

# Design of an XML Format based on DICOM for Analytical Cytology Data

Robert C. Leif, Suzanne B. Leif, and Stephanie H. Leif

XML\_Med, a Division of Newport Instruments, 5648 Toyon Road, San Diego, CA 92115; rleif@rleif.com  
www.newportinstruments.com

## INTRODUCTION

A major success of Analytical Cytology is that it now produces clinically useful information, which should be accessible to the pathologist and to the rest of the clinical laboratory information system. There are two ways to do this: translate the International Society of Analytical Cytology, ISAC, Flow Cytometry Standard, FCS<sup>1,2</sup>, data into the format of the clinical information system or extend the clinical information system to include the functionality of FCS.

Digital Imaging and Communications in Medicine, DICOM<sup>3,4</sup>, is a clinical information system standard, which already includes the functionality required for digital slide microscopy and can easily be extended to include flow cytometry. DICOM has been sponsored for

ISAC, April 28, 2002 6:29 PM. Page 1 of 31. Copyright Newport Instruments. All rights reserved.

data types from each of these standards can be contained in one or more XML schema documents, "schemas".

ISAC, April 28, 2002 6:29 PM. Page 3 of 31. Copyright Newport Instruments. All rights reserved.

microscopy by the American College of Pathologists. Since DICOM does include the functionality of FCS, it is possible to translate FCS into DICOM; unfortunately, this suggestion has produced a very strong **not invented here** response(5). Fortunately, a simple compromise exists. The DICOM data types can be expressed as XML schema and the actual data can be saved and shipped as XML data. This would also follow the present approach of Health Level 7, HL7<sup>6</sup>.

**Objective:** Map Analytical Cytology data to the Digital Imaging and Communications in Medicine, DICOM, standard data types and demonstrate the feasibility of expressing DICOM including its coding schemes in the Extensible Markup Language, XML.

**Method:** Previously(7,8,9), the existing DICOM Standard documents have been used to create a mapping of the existing Flow Cytometry Standard (FCS) data types to DICOM data types in particular those from Slide-Coordinates Microscopic Image Information Object. This mapping follows the present design of DICOM. As is shown in Figure 1, CytometryML is composed of data types from multiple standards. The

ISAC, April 28, 2002 6:29 PM. Page 2 of 31. Copyright Newport Instruments. All rights reserved.

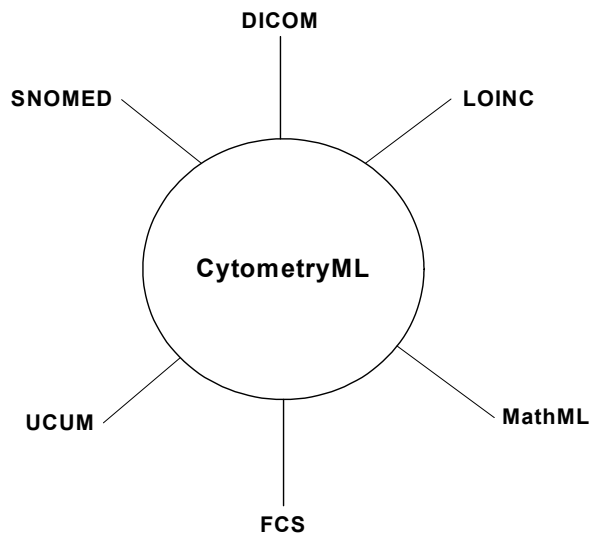


Figure 1. Schema organization. The various parts of the Cytometry schema import data types from DICOM including types from DICOM foreign Coding schemes, such as LOINC, SNOMED, and UCUM. Each of these data type sources and others such as MathML and FCS are or can be expressed as an XML schema.

- DICOM = Digital Imaging and Communications in Medicine.

ISAC, April 28, 2002 6:29 PM. Page 4 of 31. Copyright Newport Instruments. All rights reserved.

- LOINC = Logical Observation Identifier Names and Codes.
- SNOMED = Systematized Nomenclature of Medicine.
- UCUM = Unified Code for Units of Measure.
- MathML = Mathematical Markup Language.
- FCS = Flow Cytometry Standard.

