Supplemental VCDE Requirements Elicitation Initiative 2009 - 2010

Contents of this Page

- VCDE Requirements Elicitation Initiative
- Vocabulary Knowledge Center Semantics Requirements Forum (VKC)
- Background
- Metadata Management for Semantics Support
- Interested Parties and Stakeholders in New Semantic Infrastructure
- Primary Users
- Requirements in the Project Pages
- Completing Use Case and Interview Questionnaire templates

Quick Links

- Semantic Infrastructure Concept of Operations Initiatives - Requirements Master List
- Requirements Elicitation - Lessons Learned
- Requirements Elicitation Team Meeting Notes
- Semantic Requirements Use Cases
- User Stories for Semantic Requirements
- Requirements Questionnaires
- Requirements Gathering Process
- VCDE Requirements Elicitation Team Charter 2009
- VCDE Requirements Elicitation Final Report - Zoran Milosevic

Requirements planning for next generation semantics infrastructure requires ongoing community input and visionary leadership to move into the next era of semantic computing. Refer to the Semantic Infrastructure Roadmap for further evolution of the Semantic Infrastructure.

This page is linked to from the Semantic Infrastructure Concept of Operations Initiatives - Requirements Master List, which is a gateway into requirements gathering and related materials from Semantic Infrastructure Requirements Elicitation from 2009 and early 2010. The semantics computing capabilities to support interoperability in a Service Oriented Architecture demand both enhancements and new infrastructure to support the needs described below. This page contains links to the related documents and projects.

VCDE Requirements Elicitation Initiative

A small group from VCDE was formed to review and formally document requirements for the semantic infrastructure II. Refer to the VCDE Requirements Elicitation Team Charter 2009.

In eliciting the software requirements for the semantic infrastructure, we have created common templates for collecting requirements and related use cases, and are following a common process.

Vocabulary Knowledge Center Semantics Requirements Forum (VKC)

The community was asked to tell us what their requirements are for the next generation infrastructure and the VKC Semantics Requirement Forum was created for this purpose. Please feel free to visit the site and comment or contribute additional ideas.

ICR Requirements are on the Vocabulary Knowledge Center forum.

Background

The **Semantic Infrastructure and Operations Group** is responsible for the semantic aspects of the CORE Program Area, the caBIG® Vocabulary and Common Data Elements Workspace, and certain aspects of the caBIG® Architecture Workspace, and caGrid®.

The activities of the NCI CBIIT Semantics and Operations Group fall into three areas:

- Content Management - the processes and procedures that ensure the breadth and quality of the metadata and terminology used to record the semantics of data meet the needs of the caBIG® community.
- Semantics Infrastructure - design and development of software resources and operations including producing reference implementations of platform independent models

- Semantics Architecture and Management - defining the platform independent (as also called "implementation Independent") specification for systems and processes required to meet the semantics needs of the CBIIT/caBIG® enterprise, and for assuring that operational requirements for semantics support are met in a timely and reliable way.

Our **vision** is to provide computational and human interpretable representation of the meaning and context of data and services. Realization of this goal is a vital to enable the caBIG® community to revolutionize biomedical research, personalized medicine, and integrated care. To achieve this vision, the semantic infrastructure must:

- Continue to provide caBIG® with *computationally tractable representations of the meaning and representation of data*, and to extend semantic support to analytic and other services so that they can be discovered, understood, and securely utilized.

- Utilize a *consistent, comprehensive information management discipline* and software engineering standards such as ISO 10746 RM-ODP and its companion standard UML4ODP ISO 19793 to define both enterprise semantics needs and implementation neutral solutions to meet those needs.

- Provide *reference implementations* of enterprise-level platform independent models addressing semantic needs, especially the need for behavioral semantics.

- *Reduce the level of effort* associated with creation of semantic information, in part by leveraging to the greatest extent possible automated approaches to harvest semantics information from line of business and software engineering activities.

Metadata Management for Semantics Support

The semantic model and infrastructure forms a key component of the caBIG® collaborative infrastructure. The current semantic infrastructure uses a modified version of ISO 11179 Ed2, which formed the seeds for the development of ISO 11179 Ed3, and is the central component in allowing data elements and models to be annotated with concepts, and curated and registered in a repository allowing lookup and retrieval by both end-users and applications. It enables the automatic integration and transformations of data for sharing and collaboration by provisioning the infrastructure with clear, computable, and unambiguous data descriptors *for those who would create software* that can use and interpret the data in the service of cancer research.

The goal of the next generation infrastructure is to make these capabilities available to everyone with coarser-grained services that require little or no knowledge of the complexities of the infrastructure in order to gain its benefits.

Two additional opportunities for improvement in the current infrastructure approach are:

1) simplify the creation of this metadata which is currently labor intensive and therefore does not scale well

2) expand the approach beyond data discovery and interchange, to include services interoperability

Interested Parties and Stakeholders in New Semantic Infrastructure

The key to success for the next generation infrastructure is defining the critical, unmet semantic interoperability usage scenarios to ensure that the right next generation of services and infrastructure are provisioned. With the help of Mayo Clinic through the Vocabulary Knowledge Center this wiki to support requirements gathering was organized to collect and report on requirements.

Interested parties are encouraged to become involved over the next few months as we attempt to characterize in more detail the requirements to achieve the interoperability vision.

Primary Users

The primary direct and indirect end users include:

- Software and Application designers and architects
- Software and Application engineers and developers
- Scientific and medical researchers
- Medical research protocol designers
- Clinical and scientific research data managers
- Clinicians
- Patients
- Medical research study participants

Refer to Semantic Infrastructure Concept of Operations Stakeholders for a more complete description of stakeholders.

