Cytometry Standards Continuum

Robert C. Leif[1], Josef Spidlen[2], Ryan R. Brinkman[2]
[1]XML_Med, a Division of Newport Instruments,
5648 Toyon Road, San Diego, CA, USA 92115-1022
E-mail: rleif@rleif.com; phone 1 619 582-0437
www.newportinstruments.com
[2]Terry Fox Laboratory, British Columbia Cancer Research Centre, Vancouver, Canada

ABSTRACT

Introduction: The International Society for Analytical Cytology, ISAC, is developing a new combined flow and image Analytical Cytometry Standard (ACS). This standard needs to serve both the research and clinical communities. The clinical medicine and clinical research communities have a need to exchange information with hospital and other clinical information systems.

Methods: 1) Prototype the standard by creating CytometryML and a RAW format for binary data. 2) Join the ISAC Data Standards Task Force. 3) Create essential project documentation. 4) Cooperate with other groups by assisting in the preparation of the DICOM Supplement 122: Specimen Module and Pathology Service-Object Pair Classes.

Results: CytometryML has been created and serves as a prototype and source of experience for the following: the Analytical Cytometry Standard (ACS) 1.0, the ACS container, Minimum Information about a Flow Cytometry Experiment (MIFlowCyt), and Requirements for a Data File Standard Format to Describe Flow Cytometry and Related Analytical Cytology Data. These requirements provide a means to judge the appropriateness of design elements and to develop tests for the final ACS. The requirements include providing the information required for understanding and reproducing a cytometry experiment or clinical measurement, and for a single standard for both flow and digital microscopic cytometry. Schemas proposed by other members of the ISAC Data Standards Task Force (e.g, Gating-ML) have been independently validated and have been integrated with CytometryML. The use of netCDF as an element of the ACS container has been proposed by others and a suggested method of its use is proposed.

KEYWORDS: Analytical Cytometry, Standard, FCS, CytometryML, netCDF, ISAC, DICOM, HL7

1. INTRODUCTION

The completion of a Requirements Document[1] has resulted in a consensus of a broad overview of what is needed in a cytometry standard. This has been augmented by the Minimum Information about a Flow Cytometry Experiment (MIFlowCyt) Standard[2]. The first requirement is to "Encode the required information to interpret a cytometry experiment." The description of this requirement states, "A cytometry standard shall provide a mechanism to record all the essential information for a person knowledgeable in the field to interpret a cytometry experiment." This interpretation includes the relationships between the information items. The creation of the descriptions of these the relationships could require some artificial intelligence capabilities.

The second requirement is to "A cytometry standard should provide a mechanism to record all the essential information to reproduce a cytometry experiment. The description of this requirement states, "A cytometry standard should provide a mechanism to record all the essential information to reproduce a cytometry experiment." A major reason that in the second requirement employed the weaker word, should, rather than the stronger word, shall, was the belief that it was not presently possible to compensate for the inherent differences in instruments that are not identical production models. Since reproducibility is the core of experimental science, it is essential to fulfill this requirement as much as is presently possible. The second requirement can and should be met by standard informatics technology. Instrument settings have already been expressed in XML documents, which were validated by the CytometryML[3,4] schemas (http://www.newportinstruments.com/cytometryml/cytometryml.htm). Previously, this metadata has been stored in a relational database[5]. The use of XML and XML schema also permits the use of XForms for validated data entry[6] and document creation, such as with Microsoft Word[7] and Open Document[8]. The first and second requirements have significant overlap, in that much of the information required to reproduce a cytometry experiment is also required to interpret a cytometry experiment.

The third requirement is, “Efficiently store cytometry list mode data.” The description of this requirement states, “A cytometry data file standard should provide an efficient way to store multi-parametrical list mode data and images. For example, binary information shall be encoded in a file format suitable for binary data.” Binary data should not be converted to XML (base64Binary) because this would result in a significant increase in file size and complicates numeric programing by requiring forward and backward conversions. Thus, the binary data (list mode and image) should be kept in a separate file from the XML metadata.