## Mapping of FCS List Mode to DICOM Waveform

This utilization of data types from existing standards will be demonstrated by a mapping of the FCS list mode to the DICOM Waveform Information Object. The form of data presentation will be XML documents, which exist in the form of documents or “streams of bytes sent between applications, as fields in a database record, or as collections of XML information items, such as those found in the materials and methods of a Cytometry paper.

A flow waveform can be extended to include the closely related microscopic image list mode object. List mode for flow cytometry and slide microscopy data differ in that 1) for flow, time is relevant; and for microscopy, the

slide x and y coordinates are relevant; and 2) Individual cell images can be elements in an image list mode file.

The new DICOM Waveform Information Object<sup>10</sup> provides a means to describe List Mode data with functionality that is superior to FCS. The figure below describes how to map FCS list mode data to the elements of the DICOM Waveform Information Objects.

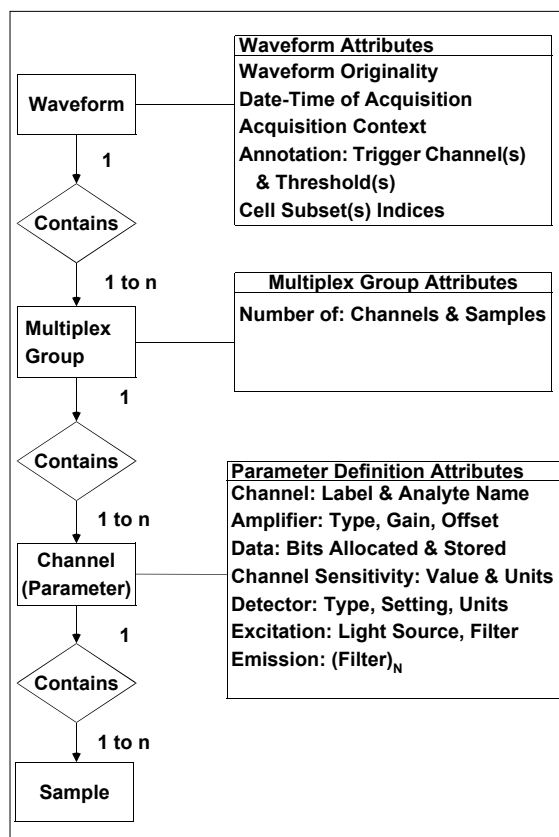


Figure 2. Description of List Mode based upon DICOM PS 3.3 - 2001, Figure J.5-1, DICOM Waveform Information Model (list mode).

Analytical cytology requires only 1 Multiplex Group (list mode file) for each data set. A single analytical cytology parameter is equivalent to a Waveform Channel. The Sample is the multidimensional vector or record that describes the data obtained from a cell.

In order to completely identify the data types and facilitate interoperability between standards, the DICOM Data Element and Value Representation (VR) Tags as well as the FCS Keywords will be provided as XML fixed attributes. Since these fixed attributes are constants, their presence will be limited to the schemas and they will not be included in the XML documents.

### Waveform Attributes

The mapping of the DICOM data types of the attributes to XML and their correspondence with FCS equivalents has been documented as Excel spreadsheets. These and other supplemental materials are available at [www.newportinstruments.com](http://www.newportinstruments.com).

The XML documents shown below are the result of the creation of XML schemas, which primarily consist of the mapping of DICOM to XML data types. The individual schema were then automatically translated into XML documents, which were subsequently edited to remove the constant attributes. The waveform XML document is shown below.

