**caBIG® Platform Independent Model and Service Specification**

**CS13: Specimen Identifier Management Service**

**2/21/2010**

|  |  |
| --- | --- |
| **Authors** | Carlos Perez |
| **Reviewers** |  |
| **Architecture Inception Team** | SemanticBits, LLC |

Document Change History

|  |  |  |  |
| --- | --- | --- | --- |
| **Version Number** | **Date** | **Contributor** | **Description** |
| 0.0.1 | 9/22/2010 | Carlos Perez | Initial Document |
| 0.0.2 | 9/22/2010 | Carlos Perez | Update |
| 0.0.3 | 11/10/2010 | Carlos Perez | Update to reflect CIM 0.1.0 changes. |
| 0.1.0 | 11/12/2010 | Carlos Perez | Cleanup |
| 0.2.0 | 11/17/2010 | Carlos Perez | Aligned with LSDAM |
| 1.0 | 11/29/2010 | George & Carlos | Base lined and send out for design review |
| 1.0.1 | 12/01/2010 | Carlos | Updates from Srikalyan comments |
| 1.0.2 | 12/01/2010 | Srikalyan | Added registerSite |
| 1.0.3 | 12/06/2010 | Carlos | Updated based on Denis comments |
| 1.0.4 | 12/21/2010 | Jithu Ramesh | Formatting update. |
| 1.0.5 | 2/21/2010 | Carlos Perez | Update based on Reviews |

Table of Contents

[1 Introduction 5](#_Toc316998531)

[1.1 Overview 5](#_Toc316998532)

[1.2 Relationship to Standards 5](#_Toc316998533)

[1.3 Relation to the Specimen Identifier Conceptual Functional Service Specification 6](#_Toc316998534)

[1.3.1 Conformance and Compliance 6](#_Toc316998535)

[2 Platform Independent Model and Service Specification 7](#_Toc316998536)

[2.1 Overview and Architecture 7](#_Toc316998537)

[2.2 Implementation Considerations 7](#_Toc316998538)

[2.2.1 Assumptions 7](#_Toc316998539)

[2.2.2 Deployment Considerations 7](#_Toc316998540)

[2.2.3 Jurisdictional Domains 7](#_Toc316998541)

[2.3 Information Model 8](#_Toc316998542)

[2.3.1 MaterialIdentifier Type 8](#_Toc316998543)

[2.3.2 Material Type 9](#_Toc316998544)

[2.3.3 MaterialIdentifier State Model 9](#_Toc316998545)

[2.4 Control Data Type Definitions 9](#_Toc316998546)

[2.4.1 Status/Return Values/Exceptions 9](#_Toc316998547)

[2.4.2 Error Codes 10](#_Toc316998548)

[2.4.3 Error Implementation Model 10](#_Toc316998549)

[2.5 Control Data Type Definitions 11](#_Toc316998550)

[2.5.1 Error Codes 11](#_Toc316998551)

[2.5.2 Query Parameters 11](#_Toc316998552)

[3 Interfaces 11](#_Toc316998553)

[3.1 UML Model of Interfaces 11](#_Toc316998554)

[3.2 Operations (Enumeration) 12](#_Toc316998555)

[3.3 Operation Behavior Descriptions 12](#_Toc316998556)

[3.3.1 SpecimenIdentifierAuthority Interface 12](#_Toc316998557)

[4 Dynamic Model 17](#_Toc316998558)

[5 Profiles 18](#_Toc316998559)

[5.1 Functional Profiles 18](#_Toc316998560)

[6 Relationship with other services 18](#_Toc316998561)

[7 Conformance Statements 19](#_Toc316998562)

[8 Appendix A – References 21](#_Toc316998563)

[9 Appendix B – Glossary 21](#_Toc316998564)

[10 Appendix C - Cross Reference Tables 21](#_Toc316998565)

[10.1 Interface to Profile Mapping 21](#_Toc316998566)

[10.2 Interface Operations to Profile Operations 21](#_Toc316998567)

# Introduction

## Overview

This service will manage relationships between specimens across multiple sites/applications to facilitate interoperability. In general, one should be able to get further information about some biospecimen from some distributed set of information services by resolving an identifier that is globally unique. A globally unique identifier is required in order to track a biospecimen’s information as it is processed by different information services.