The fourth requirement is, “Facilitate support for other analytical cytology data.” The description of this requirement states, “Use a common (shared) mechanism/approach/methodology/terminology to store/analyze/transport data and metadata from flow cytometry as well as image and other analytical cytology technologies. (This does not imply that an initial implementation is required to support all types of cytometry data. An initial implementation could be directed to only one modality, such as flow cytometry).” As described previously with CytometryML, it was possible to derive both a flow cytometer and a digital microscope from a common, generic ancestor. List mode is commonly used to store data produced by a digital microscope and a commercial flow cytometer records images.

Once one has a set of requirements, the problem arises how to apply them in a manner consistent with real world constraints. One of these constraints is the reuse of existing standards. For example, ideally the Simple Format described below would be employed for Analytical Cytometry Standard (ACS); however, there is an existing netCDF[9] standard, which is used by a large part of the research community.

2. BINARY DATA IN ACS

2.1. Simple Format

Although the ease of use of separate binary files has been demonstrated[10], there is a real possibility that the XML and binary files could be separated. The utility of the information provided by the combination of both files is much greater than the utility of either the XML or binary file. Conversely, when the two types of files are combined their utility is minimal if they can not be separated. One standard approach to combining XML metadata with binary data is the use of Zip files[3] and, which according to Microsoft[11], “The XPS (XML Paper Specification) physical format understood in Windows Vista™ is a ZIP file. It has been proposed that the binary data files and XML metadata files be kept in a common container[3,12]. The simplest and most common way to accomplish this is to put them together in a zip file.