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <!--Sample XML file generated by XML Spy v4.3 U (http://
www.xmlspy.com) and then edited by RCL-->
3 <w:Waveform
  3a xmlns:w="file://C:\Stds\DICOM_XML\waveform.xsd"
  3b xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  3c xsi:schemaLocation="file://C:\Stds\DICOM_XML\waveform.xsd
C:\Stds\DICOM_XML\waveform.xsd">
4 <Modality>FLOW</Modality>
5 <Waveform_Originality>ORIGINAL</Waveform_Originality>
6 <Acquisition_Date_Time>2002-05-04T13:30:47-05:00
</Acquisition_Date_Time>
7 <Acquisition_Context>
8 <Acquisition_Context_Description>1 to 1024 Chars
</Acquisition_Context_Description>
9 <!--Chars=Characters-->
10 <Triggers Trigger_Min_Value="1" Trigger_Max_Value="255">
11 <Trigger_Source>1 to16 Chars</Trigger_Source>
12 <Trigger_Source_Long_Name>1 to 64 Chars
</Trigger_Source_Long_Name>
```

ISAC, April 28, 2002 6:29 PM. Page 9 of 31. Copyright Newport Instruments. All rights reserved.

space (xmlns) and location are given in S3a. Similarly S3b gives name space and location of the parent XML schema. S3c specifies the schema location for the above XML document. This schema, which is too long to be shown here, specifies the XML document's types, elements and attributes.

S4 specifies the Modality (FLOW or IMAGE). XML syntax specifies beginnings <element> and ends </element>. The Originality of the data (ORIGINAL or PROCESSED) is specified in S5. PROCESSED data is produced by some analytical process. Although FCS 3.0 employs an analysis section for this type of data, it does not provide an audit trail suitable for today's medical-legal environment.

The third element (S6) is the start Date-Time of acquisition of the flow cytometry list mode data or image data from which list mode can be derived. The Channel Label(s) for the Trigger parameter(s) and their threshold(s) are also attributes of the waveform.

The Acquisition\_Context, which includes multiple elements starts at S7 and ends at S23. It starts (S8) with an optional description of up to 1,024 characters.

ISAC, April 28, 2002 6:29 PM. Page 11 of 31. Copyright Newport Instruments. All rights reserved.

```
13 </Triggers>
14 <Index_File_Info>
15 <Index_File_Location>
http://www.newportinstruments.com/IF1
</Index_File_Location>
16 <Indexing_Parameters_Name>1 to 64 Chars
</Indexing_Parameters_Name>
17 </Index_File_Info>
18 <Index_File_Info>
19 <Index_File_Location>
http://www.newportinstruments.com/IF2
</Index_File_Location>
20 <Indexing_Parameters_Name>1 to 64 Chars
</Indexing_Parameters_Name>
21 </Index_File_Info>
22 <!--0 to 100 index files-->
23</Acquisition_Context>
24</w:Waveform>
```

Statement numbers have been added at the left. These will be abbreviated as Sn. S1 tells the browser that this is an XML file and that the UTF-8 character set is used. UTF-8 is an 8 bit format Unicode character representation. Latin-1 encoding would be specified as ISO-8895-1. S2 is a comment.

Waveform (S3), which is a global element of the schema, requires the schema prefix, w:, whose name

ISAC, April 28, 2002 6:29 PM. Page 10 of 31. Copyright Newport Instruments. All rights reserved.

S10 gives the range upon which to trigger the instrument for the parameter named in S11. Up to 100 parameters can be used in combination to trigger the instrument.

The addition of cell subsets to FCS<sup>11</sup> has been an important improvement and can be enhanced by employing the functionality provided by the DICOM Waveform to express subsets as index files. The locations of two of these index files are given in S15 and S19. The names are specified in S16 and S20. The indices are based on DICOM Referenced Sample Positions. These are positions in one or more Channels in the Multiplex Group, which corresponds to flow cytometry list mode parameters. These positions are numbered starting with 1 and are equivalent to the indices of an array. Specific ranges or SEGMENTS of the array can be addressed. This capacity to specify a collection of individual events permits the identification of these events as members of a subset.

The use of an index in DICOM as opposed to the addition of a parameter in the FCS list mode data both simplifies the software and increases its execution speed. Since the software can index through all of the

ISAC, April 28, 2002 6:29 PM. Page 12 of 31. Copyright Newport Instruments. All rights reserved.

data that applies to a specific cell subset, the subsets can be analyzed or rendered sequentially rather than simultaneously. These Referenced Sample Positions can also be applied to single channels and employed to gate the list mode data.

The elements of both the FCS Fluorescence Compensation Matrix and the DICOM Frame of Reference Transformation Matrix are listed in row-major order. However, there is no reason to be limited to DICOM for mathematical formulae, since this is the domain of the XML Mathematical Markup Language (MathML) Version 2.0. The description of the compensation matrix can be based on Section 3.5.1 Table or Matrix (mtable) [http://www.w3.org/TR/2001/REC-MathML2-20010221\(2001\)](http://www.w3.org/TR/2001/REC-MathML2-20010221(2001)), p. 91.

For Image data specific to the Acquisition Context, parts of the Visible Light Slide-coordinates Microscopic Image information object will be employed.

## Multiplex Group

A Multiplex Group is a collection of channels which are acquired synchronously. Although flow parameters can be acquired in sequence, all of the parameters for an

individual particle (cell) are grouped together in a single list mode file. Thus, they can be treated as one Multiplex Group. Similarly, in the case of Slide Microscopy, the data describing individual cells, which was derived from a large number of sequentially collected images, is stored as a single list mode file or one multiplex group.

A selection from the XML document that describes the data in a Multiplex Group is given below.

