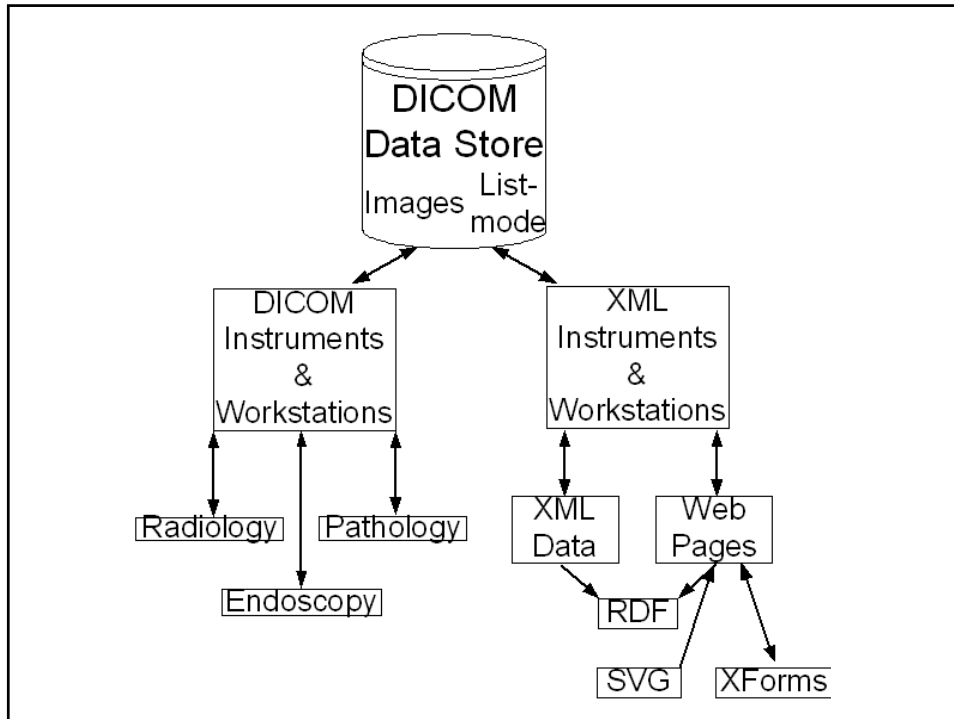


**CytometryML, a data standard,
which has been designed to
interface with other standards**

**Robert C. Leif,
Newport Instruments
rleif@rleif.com
www.newportinstruments.com**

Outline

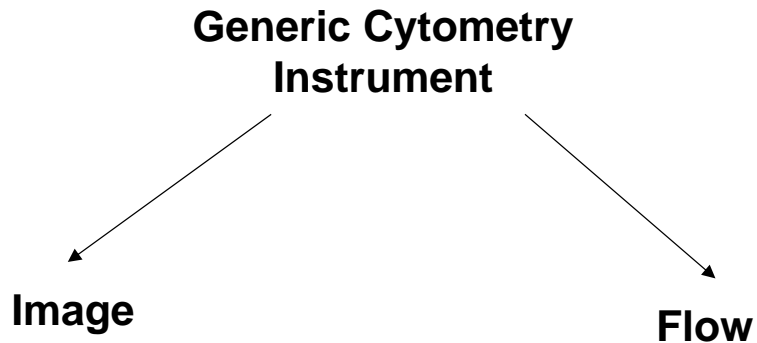
- Model of Analytical Cytology Informatics
- One Cytology Standard
- Resource Description Framework (RDF)
- Health Level 7 (HL7)



Benefits

1. The combination of image and list-mode data obtained with either a digital microscope or an imaging flow cytometer will be kept together.
2. Laboratories that use both modalities will be able to store their data in the same place.
3. The storage of the data in a common server system will satisfy the needs of many of the pathologists and physicians from other specialties and will be consistent with the interests of the hospital informatics department.
4. The part of the data that is created by a human can be entered using a XML standard form description language, XForms, which has provisions to minimize data entry errors.

Only 1 Cytometry Standard



Simple (helper) Datatype

```
<simpleType name="Platform_Type"
  id="Platform_Type">
  . <restriction base="token">
  .   <enumeration value="Upright"/>
  .   <enumeration
  .     value="Inverted"/>
  .   <enumeration value="Plainer"/>
  . </restriction>
  . </simpleType>
```

Generic Instrument

```
<complexType name="Instrument_Type">
. <sequence>
. <element name="Item_General_Info"
  type="item:Item_General_Info_Type"/>
. <element name="Objective"
  type="optics:Optic_Type" minOccurs="1"
  maxOccurs="7"/>
. <element name="Condenser"
  type="optics:Optic_Type" minOccurs="0"
  maxOccurs="10"/>
. <element name="Platform"
  type="instr:Platform_Type"/>
```

```
<choice minOccurs="1" maxOccurs="1">
. <element name="Stage"
  type="stage:Stage_Type"/>
. <element name="Fluidics"
  type="fluid:Fluidics_Type"/>
</choice>
. <element name="Viewing"
  type="instr:Viewing_Type"/>
. <element name="Sorts"
  type="boolean"/>
  <element name="Comments"
  type="dicom:Bd_1024_Type"
  minOccurs="0"/>
</sequence>
```

```
<attribute  
  ref="optics:Objective_List"/>  
</complexType>
```