2.2. Proposed ACS netCDF File

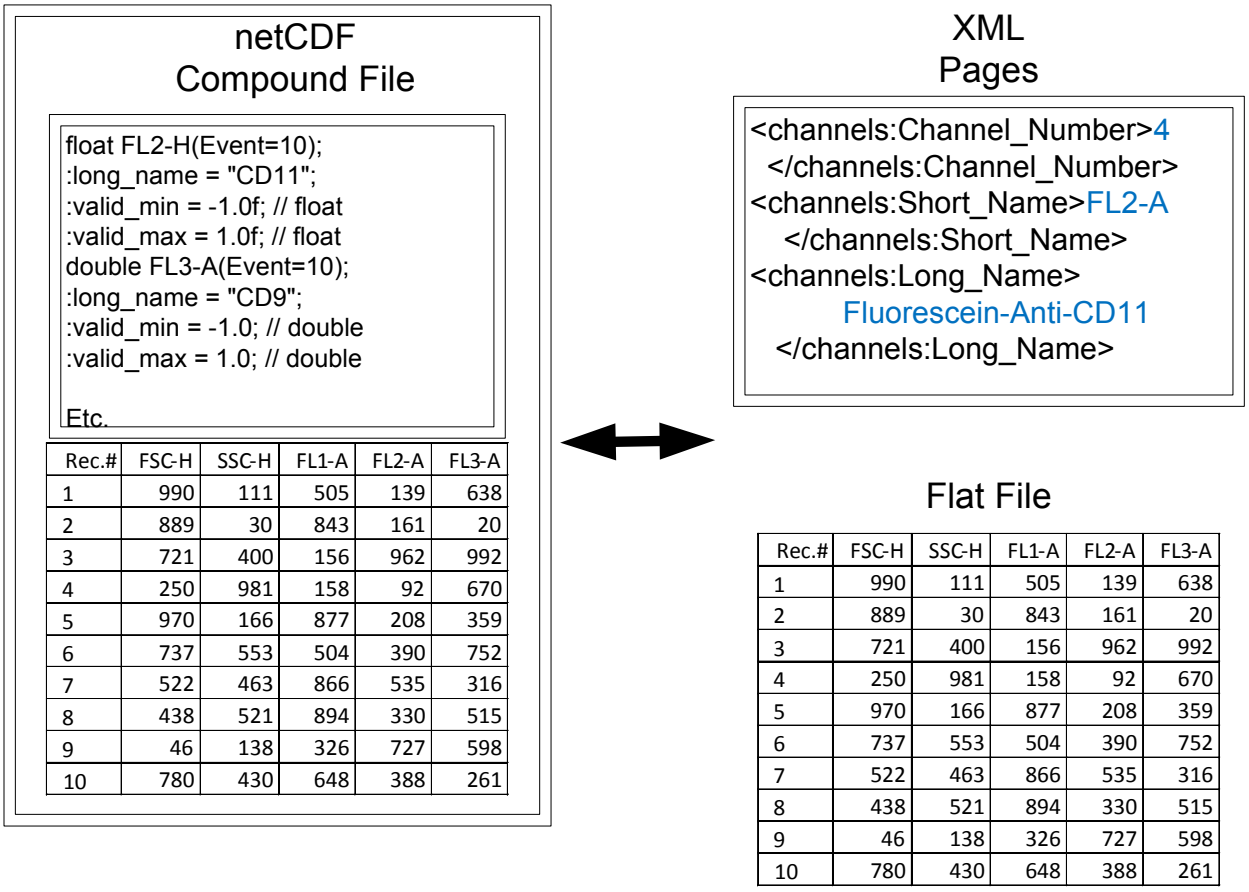
The netCDF format has been proposed by Spindlen et al[13]. According to Unidata Overview[14], “NetCDF (network Common Data Form) is an interface for array-oriented data access and a library that provides an implementation of the interface. The netCDF library also defines a machine-independent format for representing scientific data. Together, the interface, library, and format support the creation, access, and sharing of scientific data.” “Unidata is a diverse community of over 160 institutions vested in the common goal of sharing data, and tools to access and visualize that data”, which is funded primarily by the US National Science Foundation. The netCDF format includes the inclusion of a small part of the metadata in with the binary data. This file will be referred to as a compound-file. Neither the format of this metadata nor that of the included binary data is XML. The rest of the metadata is to be in the form of one or more XML files and the container at present will be probably be a zip file. It has been argued that the availability of the netCDF application programming interface (API) will facilitate the use of the netCDF format and therefore expedite its acceptance. The advantage of the existence of an interface to the R software environment for Statistical Computing[15] is also perceived as an advantage. Unfortunately, the proposed use of netCDF technically violates the following requirements that have already been specified. These include:

> Req. 4) Transparently store text-based data, which is described as, “Text-based information shall be encoded in a file format suitable for text data.” The metadata that has been included in the compound-file with the binary data violates this requirement because most of the file is composed of binary data and the text data is not in XML.
> Req. 8) “Each type of information shall only be stored in one file format.” This is described as, “There shall be no competing standard ways to store the same type of information, for example, if compensation description is stored in an XML-based file, it shall not be stored as part of an FCS data file.” Since metadata that is equivalent to that in the XML schemas is stored as part of the compound-file, Req. 8 is violated by the suggested use of netCDF. However,

this has largely been mitigated by the capacity to convert this data to XML that will then be validated against one or more schemas, as described below.

Req. 21) "Use a simple format oriented on data interoperability. This is described as, "The format shall be as simple as possible to support interoperability, unnecessary options shall be avoided." netCDF violates this by incorporation of the metadata together with the binary data. This removes the possibility of using tools that are supplied with the compilers or operating system APIs. The simplest format for the metadata is XML and the simplest data structure to represent list-mode data is a flat file, which is otherwise known as an indexed sequential file, and often referred to as ISM indexed Sequential Method. The other standard method to store this data is a B tree, which is often the basis of a relational database. However, this has to a large extent been mitigated by the capacity to extract the binary data into a flat file, as described below.