Requirements in the Project Pages

Requirements materials attached to this wiki page have been received from the following groups and are being used to supplement the Semantic Infrastructure Concept of Operations Initiatives - Requirements Master List:

Stakeholder	Contact	Analyst	Area of Interest	Usage/Primary Interaction Scenarios
Clinical Governance Group	John Speakman CTMS Wiki - Storyboards and Semantic Profiles for services interoperabilty	Patrick McConnell (21090)	ScenPro Analyst and 5AM, ISO Datatype Documentation of ISO use for COPPA (all projects have to use the datatypes) Implementation of guidelines must be complete in June	-'Operationalized' ISO 21090 Datatypes - Discover and share/reuse models - Rules engine and repository: Scenario #8: Management of Routine Non-Laboratory-Based Adverse Events (caAERS and CDMS) - Protocol Metadata and Rules Engine

				- C3PR FR-230 The system must integrate with the Cancer Data Standards Repository (caDSR) and Enterprise Vocabulary Services (EVS) (e.g. for eligibility criteria)
MediData	(Dianne Reeves)	---	Data Elements on Forms: Metadata registry information and semantic metadata, specifically data elements (CDEs) to record and share centrally defined forms variables that can be used to customize new protocols for clinical trials	- Share Forms including form structure, behavior and variables - similar to the way C3D uses caDSR but may have additional metadata that needs to be stored/shared
Genzyme	Sue Dubman Julie Smiley Director, Data Management Genzyme Oncology)	---	MDR metadata exchange -->Implement global data standards to collect, process, analyze and report clinical research data throughout the entire product lifecycle (see GetSmart attachment)	- share data standards and variables (CDEs) - Get Smart Goal #2 Data Collection and Processing: Technology enabled standard Case Report Forms and edit checks. - Get Smart Goal #5 Core Infrastructure: Data Elements Dictionary, Validation Tools, Systems Interoperability Guidelines and Architecture, Metadata Repository, Mapping and Conversion Tools and Structured Authoring Tool
Medical and Scientific Researchers	Yolanda Gill	---	Workflow: Metadata and rule support	- Metadata to support workflow
MD Anderson	Mike Riben	---	Alignment/Interoperability between NCI and MD Anderson's metadata and vocabulary solutions; Possible UAT(See MD Anderson Attachment)	- Metadata Registry Interoperability
Emory	Stuart Turner, Eliot Seigel and Joel Saltz	---	Semantic interoperation requirement stemming from the TCGA Radiology and In Silico projects. See Attached use case with 4 semantic requirements identified.	---
Novartis	Mehta Saurin	---	MDR tools ie. 1) Alternate names for permissible values i.e. currently you can register 'M', 'MALE' but we would like to register additional name such as 'm', '1', 'Male' etc. 2) Additional attributes for a data element - although there is a reference field available we would want (for operational purposes) additional attributes to define items such as 'SAS format', display format etc. 3) if the list of permissible values is Extensible flag 4) A way to relate data elements to each other 5) Distributed repository	- Enhanced/standardizxed Metadata attributes - Metadata Registry Interoperability
CDISC and SHARE	Margaret Haber	---	Pilot of Semantic Media Wiki for harmonizing and updating data elements; Input on tooling and metadata extensions	---
Mayo	Robert Friemuth	---	new metadata repository and CTS2	- Federated Terminologies
LS Governance and ICR	Juli Klemm, Baris	---	caB2B IRWG requirements, caBIG Gene Pattern and Analytical Services interoperability. Seamless interoperability, discover services and data that can be combined; construct new workflows	- Integrated system of tools (see IRWG Requirements) - Support workflow authoring tools (such as Taverna and caB2B)
Terminologists	Margaret Haber, Sheri De Coronado	---	formal requirements for the new terminology and metadata services and for assessing equivalence between pre- and post-coordinated terminology	See Semantic Infrastructure Concept of Operations Initiative 5 Compliance.]
caBIG Community	Various	---	Vocabulary Knowledge Center Semantic Requirements Wiki Forum	See Semantic Infrastructure Concept of Operations Initiatives - Requirements Master List of requirements organized by Semantic Infrastructure initiatives
Software Architects and Designers	Anand Basu Charlie Mead	Patrick McConnell	Discover and integrate services, on the fly, to perform scientific research	1. build new services that can interact with other existing services using workflow authoring tools such as Taverna 2. Support for "Conformance Profiles" --> Profiles are a mechanism used to constrain broader service capabilities to meet specific functional needs identified within a domain or locality (See Conformance Profiles attachment)
Metadata Curators	Dianne Reeves	---	Creating new content in caDSR	1. Customizable metadata download 2. Clinician friendly browsing: improve search and browsing functions, leveraging existing semantics and metadata 3. Improve ability to organize and reuse content (Classifications) with batch upload/editing
FDA	Margaret Haber Ed Helton	---	Metadata and Terminologies	TBD
caEHR	Gilberto Fragoso	---	---	---
NHIN	?	---	Interoperability in standards development	---
HL7/CIC	Dianne Reeves	---	HL7 ISO 21090 Datatypes; CIC MDR requirements	---

Completing Use Case and Interview Questionnaire templates

Template documents have been created and made available to those helping gather requirements to ensure consistency. To utilize one of the templates:
1. On this project page, click "Add Page"
2. Underneath the field for the page name, click "Select a page template to start from"
3. Select from:

- "2009-10 Requirements Refinement Script" for decomposition of Forum entries
- "2009-10 Requirements Use Case Template"
    4. Complete the template
    5. Click "Save"
    6. If there are attachments related to the item that have been posted to the Supplemental VCDE Requirements Elicitation Initiative 2009 - 2010 page, move or link them to the new page.