```
1 <Num_Waveform_Channels>10</Num_Waveform_Channels>
2 <Num_Samples>50000</Num_Samples>
```

S1 and S2 state respectively that data from 10 parameters and from 50,000 cells have been acquired.

A slightly abbreviated schema for this short XML document is given below.

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <schema targetNamespace=
   "file://C:\Stds\DICOM_XML\multiplexed_groups.xsd"
3a xmlns="http://www.w3.org/2001/XMLSchema"
2b xmlns:fcs="file://C:\Stds\DICOM_XML\fcs_3_0.xsd"
2c xmlns:multi=
   "file://C:\Stds\DICOM_XML\multiplexed_groups.xsd"
2d xmlns:dicom="file://C:\Stds\DICOM_XML\dicom.xsd"
```

```
3 elementFormDefault=
   "unqualified" attributeFormDefault="unqualified">
4a <import namespace="file://C:\Stds\DICOM_XML\dicom.xsd"
4b schemaLocation="file://C:\Stds\DICOM_XML\dicom.xsd"/>
4c <import namespace="file://C:\Stds\DICOM_XML\fcs_3_0.xsd"
4d schemaLocation="file://C:\Stds\DICOM_XML\fcs_3_0.xsd"/>
5 <annotation> <documentation>
Copyright 2002 Newport Instruments. The DICOM Number of Wave-
form Channels Type includes one element, Number, and two con-
stant attributes, Tag and VR. The use of an attributes with a fixed
value eliminates the need to enter a value into the Web page; yet,
permits the value to be accessed by the application. There is no
FCS keyword for this type. It is the 'n' in many FCS keywords.
</documentation> </annotation>
6 <complexType name="Num_Waveform_Channels_Type">
7 <simpleContent>
8 <restriction base="unsignedShort">
9a <minInclusive value="1"/>
9b <maxInclusive value="100"/>
10a <attribute name="Tag" type="dicom:Tag_Type"
    fixed="003A,0005"/>
10b <attribute name="VR" type="dicom:VR_Type" fixed="US"/>
11 </restriction>
12 </simpleContent>
13</complexType>
```

```
14<annotation> <documentation>
(50xx,2006) Number of Samples UL (unsigned long) appears 1
time. FCS Keyword is $TOT, which "specifies the total number of
events in the data set."
```

```
</documentation> </annotation>
15<complexType name="Num_Samples_Type">
16 <simpleContent>
17 <restriction base="unsignedInt">
18 <minInclusive value="1"/>
19a <attribute name="Tag" type="dicom:Tag_Type"
    fixed="50xx,2006"/>
19b <attribute name="VR" type="dicom:VR_Type" fixed="UL"/>
19c <attribute name="FCS_Keyword"
    type="fcs:FCS_Keyword_Type" fixed="$TOT"/>
20 </restriction>
21 </simpleContent>
22</complexType>
23<complexType name="Multiplex_Groups_Type">
24 <sequence>
25 <element name="Num_Waveform_Channels"
    type="multi:Num_Waveform_Channels_Type"/>
26 <element name="Num_Samples"
    type="multi:Num_Samples_Type"/>
27 </sequence>
28</complexType>
29<element name="Multiplex_Group"
    type="multi:Multiplex_Groups_Type"/>
30</schema>
```

The targetNamespace attribute for this schema (S2) is a file. S2a-d provide the namespaces and prefixes of the 3 other schemas and the multiplexed\_groups.xsd

schema itself (S2c). These all can supply elements, data types, and attributes. Since the preponderance of these come from the XMLSchema (S2a), it was chosen as the one schema without a prefix. S3 permits the elements and attributes to exist without prefixes. An import statement (S4a and S4c) is required to make the dicom and fcs\_3\_0 (FCS3.0) schemas visible. Their locations are given in S4b and S4d.