The problem of the trade-off, which exists between the absolute fulfillment of a set of requirements and real-world considerations based on the strong academic and research backing for the use of netCDF and its extensive use has been solved. The solution is 1) to use a version of netCDF that will only include a very limited number of metadata entities[13] and 2) create free conversion software that produces the pure binary and XML files and conversely there is similar conversion software that converts the pure binary and XML files into netCDF. This software will be made freely available and will not be covered by a license that in any way inhibits its commercial use. One of these programs has been already specified in the Analytical Cytometry Standard NetCDF Conventions for List Mode Binary Data File Component[13].



Rec.#	FSC-H	SSC-H	FL1-A	FL2-A	FL3-A
1	990	111	505	139	638
2	889	30	843	161	20
3	721	400	156	962	992
4	250	981	158	92	670
5	970	166	877	208	359
6	737	553	504	390	752
7	522	463	866	535	316
8	438	521	894	330	515
9	46	138	326	727	598
10	780	430	648	388	261

Figure 1. Left, Contents of the proposed netCDF file. A small text file is combined with the binary file. Right, Simple Format. The textual information from the netCDF file is in the form of one or more XML pages and the binary data is present as a separate flat file or possibly a B tree. the double headed arrow indicates the availability of interconversion software.

2.2.1"NetCDF to Simple Format Converter"

This converter (Figure 1) will be an open source software application that can convert a NetCDF file, such as the one shown on the left side of Figure 1 to two or more files. One of these will be the flat file shown on the right side of Figure 1. This file will be a simple binary sequential file, as previously described[10], which is the same format, which is "currently used for the binary data segment in FCS[16], i.e., a file with measurements stored as a list of events, one event sequentially after another; each event stored as a list of measurement values; the number and order of measurement values corresponding to the number and order of variables within the NetCDF data file[13]." As shown on the right side of Figure 1, one or more XML pages (files) that include all of the metadata that is stored within the NetCDF data file. This file or files will conform to the ISAC ACS XML and/or RDF schema conventions.

The converse of the converter described above will also be created. This application will combine and convert the two or more files shown on the right side of Figure 1 (the flat file with the XML and RDF pages) into the NetCDF file.

The availability of these conversion programs would also serve to decrease the possible anxiety of corporate intellectual property attorneys about entanglement with Free Software Foundation (GNU) licenses. Although the use of the netCDF specification is covered by a reasonable licence, and there are several programming language interfaces that are not covered by GNU licenses, the R interface[17] to and Java library[18] for netCDF are respectively covered by the GNU General Public License (GPL) and the GNU Library or Lesser General Public License, LGPL (http://www.gnu.org/copyleft/lgpl.html), which includes many of the conditions of the GNU General Public License, GPL (http://www.gnu.org/copyleft/gpl.html). ISAC should hold the copyright on the components of the standard that it creates and should avoid contaminating the copyright with any material that will in any way interfere with the rights to intellectual property that is owned by cytometry manufacturers.

ISAC should place the above specified conversion utilities on its web-site, where they can be freely downloaded. It is hoped that the above solution of providing conversion programs will provides a precedent for the solution of the problem of image files where entrenched formats such as RAW, TIFF, JPEG, and DICOM cannot be avoided.

3. USES

Before a software standard is created, its use cases should be defined. After its presumed uses have been described, the requirements document[1] can be revised and a hazard analysis[19] should be, in this case, updated and extended. The following uses for the ACS are suggested

1) Data transmission

2. Data storage

3. Data entry

4. Data retrieval

5. Data mining (establishing relationships)

6. Creation of Documents

    6.1. Reports

    6.2. Publications

    6.3. Grant proposals

    6.4. Web sites

In the case of cytometry and many other areas of science and medicine, all of the above uses can be handled by XML in combination with RDF except for large binary objects, such as images and list-mode data[10]. Since both humans and software share the common ability to make mistakes, the XML metadata should be validated. The use of XML schema to validate defined datatypes also facilitates and increases the accuracy of data entry and software creation. The more that the software environment knows about an object; the less the programmer needs to describe. XML schemas based on the XML Schema Definition Language, XSDL[20,21], have been employed in the cytometry markup language, CytometryML,

to define many of the datatypes used in cytometry[3,22]. Of greater significance, XSDL is used by Health Level 7, HL7, to define the datatypes for HL7 Version 3, which includes the infrastructure necessary to build a clinical messaging system (see below). The use of the Resource Description Framework (RDF) will greatly facilitate the analysis of data and data-mining of the XML produced for the other uses.