Flow & Image Instruments

```
M.<complexType name="Microscope_Type">  
F.<complexType name="Flow_Cytometer_Type">  
  <complexContent>  
    <restriction base="instr:Instrument_Type">  
      <sequence>  
        <element name="Item_General_Info"  
          type="item:Item_General_Info_Type"/>  
M.<element name="Objective"  
  type="optics:Optic_Type" maxOccurs="7"/>  
F.<element name="Objective"  
  type="optics:Optic_Type" maxOccurs="1"/>  
M.<element name="Condenser"  
  type="optics:Optic_Type" minOccurs="0"  
  maxOccurs="1"/>  
F.<element name="Condenser"  
  type="optics:Optic_Type" minOccurs="0"  
  maxOccurs="10"/>
```

```
<element name="Platform"
  type="instr:Platform_Type"/>
M. <element name="Stage"
  type="stage:Stage_Type" minOccurs="1"
  maxOccurs="1"/>
F. <element name="Fluidics"
  type="fluid:Fluidics_Type"
  minOccurs="1" maxOccurs="1"/>
  <element name="Viewing"
    type="instr:Viewing_Type"/>
  <element name="Sorts" type="boolean"/>
  <element name="Comments"
    type="dicom:Bd_1024_Type"
    minOccurs="0"/>
```

```
</sequence>
  M. <attribute
    ref="optics:Objective_List"/>
  </restriction>
</complexContent>
</complexType>
```

Resource Description Framework (RDF)

Web Ontology Language (OWL)

Essential difference between RDF-OWL and Schema based XML

- RDF is used for showing connections and making inferences (artificial intelligence)
- XML based on schemas is used for standard information technology purposes.

Uses for XML metadata from cytometry measurements Schemas

1. To be stored in either a standard SQL relational database and/or an XML database or a combination thereof;
 - Schema tools can be used to create databases.
2. To be transmitted from one computer to another;
 - Schema and copy of schema validate at both ends & can be used for compression.

Uses of cytometry metadata (Schemas)

Cont.

3. To permit validation of data, particularly that entered by humans;
 - Data entry (dictation) errors reduced by strong typing. Ontologies (RDF-OWL) can provide the values for enumerations. XForms, works with schema.
4. To publish the data as reports or papers;
 - Microsoft Open XML Format
5. To define control files, such as test definitions;
 - These files require details.

Uses of cytometry metadata (RDF-OWL)

Cont.

6. To be used to find relationships between data.
 - If the determination of these relationships is to be accomplished outside of querying a database, then the functionality of RDF is definitely useful.

RDF-OWL example (from Flowcyt)

```
<rdfs:Class
  rdf:ID="Flow_Cytometer_Descriptor">
  <rdfs:comment
    rdf:datatype="http://www.w3.org/2001/
    XMLSchema#string">
    Contains all information that describes parts of a
    flow cytometer. </rdfs:comment>
  <rdfs:subClassOf>
    <owl:Class
      rdf:ID="Instrumentation_Descriptor"/>
    </rdfs:subClassOf>
  </rdfs:Class>
```

Health Level 7 (HL7)

- HL7 version 3 -- Reference information model -- Release 1 Standard
 - ISO/HL7 21731:2006 Health informatics
 - <http://www.hl7.org/v3ballot/html/welcome/downloads/downloads.htm>
 - Full Site
 - This download (225 Meg ZIP file) includes all materials you find in this Version 3 Ballot site.
- The material in CytometryML corresponds to HL7 entities.

Navigation thru HL7 Version 3 Standard

- Ver3_Ballot/html/welcome/environment/index.htm (Top left)
- Introduction
- Foundation
- Specification Infrastructure
- Implementation Technology Specifications
- Services
- Domains
- Background Documents
- Support Files
- Known Issues

Coverage (32 items)

Transport Specifications	Implementation Guide: Using SNOMED CT
Accounting and Billing	Claims & Reimbursement
Clinical Genomics	Diagnostic Imaging
Laboratory	Specimen
Orders and Observations	Regulated Product Submission

Specimen lite

- COCT_RM080200UV.htm
- http://www.hl7.org/v3ballot/html/domains/ct/editable/COCT_RM080200UV.htm

Present Problems with HL7 Schemas

- Were not designed to be read by humans!
 - Meaningless file names
 - Opaque namespaces
- Foundation Schemas not object oriented
 - Voc.xsd 19,366 lines of code
 - All simpleTypes;
 - Everything from RealmOfUse (Country) to UnitsOfMeasureCaseInsensitive
- Overuse of abbreviations
 - bl,BL,bn,ED,st,ST,cs,CD,CE,CV,CS,CO,CR,S
C,Il,,ts,TS,ADXP,...

Partial Example

```
<xs:complexType name="COCT_MT080200UV.S  
specimen">  
  <xs:sequence>  
    <xs:group ref="InfrastructureRootElements"/>  
    <xs:element name="id" type="II" minOccurs  
      ="1" maxOccurs="1"/>  
    <xs:element name="code" type="CE"  
      minOccurs="1" maxOccurs="1"/>  
    <xs:element name="specimenNatural" type=  
      "COCT_MT080200UV.Natural" minOccurs="1"  
      maxOccurs="1"/>
```

Conclusions

1. Only need one standard for Flow and Image Cytometry.
2. Functionality of RDF is useful.
3. Incompatibility of RDF with XML Schema is a big problem.
4. HL7 would be acceptable if their schema generator were fixed.
5. ISAC should collaborate with HL7 and DICOM

Reference

- Andrew Hinchley, “Understanding Version 3, A primer on the HL7 Version 3 Communications Standard”, ISBN 3-933819-19-9, <http://www.moench-medien.de>