Biospecimens are commonly shipped from one location to another. The receiving location must be able to electronically acquire relevant biospecimen data (class, type, pathology status, clinical annotations) about the received biospecimens or their parents from the shipping or any other site.

As researchers query for relevant biospecimens globally, biospecimens, their parent and their children must be associated to one or more biospecimen informatics systems and their annotations must be readily available from all of these systems in order to determine which biospecimens are most relevant to the investigator.

Microarray, proteomics, and next generation sequencing cores are processing and analyzing biospecimens using genome- and proteome-wide profiling technologies. Each of these cores utilizes one or more laboratory information management systems (LIMS) and data delivery systems (e.g. caArray). In the course of experiments performed by these cores, it is also likely that biospecimens will be pooled and combined. These systems must be able to track the provenance of biospecimens that are received, processed, and analyzed.

## Relationship to Standards

The GSID service conforms with existing web standards and protocols, providing a natural approach to identifier adoption and resolution.

|  |  |
| --- | --- |
| **Standards** | **Description** |
| NCI’s LSDAM v2.0 | The caBIG Life Sciences Domain Analysis Model (LS DAM) is a shared view of the semantics of the Life Sciences domains: Integrative Cancer Research (ICR), Tissue Banking and Pathology Tools (TBPT), Imaging, is aligned where appropriate, with the Clinical Sciences workspace through BRIDG. |
| Open Provenance Model | The Open Provenance Model is a model of provenance that is designed to allow provenance information to be exchanged between systems. |
| PURL | PURLs (**P**ersistent **U**niform **R**esource **L**ocators) are Web addresses that act as permanent identifiers in the face of a dynamic and changing Web infrastructure. |
| UUID | Universal Unique Identifier (UUID) DCE Version 3 and 5. |

## Relation to the Specimen Identifier Conceptual Functional Service Specification

|  |  |  |
| --- | --- | --- |
| **Conceptual Functional Service Specification Name** | **Conceptual Functional Service Specification Version** | **Description & Link to the Conceptual Functional Service Specification** |
| CS13: Specimen Identifier Management Service | 1.0.5 | The Global Specimen Identifier Conceptual Information Model. |

|  |  |
| --- | --- |
| **Deviation from the Conceptual Functional Service Specification** | **Reason for Deviation** |
| None | Not Applicable |

### Conformance and Compliance

|  |  |
| --- | --- |
| **Conformance Profile No.** | SIDM-CP1 |
| **Conformance Profile Name** | GSID Query |
| **Functional Profiles** | |  |  | | --- | --- | | **Functional Profile No.** | **Functional Profile Name** | | SIDM-FP1 | GSID Query | |
| **Semantic Profiles** | |  |  | | --- | --- | | **Semantic Profile No.** | **Semantic Profile Name** | | SIDM-SP1 | GSID v1.0 Query | |

|  |  |
| --- | --- |
| **Conformance Profile No.** | SIDM-CP2 |
| **Conformance Profile Name** | GSID Edit |
| **Functional Profiles** | |  |  | | --- | --- | | **Functional Profile No.** | **Functional Profile Name** | | SIDM-FP2 | GSID Edit | |
| **Semantic Profiles** | |  |  | | --- | --- | | **Semantic Profile No.** | **Semantic Profile Name** | | SIDM-SP2 | GSID v1.0 Edit | |

# Platform Independent Model and Service Specification

## Overview and Architecture

The *Specimen Identifier Management (SIDM)*is a service that issues and manages biospecimen identifiers, their derivations from other biospecimens and reference information regarding the external systems that contain information about the biospecimen identifier.

The *Specimen Identifier Management* service provides a caBIG grid interface for administration and identifier resolution purposes.