4. COOPERATION WITH OTHER ORGANIZATIONS

Since no single group or organization has sufficient knowledge or resources to produce a single standard that describes a cytometry experiment or clinical test in its entirety, each group or organization should produce works in their own area of expertise. However, all of the standards shall be specified employing a common modality; the datatypes employed in the standards should wherever possible be part of the same inheritance hierarchy; common datatype definitions should be employed; and the source of a datatype and its definition should be a major consideration in its selection. In the context of cytometry, this means that Health Level 7 (HL7) and Digital Imaging and Communications in Medicine (DICOM) standards provide the clinical information infrastructure; and ISAC provides the cytometry instrumentation, reagent, and analysis details; one standard is used for both flow cytometry and digital microscopy; and XML schemas as well as the Resource Definition Framework (RDF) are used as the primary modality for the standards.

The International Society for Analytical Cytology, ISAC, has two standards, an established one for flow analyses, FCS[16,23], and an infrequently used other for digital microscopy[24]. The Flow Informatics and Computational Cytometry Society, FICCS, (http://flowcyt.sourceforge.net/) has developed schemas that describe gating[26], transformation[27] and fluorescence compensation[28] and has proposed their use to ISAC. The Digital Imaging and Communications in Medicine (DICOM) Working Group 26 is developing Supplement 122[29], which includes pathology specimen identification and revised pathology storage classes. In the present preliminary version of Supplement 122, the specimen is a "discrete physical object that is the subject of pathology examination". It can be an organ part, block with embedded tissue, or a slide with something (sections, cells, etc.) on it. The physical characteristics of microscope slides and coverslips have previously been described in DICOM. The harvesting procedures, conservative conditions (fixation, freezing, etc.), organ and location within the organ, procedure step (sectioning, hisopathological or cytological examination, fixation, and staining are also to be included in Supplement 122. Tissue microarrays are presently also included in Supplement 122. Whole slide imaging is an important new technology, which is being introduced into DICOM with this supplement. Accession numbers and containers are also included. DICOM already includes key complex datatypes like patient and physician. Presently in DICOM, the number of colors (measurements) is being extended to be greater than 3. Neither list-mode nor flow cytometry are mentioned in DICOM.

The development of Supplement 122 is being performed in conjunction with the Anatomic Pathology Special Interest Group of HL7 (http://www.hl7.org/), which includes in its Charter the following statement:

> "Formal Relationships with Groups Outside of HL7"
> "This SIG acknowledges and cooperates with the DICOM WG 26 (Pathology Imaging), includes representation from the existing College of American Pathologists and will seek participations from specialty societies in the field of anatomic pathologist, surgeons, cancer researchers, and other users of surgical pathology reports. These liaisons will be with the approval of the HL7 Board in accordance with Policy and Procedures."

Since none of the directly or loosely related analytical cytology groups that are trying to create standards have the capacity or interest and probably do not have the required domain knowledge to create the datatypes that have or are being created by the DICOM-HL7 cooperative, it will greatly increase our chances of success if we limit our works to those areas where we have special expertise and reuse[30]and/or better yet interoperate with the products of the DICOM-HL7 cooperative.

4.1. HL7

The XML version of HL7 is version 3 -- Reference information model -- Release 1 Standard. It is a joint international standard with ISO, ISO/HL7 21731:2006Health informatics. This means that one needs to pay for the standard document or be a member of HL7. Since the standard is continuously being extended and modified, the necessary information is available at http://www.hl7.org/v3ballot/html/welcome/downloads/downloads.htm

The instructions recommend downloading the full site because that will maintain the hyperlinks. This download of very

large ZIP file includes all materials one will find in this Version 3 Ballot at the site.