S5 is an annotation element that encloses documentation. Although this is essentially a typed comment, its use provides information to the XML parser.

S6 through S13 describe a type, which is called complex because it includes attributes. However, since the actual element, which is based on an unsignedShort (0 to 65,535) is explained in the syntax of a simpleType, the content is described a simple (S7). Since there has to be at least one measured parameter, the range starts at 1 (S9a) and ends with a number, 100, (S9b) that would permit the acquisition of a more than a complete fluorescence spectrum at a resolution of 10nm. Since this is a DICOM type, both its tag (10a), which is the concatenation to two pairs of hexadecimal

4 digit numbers, and its value representation, VR, which is a type or class; are given as fixed (constant) attributes. This provides both a 1 to 1 correspondence with DICOM and will facilitate interconversion. The same structure is used with the Num\_Samples\_Type (S15 to S22). However since this type is present in FCS3.0, an attribute which specifies the FCS keyword, \$TOT, has been included.

Elements of the two types are combined S23 to 28 into one global data structure Multiplex\_Groups\_Type. This combination is a sequence (S24 to S27). A sequence is the DICOM and XML format for an aggregate of objects or an element that contains one or more other elements. It is similar to a Pascal or Ada record, or a C struct. The use of a sequence in an XML document permits the number of times each element can be used to be specified and permits XMLSpy to automatically produce an XML document. In the element declaration (S29), the multi prefix is required to identify the Multiplex\_Groups\_Type.

XML Schema is a World Wide Web Consortium recommendation. As shown above, XML documents can be defined by schemas<sup>12</sup>, which employ XML

syntax. These schemas provide a very precise means for defining the structure, content and semantics of XML documents. Schemas permit groups or organizations to create classes of documents which include shared vocabularies, a common thesaurus. Schemas include the definitions of: Elements, Attributes, and Types.

Schemas are a software construct which includes:

- maximization of readability,
- control and renaming of namespaces,
- type extensibility
- and type restriction by
  - range checking, which is one type of assertion,
  - and pattern, which permits string formats that describe the order of characters as well as membership in character sets.

Since an XML page that employs types requires the public specification of the location of its schema(s), any additions by a software vendor have to be public. This should permit the investigator to have complete ownership and use of his/her data.

DICOM encompasses other standards by treating them as foreign Coding schemes. The simple approach shown in Figure 1 is to represent the data types of these other standards as schema. As was shown in the schema for the multiplexed\_groups, the multitude of coding schemes for the other standards can be mapped to fixed attributes. These fixed attributes are only included in the schema(s) for each standard and thus do not have to be referenced in schemas that import their data types or XML documents that are based upon them. The use of schemas also permits the inclusion of extremely useful XML Standards, such as MathML, which are thoroughly unrelated to DICOM.

### Parameters (Channels)

Each of the parameters (channels) is described by a sequence that includes the equivalent information used in FCS to describe a parameter. An XML document that describes a parameter is shown below.

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <p:Parameters
3a xmlns:p="file://C:\Stds\DICOM_XML\parameters.xsd"
```

```

3b xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3c xsi:schemaLocation="file://C:\Stds\DICOM_XML\parameters.xsd
C:\Stds\DICOM_XML\parameters.xsd">
4 <Parameter>
5 <Waveform_Channel_Number>1
  </Waveform_Channel_Number>
6 <Short_Name>FL1</Short_Name>
7 <Analyte_Info>
8 <Binding_Species>IgG</Binding_Species>
9 <Binding_Species_Name>Anti5BrdU
  </Binding_Species_Name>
10 <Tag_Name>Fluorescein</Tag_Name>
11 <Vendor>Phoenix Flow Systems</Vendor>
12 <Product_Number>PRB-1</Product_Number>
13 <Comment>Best there is.</Comment>
14 </Analyte_Info>
15 <Dectector_Info>
16 <Detector>PMT</Detector>
17 <Detector_Setting>600</Detector_Setting>
18 <Detector_Units>
19 <Prefix_Type>none</Prefix_Type>
20 <Unit_Abbreviation>V</Unit_Abbreviation>
21 </Detector_Units>
22 <Measurement>FLOURESCENCE</Measurement>
23 <Beam_Splitter_Info Prefix="nano" Unit="m">
24 <Manufacturer>Unknown</Manufacturer>
25 <Part_Num>D400</Part_Num>
26 <Beam_Splitter>DICHROIC_REFLECT_LOW
  </Beam_Splitter>
27 <Low_Cut_Off_1>500</Low_Cut_Off_1>

```

ISAC, April 28, 2002 6:29 PM. Page 21 of 31. Copyright Newport Instruments. All rights reserved.