## Implementation Considerations

Some of the framework components are optional. The Global Specimen Identifier service is the only required component, which provides, at minimum, resolution services for identifiers via HTTP.

|  |  |
| --- | --- |
| **Implementation Consideration** | **Impact** |
| ISO data types may need to be constrained for the implementation | The service developed might not utilize all ISO 21090 data types |
| Service to use NCI’s version of LSDAM v2.0 | The service developed might not be fully compliant with other LSDAM v2.0 systems |

### Assumptions

|  |  |
| --- | --- |
| **Assumptions** | **Affects** |
| Audit & Logging is assumed to be handled by supporting infrastructure | It is assumed that any requirements for logging or auditing of operation calls will be handled by supporting infrastructure, so this has not been referenced in the individual operations. |
| Authorization checks are enforced | Only authorized users will be allowed access |

### Deployment Considerations

|  |  |
| --- | --- |
| **Deployment Consideration** | **Impact** |
| Enterprise | The Global Specimen Identifier Service is assumed to be run at the NCI data center and should be accessible external systems through the World Wide Web via HTTP. |

### Jurisdictional Domains

|  |  |
| --- | --- |
| **Jurisdictional Domains** | **Boundary** |
| NCI’s OCIO | The service will be deployed within NCI network |

*Note: Since the GSID service is hosted at NCI, it falls under the jurisdictions of the office of the Chief Information Officer (CIO). Since NCI it a federal agency, it is also under the jurisdiction of US Federal government.*

## Information Model

### MaterialIdentifier Type



***Figure 1. Diagram of Informational Model***

### Material Type



***Figure 2. Diagram of Material***

### MaterialIdentifier State Model



***Figure 3. Material Identifier Model***

## Control Data Type Definitions

### Status/Return Values/Exceptions

|  |  |  |
| --- | --- | --- |
| **Return** | **Operation** | **Exceptions** |
| N/A | addSite | NamingAuthorityConfigurationException  NamingAuthoritySecurityException |
| ST | registerGSID | NamingAuthorityConfigurationException  NamingAuthoritySecurityException |
| COLL<ST> | generateIdentifiers | NamingAuthorityConfigurationException  NamingAuthoritySecurityException |
| BL | validateIdentifier | NamingAuthorityConfigurationException  NamingAuthoritySecurityException |
| Tree | getParentHierarchy | InvalidIdentifierException  NamingAuthorityConfigurationException  NamingAuthoritySecurityException |
| Tree | getChildHierarchy | InvalidIdentifierException  NamingAuthorityConfigurationException  NamingAuthoritySecurityException |

### Error Codes

### Error Implementation Model



***Figure 4. Diagram of Error Code***

|  |  |  |
| --- | --- | --- |
| **Attribute Name** | **Type** | **Attribute Description** |
| code | ST | The actual error code which the exception class represents |
| message | ST | The detailed error message |
| severity | ST | The severity of this error. It can be “FATAL”, “ERROR”, “INFO” or “WARNING” |
| type | ST | Indicates what type of error has occurred. E.g. Business, System etc. |

## Control Data Type Definitions

### Error Codes

|  |  |
| --- | --- |
| **Error Code** | **Error/Exception Description** |
| NamingAuthorityConfigurationException | A configuration error has been detected by the naming authority while serving a request. |
| InvalidIdentifierException | The identifier provided is invalid or does not exist. |
| InvalidIdentifierValuesException | The identifier metadata provided is invalid. |

### Query Parameters

|  |  |
| --- | --- |
| **Operation** | **Parameters** |
| getKeyNames | URI Identifier |
| getKeyData | URI Identifier, ST KeyName |
| resolveIdentifier | URI Identifier |

# Interfaces

## UML Model of Interfaces



***Figure 5. Diagram of UML Model.***

|  |  |  |
| --- | --- | --- |
| **Interface No** | **Interface Name** | **Interface Description** |
| SIDM-INF1 | SpecimenIdentifierAuthority | Provides functionality to manage Global Specimen Identifiers. |

