ISO Datatype Extended Support

Scope Document

Version No: 0.3

Last Modified: 3/10/2010

Team : caCORE SDK, caGrid, caAdapter

Client : National Cancer Institute -   
 Center for Bioinformatics and Information Technology,

National Institutes of Health,

US Department of Health and Human Services

Document History

Revision History

| **Version Number** | **Revision Date** | **Author** | **Summary of Changes** |
| --- | --- | --- | --- |
| 0.1 | 2/19/2010 | Satish Patel, John Eisenschmidt, Ye Wu | Initial Draft |
| 0.2 | 3/3/10 | Scott Oster | Copy edits, comments, and refinements |
| 0.3 | 3/10/10 | Satish Patel | Reconciled comments, edit content |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Review

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Name** | **Team/Role** | **Version** | **Date Reviewed** | **Reviewer Comments** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Related Documents

More information can be found in the following related documents:

|  |
| --- |
| **Document Name** |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

Table of Contents

1. Introduction 4

2. Problem Statement 4

2.1 Current Solution 4

2.2 Objectives 4

3. Stakeholder and User Descriptions 6

3.1 Stakeholder Summary 6

4. Timeline Requirement 7

5. Assumptions 7

6. Development Process 7

7. Tools and Features 7

7.1 Semantic Infrastructure Tooling 7

7.2 ISO21090 Java Library 8

7.3 caGrid Support 8

7.4 caAdapter Support: 9

7.5 caCORE SDK Support: 9

8. Release Non-Functional Requirements 10

Scope for ISO Datatype Support

# Introduction

The purpose of this document is to identify

1. Responsible teams to support the ISO datatype in the core infrastructure tooling
2. Enhancements needed in current infrastructure tools

# Problem Statement

NCI Center for Biomedical Informatics and Information Technology (CBIIT) and the entire health care community envision achieving working semantic interoperability through Services-Aware Enterprise Architecture Framework (SAEAF). The working semantic interoperability includes collections of structures, processes and components that support computable Semantic Interoperability. The working semantic interoperability will be achieved through an approved ECCF Conformance/Compliance framework that facilitates explicit and layered expression. ISO 21090 data type is one of the center pieces to the semantic interoperability which will define how data/information will be exchanged.

## Current Solution

The infrastructure tooling teams are working on providing a short-term solution of ISO datatypes implementation. The current short-term solution involves providing support for the limited set of datatypes by leveraging localized ISO 21090 library. This library includes 1) Java API, 2) XML serialization framework and 3) Hibernate mapping for thirty ISO 21090 based datatypes with limited set of attributes.

During the first phase of implementation, limited set of functionalities were added in the infrastructure tools to meet the short-term requirements of PODS project. In order to provide support for additional features, current tooling support needs to be enhanced.

## Objectives

Primary objective of this project is to establish ISO 21090 compatible infrastructure along with comprehensive ISO 21090 enabling and supporting tools. The developed infrastructure should be ready to be adopted by any organizations to establish their own health information system, or to be integrated with nationwide BIGHealth system.

**Semantic Infrastructure Support:**

The core infrastructure includes the center’s new semantic infrastructure, tools to manage, integrate and process semantic metadata, workflow and etc.

**ISO 21090 Tooling Support:**

With support of enhanced caCORE SDK, caGrid, and caAdapter, the user will be able to explore the full spectrum benefit of semantic interoperability. Users and applications will be able to efficiently and effectively build semantic interoperable health systems; register and discover health semantic information; exchange health related information; etc.

# Stakeholder and User Descriptions

## Stakeholder Summary

|  |  |  |
| --- | --- | --- |
| **Name** | **Description** | **Responsibilities** |
| George Komatsoulis | Deputy Director, NCI |  |
| Charlie Mead | Chief Technology Officer, NCI CBIIT |  |
| Avinash Shanbhag | Director, Core Infrastructure Engineering | Oversees CBIIT caCORE software engineering, and caGrid development |
| Sichen Liu | Associate Director, Core Infrastructure Engineering | Oversees caAdapter, and SDK product management and development |
| Libby Prince | ISO 20190 Project Manager |  |
| PODS Project Team | The PODS implementation is the reference use case for this effort |  |
| COPPA Project Team | Contributor of the ISO datatype’s localized library | Provide support to the development teams to adopt the localized library, provide guidance on the library |
| caGrid Project Team | Development team | Provide support for the ISO datatypes in their managed tools |
| caAdapter Project Team | Development team | Provide support for the ISO datatypes in their managed tools |
| caCORE SDK Project Team | Development team | Provide support for the ISO datatypes in their managed tools |

# Timeline Requirement

* TBD based on scope and requirements

# Assumptions

* TBD

# Development Process

In order to provide support for the ISO 21090 datatypes in the infrastructure tools, multiple project teams have to implement enhancements in their respective tools. The individual project teams will formulate a virtual task force to implement the project requirements. The task force will implement the project by going through following phases:

**Analysis Phase:**

As the long term solution for ISO 21090 datatypes is developed, the project teams will perform a detail analysis of the project components and features to determine enhancements needed in the infrastructure tools. The development teams have identified some of the analysis and enhancements areas; these analysis and enhancement areas are listed in the next subsection. *After the analysis phase, the project scope will be adjusted based on the findings and priorities.*

**Implementation Phase:**

After the analysis phase is complete, the development team will start the implementation plan. The implementation phase will require close coordination between multiple project teams. The developed products will be tested in individual manner an independent QA team. The individual product testing will be followed by the integration testing to ensure the infrastructure tools work in harmony with each other.