```

28 <Description>Original equipment</Description>
29 </Beam_Splitter_Info>
30 <Emission_Filter_Info Prefix="nano" Unit="m">
31 <Manufacturer>Unknown</Manufacturer>
32 <Part_Num>B520</Part_Num>
33 <Emission_Filter>BAND_PASS</Emission_Filter>
34 <Band_Width_Location>1/e</Band_Width_Location>
35 <Peak_1>520</Peak_1>
36 <Band_Width_1>15</Band_Width_1>
37 <Description>Original equipment</Description>
38 </Emission_Filter_Info>
39 </Dectector_Info>
40 <Amplifier_Info>
41 <Mode>LOG</Mode>
42 <Gain>1.0</Gain>
43 </Amplifier_Info>
44 <Data_Format>
45 <Numeric_Class>Integer</Numeric_Class>
46 <Num_Bits_Allocated>16</Num_Bits_Allocated>
47 <Num_Bits_Stored>12</Num_Bits_Stored>
48 </Data_Format>
49 <Excitation_Info>
50 <Light_Source_Info>
51 <Manufacturer>Coherent</Manufacturer>
52 <Part_Num>Ar_CW</Part_Num>
53 <Light_Source>Ar_LASER</Light_Source>
54 <Power Prefix="milli" Unit="W">15</Power>
55 <Wavelength>488</Wavelength>
56 <Description>Original equipment</Description>
57 </Light_Source_Info>

```

ISAC, April 28, 2002 6:29 PM. Page 22 of 31. Copyright Newport Instruments. All rights reserved.