Parenthetically, a good way to navigate the HL7 documentation is visit the Universal Domains, which will then lead you to the schemas. Since the schemas have unintelligible names, going to them first will be very frustrating. The path for Common Message Elements Types, CMETS, and then schemas is: Universal Domains/Common Message Elements Types/CMET Definitions for all domains. It should be noted that no guarantee can be given that this path will not be changed in a subsequent version of the ballot. For each datatype, there is a thumbnail image of an object diagram below the top center and an icon containing text at the bottom right. Activation of the thumbnail image reveals an object model; and activation of the icon containing text leads to the schema that describes the object model.

4.2. Suggested Organization for the Development of a Standard(S)

The only modality that will permit reuse of the HL7 Version 3 is XML schemas, XSDL. For those, who wish to eventually use the Resource Description Framework, RDFa[31,32] apparently supports XSDL simpleTypes that are embedded in XHTML documents and should eventually be able to use CURIES[33] to work with XSDL complexTypes.

The DICOM-HL7 infrastructure is essential to provide a continuous group of standards, since it includes diverse items, such as billing, specimen description, people, microscope slides, images, and the use of SOAP to send messages. CytometryML provides an infrastructure that describes the instrument and will describe the staining of the specimen. It is planned to have CytometryML or the new ISAC ACS include interface schemas that will permit other groups to reuse HL7 datatypes. This way the Data and Image Analysis Special Interest Group (D&IA SIG) of the Society for Biomolecular Sciences[34] could extend the DICOM-HL7 image model and the D&IA SIG could collaborate with the Flow Informatics and Computational Cytometry Society[35], which would continue to do the specialized flow informatics; however, it would do so in a manner that permits its XML software to interoperate with HL7 and the D&IA SIG.

CytometryML will be augmented by continuing to translate DICOM into XML schemas and working with HL7 to derive the CytometryML utility schemas, such as num_types and units from HL7.

5. CONCLUSIONS

A new combined standard should be based, as much as possible, on existing standards including netCDF, HL7, and DICOM. This will inevitably involve compromises. However, the reuse of any reasonable technology often outweighs the problems resulting from the necessary compromises associated with this reuse. The XSDL schemas of CytometryML (http://www.newportinstruments.com/cytometryml/cytometryml.htm), the XSDL schemas of the Flow Informatics and Computational Cytometry Society, at least some of which are becoming part of the ACS, and the image specific work of Data and Image Analysis Special Interest Group (D&IA SIG) of the Society for Biomolecular Sciences should all be harmonized as much as possible with each other and HL7.

6. ACKNOWLEDGEMENTS

This work was sponsored by Newport Instruments internal research funds and by NIH Grant EB5034 to RRB. RRB is an ISAC Scholar and a Michael Smith Foundation for Health Research Scholar.

7. REFERENCES

1. J. Spidlen, Ryan Brinkman, Robert C. Leif and other members of the ISAC Data Standards Task Force, “Requirements for a data file standard format to describe cytometry and related analytical cytology data Version 0.070920 September 20, 2007”, http://wiki.ficcs.org/ficcs/Requirements.

2. J. Lee, J. Spidlen, K. Boyce, J. Cai, N. Crosbie, M. Dalphin, J. Furlong, M. Gasparetto, E. M. Goralczyk, B. Hyun, K. Jansen, T. Kollmann, M. Kong, R. C. Leif, S. K. McWeeney, T. D. Moloshok, W. Moore, G. Nolan, J. Nolan, J. Nikolich-Zugich, D. Parrish, B. Purcell, Y. Qian, B. Selvaraj, C. Smith, O. Tchuvatkina, A. Wertheimer, P. Wilkinson, C. Wilson, J. Wood, R. Zigon, R. Scheuermann, and R. R. Brinkman, “Minimum Information about a Flow Cytometry Experiment (MIFlowCyt) Normative Standard, Version 07.5.25 – May 25, 2007, A standard for outlining the minimum information required to report the experimental details of flow cytometry experiments”, http://wiki.ficcs.org/ficcs/MIFlowCyt?action=AttachFile&do=get&target=MIFlowCyt_normative_standard_proposal_v070525.pdf (Last visited 5 Dec. 2007).