**Transition Phase:**

In this phase, the development team will also work with the technical writer team to develop necessary user documentation for their respective tools. During the transition phase, the developed products will be deployed in the target deployment environment.

# Tools and Features

## Semantic Infrastructure Tooling

Semantic infrastructure is a necessary component of the data service infrastructure. Current semantic infrastructure supports the model based registry with model elements being simple types. If the model uses complex ISO datatypes, then the registry is not capable of handling the details of the complex types. As part of this effort, the semantic infrastructure will be evaluated for fitness to support the requirements of this project. After the evaluation, the project teams will determine what enhancements are necessary and how those enhancements will be implemented. At minimum following items will be explored:

1. Registration of complex Datatypes in the semantic infrastructure as a project
2. Metadata registration tooling to support “shorthand’ notation for use of those types (e.g. They can be listed as a datatype in the model, but expanded to be a complex class association in the registered metadata model)

## ISO21090 Java Library

ISO 21090 common library was developed as an independent software project by a cross-project team, in support of the following:

1. Java Beans for CBIIT-wide Localization
2. Serialization support for to/from those Beans to standard ISO XSD
3. Hibernate support for ORM mapping of those Beans to database

In the first phase of implementation, the common library was the only implementation supported from the tooling teams. For longer term, the development team will evaluate two possible options to support the ISO datatypes:

1. Enhance the existing library to support additional datatypes and their attributes from ISO datatype standard not supported in the first release.
2. Provide a framework for creating localized ISO 21090 library and using them in the infrastructure tools.

Final recommendation on the implementation will be made after the analysis of these options. Development team will provide the tooling support for the selected option.

## caGrid Support

caGrid’s Introduce project provided initial support for the ISO 21090 datatypes in form of an extension. The initial extension included support for the 30 datatypes from common ISO 21090 library that COPPA/PODS project was using. The development team will analyze the problem space to determine workflow and necessary tooling enhancements in order to provide larger support to the ISO datatype standard. At minimum, following items will be analyzed for enhancements:

1. Enhance support for additional ISO 21090 datatypes and like common library, provide tooling support for different localizations of ISO 21090 datatypes
2. Enhance caGrid’s Metadata Model Service (MMS) to read the metadata from the target semantic infrastructure, allow for annotations on models using ISO datatypes, and generate domain models (as metadata for Data Services infrastructure) for models using ISO datatypes.
3. Enhance Data Service infrastructure to support new MMS
4. Enhance query support to query through the ISO datatypes based systems
5. Enhance Introduce to support ease of use of ISO datatypes (as provided in the short-term support, perhaps enhanced)
6. Enhance Data Service Infrastructure to leverage SDK query support for ISO datatypes

## caAdapter Support:

caAdapter project provided initial support for the ISO 21090 datatypes in the Model Mapping Service. The initial release included support for the 30 datatypes that COPPA/PODS project was using. The caAdapter team will evaluate and perform following activities in order to provide larger support to the ISO datatype standard.

1. Support ISO 21090 conformation and localization process. caAdapter will provide capabilities for users to load ISO 21090 datatypes, eliminate unused datatypes and attributes, define default values, global constant, add constraints, and etc.
2. Support evolving ISO datatypes and localization. When ISO 21090 datatype standard or the localization version changes, caAdapter will identify the change and visually present the changes. caAdapter will automatically adjust the localization. caAdapter will integrate with SDK to automatically update all subsequent mapping and transformation artifacts.
3. Extend caAdapter to support all caCORE SDK supported ISO 21090 datatypes, constraints and implemented in the MMS model.
4. For an application that keeps ISO datatype models, information models and data models, provide tooling support to perform mapping between ISO datatype model, information model and data model.
5. Extend caAdapter to support semantic integration with new semantic framework. The new semantic framework may introduce artifacts, such as artifacts to support behavioral framework, ECCF framework. caAdapter will provide capabilities to assist mapping among these new artifacts.

## caCORE SDK Support:

caCORE SDK project provided initial support for the ISO 21090 datatypes in the Java API and Code generator. The initial release included support for the 30 datatypes that COPPA/PODS project was using. The SDK team will evaluate and perform following activities in order to provide larger support to the ISO datatype standard.

1. Enhance support for additional ISO 21090 datatypes and provide tooling support for different localizations of ISO 21090 datatypes
2. Enhance SDK code generator components including Java beans, Hibernate mapping files, XML serialization files, Model validators etc for support of the ISO 21090 datatype library
3. Enhance SDK Java API to support additional datatypes and additional attributes
4. Enhance QBE translator support for additional datatypes
5. Perform integration with the CQL query translator to support additional datatypes
6. Redesign RESTful interface to support query through the ISO datatypes
7. Redesign the SDK web interface to support query building for ISO datatypes
8. Enhance security in the runtime system
9. Enhance Writable API functions to support persistence through the ISO datatypes

# Release Non-Functional Requirements

1. **Open Source**

The product will continue to use open source tools and technologies. If an appropriate open source software or tool is not available then prior approval from the product manager will be required before using the tool.

1. **Technology Stack compatibility**

The products will continue to adhere to the CBIIT technology stack when choosing the versions of the tools it’s using. Any deviation from the technology stack will be pre-approved by the product manager.

1. **Automated builds**

The products will continue to support their existing automated build processes.