```

58 <Excitation_Filter_Info Prefix="nano" Unit="m">
59 <Manufacturer>Unknown</Manufacturer>
60 <Part_Num>488_Narrow</Part_Num>
61 <Excitation_Filter>BAND_PASS</Excitation_Filter>
62 <Band_Width_Location>1/e</Band_Width_Location>
63 <Peak_1>488</Peak_1>
64 <Band_Width_1>5</Band_Width_1>
65 <Description>Original equipment</Description>
66 </Excitation_Filter_Info>
67 </Excitation_Info>
68 </Parameter>
69 </p:Parameters>

```

Familiar values have been included in this XML document. The description of the Parameter (Channel) sequence starts with the Waveform Channel Number (S5) which is equal to the FCS parameter number, n. Since the sequence construct is used, the value of n needs to be given only once; rather than, as in FCS, being included with each data element.

The Analyte\_Info (S7 to S14) can be based on the LOINC database<sup>13</sup>. For instance, 522 items are listed in the LOINC CellMark Class. These start with CELLS.CD97/100 CELLS and end with BLASTS.TERMINAL DEOXYRIBONUCLEOTIDYL TRANSFERASE/100 BLASTS. The LOINC

ISAC, April 28, 2002 6:29 PM. Page 23 of 31. Copyright Newport Instruments. All rights reserved.

nomenclature presently describes test results; however, it can be used as a basis for the Binding\_Species\_Name (S9) in a format suitable for Analytical Cytology<sup>9</sup>. The complete description of the specimen is separate from that of the Waveform (List Mode) data.

The detector information (S15 to S39) includes the type of detector (S16) and its units. In this case a PMT run at 600 Volts. Both the standard prefixes (S19): milli, micro, etc. and the abbreviation (S20) come from the Unified Code for Units of Measure (UCUM). The use of "none" to describe a value that does not have a prefix is new.

The measurement (S22) is an enumerated type, which includes: fluorescence; low angle, 45 or 90 degree light scatter; extinction; dc or rf impedance; and other. Both the beam splitter (S23 to S29) and the emission filter (S30 to S38) can have up to 3 wavelengths.

Both the amplifier mode (S41), linear or log and gain (S42) are specified.

The data format (S44 to S48) describes the numeric class (S45), integer or float, the size of the data unit (S46), and the precision of the measurement (S47). A

ISAC, April 28, 2002 6:29 PM. Page 24 of 31. Copyright Newport Instruments. All rights reserved.

character type could be added. However, this does not appear to be necessary.

The excitation information (S49 to S67) includes information on both the light source (S50 to S57) and if there is one, an excitation filter. The formats are similar to the Detector\_Info (S15 to S21) and the Emission\_Filter\_Info (S30 to S38).

## Discussion:

Any standard must describe the objects and provide methods for them including their storage and transmission. The simplest way to have two standards interoperate is to base them on the same objects and data types.

As shown in Figure 3, XML schemas can be the basis of XML documents, which can interoperate efficiently and productively with both local and distributed computer systems. XML syntax acts as a common

technology that glues systems together.

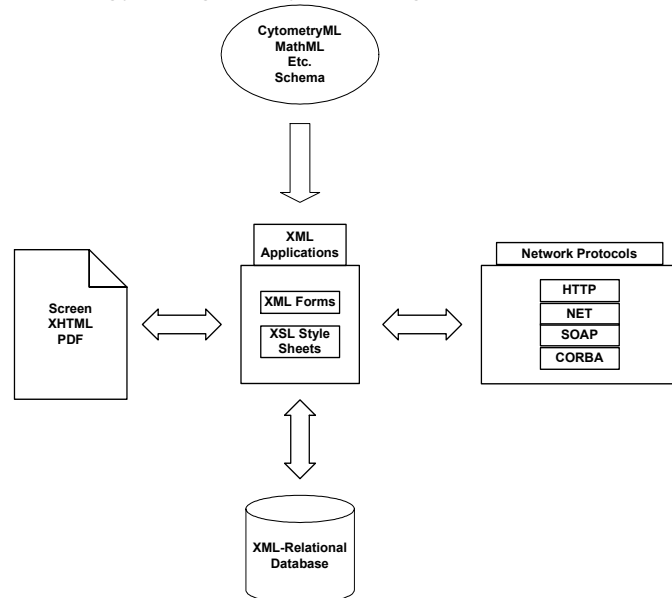


Figure 3. Interaction of XML schemas with other applications to produce screens, printed documents, store and retrieve data, and distribute information via the web.

Since commercial programs, such as Microsoft Excel, can operate in XML mode, it should be possible to have them operate directly on cytometry data.

The mapping of the large binary DICOM data types, such as visual light images and waveforms, to a textual format would result in an unacceptable increase in the data size. Therefore, the large DICOM binary data types should remain in DICOM format. However, these binary data types should be completely described in XML.

The capacity for Structured Reporting<sup>14,15</sup> in DICOM provides a very powerful means to connect the data and the pathologist or other individual who makes a clinical decision based on the data. XML could be used to generate these reports.

## Conclusions

The feasibility of expressing the data contained in a FCS file in XML employing DICOM data types has been demonstrated and a preliminary set of schema have been created. DICOM permits the creation of indices to individual cell subsets. Parts of DICOM coding schemes for other standards, such as UCUM have been represented by XML schema.

The combination of the domain knowledge of the

Cytometry community with the well thought out, tested design of DICOM and the universality of XML will have the benefits of:

1. Creating one analytical cytology standard for both flow and image data.