## Operations (Enumeration)

|  |  |  |  |
| --- | --- | --- | --- |
| **Operation No.** | **Operation Name** | **Interface Name** | **Operation Description** |
| SIDM-INF1-OP1 | registerGSID | SpecimenIdentifierAuthority | Register a new GSID. |
| SIDM-INF1-OP2 | addSite | SpecimenIdentifierAuthority | Register a new site. |
| SIDM-INF1-OP3 | generateIdentifiers | SpecimenIdentifierAuthority | Generate a list of UUIDs. |
| SIDM-INF1-OP4 | validateIdentifier | SpecimenIdentifierAuthority | Validate an identifier. |
| SIDM-INF1-OP5 | getParentHierarchy | SpecimenIdentifierAuthority | Get Parent Hierarchy of identifier. |
| SIDM-INF1-OP6 | getChildHierarchy | SpecimenIdentifierAuthority | Get Children of an identifier. |
| SIDM-INF1-OP7 | registerSite | SpecimenIdentifierAuthority | Register a new Site. |

## Operation Behavior Descriptions

### SpecimenIdentifierAuthority Interface

#### registerGSID

Register a new Specimen Identifier.

|  |  |
| --- | --- |
| **Behavior Description** | This operation registers a Specimen Identifier. |
| **Pre-Conditions** |  |
| **Security Conditions** | A user can create identifiers if all of the below conditions are met:   1. The user has been given explicit authorization to register identifiers. 2. User has registered a site with his credentials. |
| **Inputs** | * ST suggestedIdentifer * COLL<ST> parentIdentifiers |
| **Outputs** | Identifier |
| **Post-Conditions** |  |
| **Exception Conditions** | NamingAuthorityConfigurationException  InvalidIdentifierValuesException  InvalidIdentifierException  NamingAuthoritySecurityException |
| **Additional Implementation Details** |  |
| **Notes** |  |

#### addSite

Add the user as a Site that has information on the GSID.

|  |  |
| --- | --- |
| **Behavior Description** | Add the current user as a site that has information about the input specimen identifier. |
| **Pre-Conditions** |  |
| **Security Conditions** | A user can create identifiers if the below condition is met:   * The user has been given explicit authorization to register identifiers. |
| **Inputs** | * ST identifer |
| **Outputs** | None |
| **Post-Conditions** |  |
| **Exception Conditions** | NamingAuthorityConfigurationException  InvalidIdentifierValuesException  InvalidIdentifierException  NamingAuthoritySecurityException |
| **Additional Implementation Details** |  |
| **Notes** |  |

#### generateIdentifiers

Create a batch of Identifiers.

|  |  |
| --- | --- |
| **Behavior Description** | Generates a batch of valid UUIDs. |
| **Pre-Conditions** | None |
| **Security Pre-Conditions** | A user can create identifiers if the below condition is met:   * The user has been given explicit authorization to register identifiers. |
| **Inputs** | * Number of UUIDs to generate. |
| **Outputs** | BAG<II> batchOfIdentifiers |
| **Post-Conditions** | None |
| **Alternate Conditions** | None |
| **Exception Conditions** | NamingAuthorityConfigurationException  NamingAuthoritySecurityException |
| **Additional Details** | None |
| **Notes** | None |

#### validateIdentifier

Validate Identifier.

|  |  |
| --- | --- |
| **Behavior Description** | Validates if the identifier can be registered as unique by the service. |
| **Pre-Conditions** | None |
| **Security Pre-Conditions** | A user can create identifiers if the below condition is met:   * The user has been given explicit authorization to register identifiers. |
| **Inputs** | ST identifier |
| **Outputs** | BLBL |
| **Post-Conditions** | None |
| **Alternate Conditions** | None |
| **Exception Conditions** | NamingAuthorityConfigurationException  NamingAuthoritySecurityException |
| **Additional Details** | None |
| **Notes** | None |

