

FLOW CYTOMETRY DATA STANDARDS FOR THE SEMANTIC WEB

Josef Spidlen¹, Robert Gentleman², Perry Haaland³, Thomas Moloshok⁴, Michael Ochs⁴, Clayton Smith¹, Adam Treister⁵, and Ryan Brinkman¹

¹ BC Cancer Research Center, Vancouver, BC, Canada

² Fred Hutchinson Cancer Research Center, Seattle, WA, USA

³ BD Technologies, Research Triangle Park, NC, USA

⁴ Fox Chase Cancer Center, Philadelphia, PA, USA

⁵ Tree Star, Inc., San Carlos, CA, USA

Objective

The Flow Cytometry (FCM) Data File Standard (FCS) has been adopted as the common representation of FCM data. Scientists can choose among instruments and software with no major compatibility issues for fluorescence values that FCS captures. However, experimental metadata and analysis components are not recorded in a standardized fashion for use by independent parties. This prevents collaborative opportunities to recreate experimental methods and results.

Materials & Method

The recent trend in data standard development has been to use Extensible Markup Language (XML) as the preferred mechanism to define data representations. We propose using XML where appropriate based on the character of stored information. We adopt other mechanisms (e.g., Resource Description Framework, Web Ontology Language, or Unified Modeling Language) where the syntactic and document-centric XML cannot achieve the level of interoperability required for highly dynamic metadata, such as flow cytometry experiment annotation.

Results

We have brought together a cross-disciplinary international collaborative group from both academia and industry to collaborate on development of data standards for FCM. We have developed a proposal on how to form gate definitions (Gating-ML), a proposal on how to describe parameter compensation (Compensation-ML), and transformation (Transformation-ML). Furthermore, we present guidelines outlining the minimum information required to unambiguously record, report, interpret, and reproduce FCM experiments (MIFACE). FlowRDF proposal suggests a methodology to annotate experiments in a standardized and extensible way, making the annotation both human-readable and machine-processable. Finally, we are proposing a new version of FCS addressing several well known shortcomings, through the incorporation of technologies from standards bodies such as the W3C Consortium and the OMG Group, that were not available when FCS was first conceived.

Conclusion

Wide collaboration has been recognized as highly important factor for the prevention of establishment of competing or incompatible standards describing the same domain. We are working with the Functional Genomics Investigation Ontology (FuGO) development group to extend the FuGO ontology for FCM purposes. Furthermore, we closely cooperate with the International Society for Analytical Cytology (ISAC) Data Standards Task Force to ensure quality and adoption of developed standards. Wide community feedback in this stage of development is a critical success factor. Our proposed standards are open and available for download at <http://flowcyt.sourceforge.net/>. We believe that these efforts will provide consistency in the electronic recording of FCM data and analyses leading to better interoperability, and fostering collaboration in many FCM-based research fields.

Contact Information

Josef Spidlen, Terry Fox Laboratory, British Columbia Cancer Research Center
675 West 10th Avenue, Vancouver, BC, V5Z 1L3, Canada
Phone: +1-604-675-8000 x 7755, Fax: +1-604-877-0712,
Email: jspidlen@bccrc.ca