3. R. C. Leif, “CytometryML and Other Data Formats”, in Manipulation and Analysis of Biomolecules, Cells, and Tissues III, D. Farkas, D. V. Nicolau, and R. C. Leif, Editors, SPIE Proceeding Vol. **6088-0L,** 1-7 (2006).

4. R. C. Leif, S. B. Leif, and S. H. Leif, “CytometryML, An XML Format based on DICOM for Analytical Cytology Data “, *Cytometry*

54A, 56-65 (2003).

5. R. C. Leif, R. Rios, M. C. Becker, C. K. Becker, J. T. Self, and S. B. Leif, "The Creation of a Laboratory Instrument Quality Monitoring System with AdaSAGE". Advanced Techniques in Analytical Cytology, Optical Diagnosis of Living Cells and Biofluids, Ed. T. Askura, D. L. Farkas, R. C. Leif, A. V. Priezzhev, and B. J. Tromberg. A. Katzir Progress in Biomedical Optics Series Editor SPIE Proceedings Series, Vol. **2678**, 232-239 (1996).

6. "XForms 1.0 (Third Edition)", W3C Recommendation 29 October 2007, http://www.w3.org/TR/xforms/.

7. "Microsoft Office Word 2007 Technical Articles, Walkthrough: Word 2007 XML Format", http://msdn2.microsoft.com/en-us/ms771890.aspx.

8. "Open Document Format for Office Applications (OpenDocument) v1.1, OASIS Standard, 1 Feb 2007", http://docs.oasis-open.org/office/v1.1/OS/OpenDocument-v1.1.pdf.

9. "The NetCDF Users' Guide, version 3.6.2", http://www.unidata.ucar.edu/software/netcdf/docs/netcdf.html (last updated on 23 July 2007).

10. R. C. Leif, "CytometryML, Binary Data Standards", in Manipulation and Analysis of Biomolecules, Cells, and Tissues II, D. V. Nicolau, J. Enderlein, R. C. Leif, and D. Farkas, Editors, SPIE Proceeding Vol. **5699**,. 325-333 (2005).

11. "Integrate Data, XML Paper Specification (XPS)", http://msdn2.microsoft.com/en-us/windowsvista/aa904982.aspx (last visited 8 Dec. 2007).

12. "Analytical Cytometry Standard (ACS) 1.0 (Proposal Version 0.8)", http://wiki.ficcs.org/ficcs/ACS_1.0?action=AttachFile&do=get&target=ASC1_v0.8.pdf (Last visited 8 Dec. 2007).

13. J. Spidlen, R. C. Leif, W. Moore, R. Brinkman, "NetCDF Conventions for List Mode Binary Data File Component, Proposal version 071011", http://wiki.ficcs.org/ficcs/NetCDF?action=AttachFile&do=get&target=Conventions_071011.pdf (October 11, 2007)

14. "An Overview of Unidata", http://www.unidata.ucar.edu/publications/directorspage/UnidataOverview.html (last visited 16 Dec. 2007).

15. "The R Project for Statistical Computing", http://www.r-project.org/ (Last visited 4 Dec. 2007).

16. 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**, 118–122 1997.

17. P. Michna, "RNetCDF: R Interface to NetCDF Datasets", http://cran.r-project.org/src/contrib/Descriptions/RNetCDF.html (last visited 16 Dec. 2007).

18. "NetCDF Java Library, Version 2.2 (alpha)", http://www.unidata.ucar.edu/software/netcdf-java/ (last visited 10 Dec. 2007).