#### getParentHierarchy

Return parents of a Specimen Identifier

|  |  |
| --- | --- |
| **Behavior Description** | Return the parent hierarchy of the input Specimen Identifier. |
| **Pre-Conditions** | None |
| **Security Pre-Conditions** | None |
| **Inputs** | ST identifieridentifier |
| **Outputs** | Tree representing Identifiers Identifiers starting from a given Identifier. |
| **Post-Conditions** | None |
| **Alternate Conditions** | None |
| **Exception Conditions** | NamingAuthorityConfigurationException  InvalidIdentifierValuesException  InvalidIdentifierException  NamingAuthoritySecurityException |
| **Additional Details** | None |
| **Notes** | None |

#### getChildHierarchy

Return children of a Specimen Identifier.

|  |  |
| --- | --- |
| **Behavior Description** | Return the parent hierarchy of the input Specimen Identifier. |
| **Pre-Conditions** | None |
| **Security Pre-Conditions** | None |
| **Inputs** | ST identifier.ST identifier. |
| **Outputs** | Tree representing Identifiers starting from given Identifier. |
| **Post-Conditions** | None |
| **Alternate Conditions** | None |
| **Exception Conditions** | NamingAuthorityConfigurationException  InvalidIdentifierValuesException  InvalidIdentifierException  NamingAuthoritySecurityException |
| **Additional Details** | None |
| **Notes** | None |

#### registerSite

Register a site.

|  |  |
| --- | --- |
| **Behavior Description** | This operation registers a Site to the user. |
| **Pre-Conditions** |  |
| **Security Conditions** | A user can register a site if all of the below conditions are met:   1. The user has been given explicit authorization to create a site. 2. The user did not register any other site previously with his credentials. |
| **Inputs** | * ST application * ST application URL * ST application version * ST contact email * ST contact name * ST contact phone * ST organization |
| **Outputs** | None. |
| **Post-Conditions** |  |
| **Exception Conditions** | NamingAuthorityConfigurationException  InvalidIdentifierValuesException  InvalidIdentifierException  NamingAuthoritySecurityException |
| **Additional Implementation Details** |  |
| **Notes** |  |

# Dynamic Model



***Figure 6. Diagram highlights the interaction of various services with the Specimen Identifier Management Service.***

A Specimen Management System receives security credentials using the Security Framework. Using these credentials, it registers a Specimen Identifier Management Service. At a later time a client (typically another specimen management system) accesses the Specimen Identifier Service to resolve a specimen identifier. Typically the client already knows of the Specimen Identifier Management service, but in the event that it does not, it uses a prefix Authority to discover the service. Once having access to Specimen Identifier Management service, it can retrieve dependency information of a Specimen Identifier or it can access further information by using the Site information associated with a Specimen Identifier. The Site information should provide information as to which information service is aware of the biospecimen and using the Specimen Identifier, one may use it as a key to retrieve additional information from the information service.

# Profiles

## Functional Profiles

|  |  |  |
| --- | --- | --- |
| **Profile No.** | **Profile Name** | **Profile Description** |
| SIDM-FP1 | GSID Query | Contains functionality for identifier resolution. |
| SIDM-FP2 | GSID Edit | Contains functionality to administrate identifiers and their data. |

|  |  |  |  |
| --- | --- | --- | --- |
| **Interface Operation No.** | **Interface Operation Name** | **Profile No.** | **Profile Name** |
| SIDM-INF1-OP1 | registerGSID | SIDM-FP1 | Register a new GSID. |
| SIDM-INF1-OP2 | addSite | SIDM-FP1 | Add a site |
| SIDM-INF1-OP3 | generateIdentifiers | SIDM-FP1 | Generate a list of UUIDs. |
| SIDM-INF1-OP4 | validateIdentifier | SIDM-FP1 | Validate an identifier. |
| SIDM-INF1-OP5 | getParentHierarchy | SIDM-FP1 | Get Parent Hierarchy of identifier. |
| SIDM-INF1-OP6 | getChildHierarchy | SIDM-FP1 | Get Children of an identifier. |
| SIDM-INF1-OP7 | registerSite | SIDM-FP1 | Register a site |

# Relationship with other services

The Specimen Identifier Service is dependent on two other services; specifically the Identifier Framework and the Security Framework.