2. it is now possible to express all of the FCS keywords that are used to describe a flow cytometry experiment in DICOM.
3. DICOM by separating the list mode array into a separate file and incorporating the cell type data into an index provides a simpler and easier to maintain structure than the FCS monolithic structure.
4. retiring the present xenophobic FCS 3.0<sup>2</sup>;
5. basing a cytometry standard on one that is open, well designed, reliable, internationally accepted, and backed by the medical profession;
6. operating system independence and interfacing with multiple programming languages thus, eliminating the need for a reference implementation in one programming language;

7. interoperating with the existing medical informatics infrastructure;
8. increasing the amount of information, including clinically relevant material, in the data;
9. increasing the speed of analysis and display of cell subsets;
10. and the use of a World Wide Web standard, XML, as an intermediate form will greatly facilitate data exchange via the Internet. provide an open, reliable standard, and be interoperable with the rest of medical informatics.

## Acknowledgements:

We wish to thank Professor Ulysses J. Balis for the very helpful suggestion of mapping Flow Cytometry list mode to a DICOM Waveform.

## Future:

This document, the schema that were used to create the XML documents, and other ancillary information will be posted at [www.newportinstruments.com](http://www.newportinstruments.com). The authors welcome assistance in further refining and

ISAC, April 28, 2002 6:29 PM. Page 29 of 31. Copyright Newport Instruments. All rights reserved.

extending these documents. We hope to this in as open a manner as possible. Continuous peer review is a tested way to create a reliable software construct, such as an informatics standard. If ISAC will accept these documents for its web site, we will post them their.

## References

1. P. N. Dean, C. B. Bagwell, T. Lindmo, R. F. Murphy, and G. C. Salzman, (Data File Standards Committee), "Data File Standard for Flow Cytometry," *Cytometry*, **11**, pp. 323-332, 1990.
2. L. C. Seamer, C. B. Bagwell, L. Barden, D. Redelman, G. C. Salzman, J. C. Wood, R. F. Murphy, "Proposed new data file standard for flow cytometry, version FCS 3.0", *Cytometry*, **28**, pp. 118-22, 1997.
3. The official DICOM printed standard volumes are available for purchase from Global Engineering Documents at: <http://www.nema.org/standards/> and <http://global.ihs.com/>
4. <http://medical.nema.org/dicom/2001.html>
5. P. Dean (Facilitator), Data File Standards" ISAC XX Congress, Workshop, *Cytometry*, Supplement 10, p. 16, 2000.
6. P. Biron and K. Heitmann, "XML encoding rules of HL7 v2 messages - v2.xml.", *Health Level Seven (HL7)*, <http://www.hl7.org>
7. R. C. Leif and S. B. Leif, "Evolution of Flow Cytometry Standard, FCS 3.0, into a DICOM-Compatible Format". *Optical Diagnostics of Biological Fluids and Advanced Techniques in Analytical Cytology*, Ed. A. V. Priezzhev, T. Asakura, and R. C. Leif. A. Katzir Series Editor, Progress Biomedical Optics Series, SPIE Proceedings Series, **2982**, pp. 354-366, 1997.
8. R. C. Leif and S. B. Leif, "A DICOM Compatible Format for Analytical Cytology Data", in *Optical Investigations of Cells In Vitro and In Vivo*, D. L. Farkas, R. C. Leif, B. J. Tromberg, Editors, A. Katzir Biomedical Optics Series Ed. Proc. of SPIE **3260**, ISBN 0-8194-2699-7 pp. 282- 289, 1998.
9. R. C. Leif, and S. B. Leif, "A DICOM Compatible Format for Analytical Cytology Data, that can be Expressed in XML", *Optical Diagnostics of Living Cells IV*, D. L. Farkas and R. C. Leif, Editors, SPIE Proceedings Vol. 4260 pp. 238-48 (2001).
10. Digital Imaging and Communications in Medicine (DICOM) PS 3.3 - 2001, A.34 WAVEFORM INFORMATION OBJECT DEFINITIONS, pp. 109-115, 2001.

ISAC, April 28, 2002 6:29 PM. Page 30 of 31. Copyright Newport Instruments. All rights reserved.

11. D. Redelman, D.M. Coder, "Cell subset (CS) parameter to record the identities of individual cells in flow cytometric data." *Cytometry* Vol. 18, pp. 95-102, 1994.
12. XML Schema Part 0: Primer, W3C Recommendation, 2 May 2001 <http://www.w3.org/TR/xmlschema-0/>
13. LOINC<sup>®</sup> and RELMA<sup>™</sup>The Regenstrief Institute, [http://www.regenstrief.org/loinc/loinc\\_information.html](http://www.regenstrief.org/loinc/loinc_information.html).
14. DICOM Standards Committee, "Digital Imaging and Communications in Medicine (DICOM) Supplement 23: Structured Reporting Storage SOP Classes", NEMA (2000). [http://medical.nema.org/Dicom/final/sup23\\_ft.pdf](http://medical.nema.org/Dicom/final/sup23_ft.pdf)
15. D A.Clunie, "DICOM Structured Reporting", PixelMed Publishing ([www.pixelmed.com](http://www.pixelmed.com)), ISBN 0-9701369-0-0, 2000.