19. R. C. Leif and S. B. Leif, "Evolution of Flow Cytometry Standard, FCS3.0, into a DICOM-Compatible Format". in *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, Vol. **2982**, 354-366 (1997).

20. "XML Schema Part 2: Datatypes Second Edition", W3C Recommendation 28 October 2004, http://www.w3.org/TR/2004/REC-xmlschema-2-20041028/.

21. Pricilla Warmsley, *Definitive XML Schema*, Prentice Hall, http://www.phptr.com (2002).

22. R.C. Leif, S.H. Leif, S.B. Leif, "CytometryML, a Markup Language for ANALYTICAL CYTOLOGY", in *Manipulation and Analysis of Biomolecules, Cells and Tissues*, D. V. Nicolau, J. Enderlein, and R. C. Leif, Editors, SPIE Proceedings Vol. **4962**, 288-297 (2003).

23. "FCS, Flow Cytometry Standard", http://www.isac-net.org/ Then search for FCS.

24. P. Dean, L. Mascio, D. Ow, D. Sudar, J. Mullikin, "Proposed Standard for Image Cytometry Data Files", Cytometry, Vol. **11**, 561-569 (1990).

25. R. C. Leif, "CytometryML: a Data Standard Which Has Been Designed to Interface With Other Standards," in *Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues V*, D. L. Farkas, R. C. Leif, D. V. Nicolau, Editors, SPIE Proceedings Vol. **6441**, 64410P 1-11 (2007).

26. Data Standards Task Force, Bioinformatics Standards for Flow Cytometry Consortium: "Proposal for International Society for Analytical Cytology (ISAC), Gating-ML: Draft Standard for Gating in Flow Cytometry, version 1.1", http://flowcyt.sourceforge.net/gating/ (2006)

27. Data Standards Task Force, Bioinformatics Standards for Flow Cytometry Consortium: "Proposal for International Society for Analytical Cytology (ISAC), Transformation-ML: Draft Standard for Transformation Description in Flow Cytometry, version 1.0"; http://flowcyt.sourceforge.net/transformation/ (2006).

28. Data Standards Task Force, Bioinformatics Standards for Flow Cytometry Consortium:" Proposal for International Society for Analytical Cytology (ISAC), Compensation-ML: Draft Standard for Compensation Description in Flow Cytometry, version 1.0"; http://flowcyt.sourceforge.net/compensation/ (2006).

29. “Supplement 122, Specimen Identification and Revised Pathology”, DICOM, NEMA, http://www.dclunie.com/dicom-status/status.html#BaseStandard2001 (Last visited 16 Dec 07).

30. B. Boehm, K. Sullivan, “Software Economics: A Roadmap”, in *International Conference on Software Engineering, Proceedings of the Conference on The Future of Software Engineering*, ACM Press New York, NY, USA. (2000).

31. B. Adida and M. Birbeck, “RDFa Primer 1.0, Embedding RDF in XHTML,” W3C Working Draft 12 March 2007, http://www.w3.org/TR/xhtml-rdfa-primer/.

32. M. Birbeck, S. McCarron, “XHTML RDFa Modules, Modules to Support RDF Annotation of Elements, W3C Editor's Draft 2 April 2007, http://www.w3.org/MarkUp/2007/ED-xhtml-rdfa-20070402/

33. CURIE Syntax 1.0, A Syntax for Expressing Compact URIs, W3C Working Draft 26 November 2007, Editors: M Birbeck, S. McCarron, http://www.w3.org/TR/curie/ (last visited 16 Dec. 2007).

34. Data and Image Analysis Special Interest Group (D&IA SIG) of the Society for Biomolecular Sciences (http://www.ravkin.net/SBS/D&IA_SIG.htm) (last visited 16 Dec. 2007)

35. Bioinformatics Standards for Flow Cytometry, http://flowcyt.sourceforge.net/ (last visited 16 Dec. 2007)