***Figure 7. Diagram depicts the relationship between Specimen Identifier Service, Identifier Framework, and Security Framework.***

|  |  |  |  |
| --- | --- | --- | --- |
| **Service / System / Actors Name** | **Dependant / Depends** | **Data** | **Description** |
| Identifiers Framework | Depends |  |  |
| Security Framework | Depends |  |  |

# Conformance Statements

The identifiers framework conceptual model conforms with existing web standards and protocols, providing a natural approach to identifier adoption and resolution.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **No** | **Name** | **Type** | **Viewpoint** | **Description** | **Test method** |
| 1 | Query Performance | Obligation | Engineering | The GSID service should provide a response within 0.5 seconds to support a synchronous UI based client | Test cases to include performance testing. |
| 2 | Semantic Model | Obligation | Informational | The GSID service must provide traceability to classes in the LS PIM as identified in Section 2.3. | Design Review |
| 3 | Data Types | Obligation | Informational | The GSID service must confirm to NCI’s constrained list of ISO 21090 data types. | Design Review |
| 4 | Functional Profiles | Obligation | Computational | Functional Profiles shall be deployed as functional wholes. Ignoring or omitting functional behavior defined within a functional profile is not permitted, nor is diverging from the detailed functional specifications provided in Section 5. | 1. Design Review 2. Test cases |
| 5 | Functional Profiles – Conformant Implementation | Obligation | Computational | A conformant implementation of this specification must deploy at least one Functional Profile. If that Functional Profile has dependencies on other profiles, then those dependencies must be deployed as well to support the provenance of that service instance. | 1. Design Review 2. Test cases |

# Appendix A – References

|  |  |
| --- | --- |
| **Description** | **Location** |
| caGrid Identifiers Framework Design Document v1.0  Scott Oster  Calixto Melean | <https://ncisvn.nci.nih.gov/svn/cagrid/trunk/cagrid/Documentation/core/Identifiers/caBIG_IdentifierSvcFramework_Design.docx> |
| caIdentifiers Framework Conceptual Functional Service Specification v1.0 | <https://ncisvn.nci.nih.gov/svn/cagrid/trunk/cagrid/Documentation/core/Identifiers/caBIG_Conceptual_Functional_Service_Specification.doc> |

# Appendix B – Glossary

|  |  |
| --- | --- |
| **Term** | **Description** |
| GSID | Global Specimen Identifier Service |

# Appendix C - Cross Reference Tables

## Interface to Profile Mapping

|  |  |  |  |
| --- | --- | --- | --- |
| **Interface No.** | **Interface** | **Profile No.** | **Profile** |
| SIDM-INF1 | SpecimenIdentifierAuthority |  |  |

## Interface Operations to Profile Operations

|  |  |  |  |
| --- | --- | --- | --- |
| **Operation No.** | **Operation Name** | **Capability No.** | **Profile Capability** |
| SIDM-INF1-OP1 | registerGSID | SIDM-FP2-C1 | Register a new GSID. |
| SIDM-INF1-OP2 | addSite | SIDM-FP2-C2 |  |
| SIDM-INF1-OP3 | generateIdentifiers | SIDM-FP2-C3 | Generate a list of UUIDs. |
| SIDM-INF1-OP4 | validateIdentifier | SIDM-FP2-C4 | Validate an identifier. |
| SIDM-INF1-OP5 | getParentHierarchy | SIDM-FP2-C5 | Get Parent Hierarchy of identifier. |
| SIDM-INF1-OP6 | getChildHierarchy | SIDM-FP2-C6 | Get Children of an identifier. |
| SIDM-INF1-OP7 | registerSite | SIDM-FP2-C7 | Register a site |