continued

<relations_image:Object>
  <links:Curie_Link>
    <links:a href="test:File2.xml">
      Image2
    </links:a>
  </links:Curie_Link>
</relations_image:Object>

<relations_image:Object>
  <links:Curie_Link>
    <links:a href="test:File1.xml">
      Image1
    </links:a>
  </links:Curie_Link>
</relations_image:Object>
Results & Conclusions

• Multiple CytometryML schemas have been created.
  – Need to be revised according to WG 26 and WG 27 supplements.

• Solidify that DICOM VRs and Tags will be represented by attributes.
  – Provides backward compatibility.
Results & Conclusions Continued

• Feasibility of extending DICOM and eventually translating DICOM into the XML Schema Description language (XSDL) established.

• XSDL1.1 has features that will facilitate the creation of a standard and improve its quality.
  – Includes assertions and supports use of objects
  – The XSDL1.1 Validation software is immature.

• XSDL can serve as a software design language for DICOM
Sources

• DICOM
  – Existing
  – WG 26 Pathology
  – WG 27 XML and the PACS Restful Services for Input to and output from the PACS
• XML and binary data
Sources Continued

  – Compensation
  – Gating
• EPUB Zip Container (Book Standard)
  – XHTML5, XML, Binary
• XHTML5
  – Structured reports & Data Entry Forms
  – Schema Olivier Ishacian
    http://www.xmlmind.com/, olivier@xmlmind.com