# Imaging Identifier Service (IDS) Use Cases, Requirements, Analysis Draft version 0.1

Note: Version 0.1 of the IDS Use Case Document does not include an in depth analysis of security requirements of IDS.

Note: Version 0.1 of the IDS Use Case Document does not include an in depth analysis of federation requirements of IDS

## Description (ref: IMAG\_Services\_Description\_20100610\_Draft.xlsx):

“The IDS Services provide a global unique and unified Identification Information on any imaging objects from IMAG Workspace including subject, participants, study, AIM, and etc. The IDs will leverage caGRID ID Services/Management (from caGrid v1.4) and consolidate with DICOM ID. Imaging ID information will contain a set of imaging specific attributes to support the IDs searches, indexes, categories/groups, and internal and external references. The IDs services are critical in the IMAG Application Suit seamless Integration as well as various imaging service operations for the interoperability and referencing.”

### