

# Comment on the 2015 Interoperability Standards Advisory

Submitted by members of the Yosemite Project Steering Committee

<http://YosemiteProject.org>

*Yosemite Project Mission:*

*Semantic interoperability of all structured healthcare information*

## **Problem: Healthcare standards are silos**

The 2015 Advisory lists a patchwork of "best available" standards for clinical data interoperability. This is an important step. However, while each standard enables a level of interoperability within each domain, taken together the standards are silos: each standard is defined by different data formats, different data models and by different vocabularies that are *not always interoperable*. For example, many of these standards include overlapping concepts that are defined differently and in incompatible ways. Therefore implementers are required to create their own workarounds to bridge these standards and overcome the incompatibilities. However, because implementers use different approaches, data still lacks interoperability between implementers.

It is time to ***standardize the standards***. This will involve two steps: the first is technical and the second is social. The first step is to adopt a *universal information representation* for healthcare standards, and this is the step that this comment addresses. This universal information representation is crucial to enable the meaning of the data to be exposed in a uniform way across all required data domains and use cases. This will allow the meaning of any piece of any clinical data, expressed in any ONC-endorsed standard, to be accurately determined. The second step is to converge on a standard set of clinical concepts and related assertions that are shared across standards and that can support cross-cutting use cases so common in medicine. Although this is a social process of agreement, and not directly addressed by this comment, this social process will be greatly facilitated by use of a universal information representation.

## **Suggestion 1: Recommend a universal information representation**

The ONC should add a recommendation for a *universal information representation* that can accurately capture the meaning of any healthcare information, spanning all ONC-mandated

standards, regardless of the the data formats, data models, or vocabularies prescribed by those standards.

## **Suggestion 2: Recommend RDF as the best available universal information representation**

The ONC should recommend Resource Description Framework (RDF) version 1.1 as the best available candidate for a *universal information representation*, to meet the need described above. RDF is a standard produced by the World Wide Web Consortium (W3C) that has been used in a wide variety of domains -- including biomedical research -- for over 10 years. RDF -- and the RDF family, including RDFS, OWL, SKOS, etc. -- has already been used to express the meaning of medical vocabularies, such as SNOMED CT and ICD-11, but its use needs to be extended to all healthcare standards, to provide a common semantic foundation across all healthcare standards. Over 100 healthcare thought leaders and technologists have already expressed their view that RDF is the best available candidate for this purpose. (See <http://YosemiteManifesto.org/> )

### ***Signed by the following members of the Yosemite Project Steering Committee***

**David Booth, PhD**, is a senior software architect at Hawaii Resource Group and at Rancho BioSciences, using Semantic Web technology to make clinical healthcare data interoperability between diverse systems. He is also leading a joint HL7-W3C effort to defined an RDF ontology for FHIR. He previously worked at KnowMED, using Semantic Web technology for healthcare quality-of-care and clinical outcomes measurement, and at PanGenX, applying Semantic Web technology to genomics in support of personalized medicine. Before that he worked on Cleveland Clinic's SemanticDB project, which uses RDF and other semantic technologies to perform cardiovascular research. Prior to that was a software architect at HP Software. He was also a W3C Fellow from 2002 to 2005, where he worked on Web Services standards before becoming involved in Semantic Web technology. He holds a Ph.D. in Computer Science from UCLA.

**Conor Dowling** is CTO of Caregraf, which uses Semantic Web technologies to help health-care providers gather and analyze the information they create during the course of a patient's care. He is a specialist in clinical-data analytics with a focus on how the definition of clinical know-how and institutions shapes the description of patient care.

**Michel Dumontier, PhD**, is an Associate Professor of Medicine at Stanford University in the Stanford Center for Biomedical Informatics Research. His research focuses on methods to integrate large, heterogeneous clinical and biomedical data for discovery. His research interests include (1) developing novel therapeutics for rare and complex diseases, (2) elucidating the mechanism of drug-induced side-effects, and (3) optimizing multi-drug therapies to minimize undesirable side effects.

**Claude Nanjo, MA MPH**, is a Software Architect at Cognitive Medical Systems. He is also an active contributor to a number of HL7 and S&I clinical modeling initiatives including Health eDecision (HeD), the Clinical Quality Framework (CQF), and Fast Health Interoperability Resources (FHIR). At both Cognitive Medical Systems and Zynx Health, Claude has been involved in a number of research projects exploring the intersection between Clinical Decision Support and the Semantic Web. Prior to joining Zynx Health, Claude was engaged in research developing machine learning solutions to mine information on the Web.

**Josh Mandel, MD**, is a physician and software engineer at Children's Hospital Informatics Program at Harvard-MIT interested in improving clinical care through information technology. After earning an S.B. in computer science and electrical engineering from the Massachusetts Institute of Technology and an M.D. from the Tufts University School of Medicine, he joined the faculty of the Boston Children's Hospital Informatics Program and Harvard Medical School, where he serves as lead architect of the SMART Project (<http://smartplatforms.org>). Josh has a special interest in tools and interfaces that support software developers who are new to the health domain.

**Rafael Richards, MD MS**, is a Physician Informaticist in the Office of Informatics and Analytics at the Department of Veterans Affairs. He is currently the co-director of the VA-DoD interoperability research collaboration lab. Prior to this Dr. Richards was an Assistant Professor in the Division of Health Science Informatics at Johns Hopkins University. Rafael's current interests and activities are in data standards, medical device data integration, and modernizing VA databases to enable enterprise data federation with Linked Data sources. Prior to medicine, Rafael earned a B.S. Engineering from Swarthmore, M.S. Applied Mathematics from Lehigh University, and worked several years in industry in scientific software applications with IBM and in the UK office of Wolfram Research on Mathematica. Rafael is an invited expert in the Semantic Web in Healthcare and Life Sciences of W3C.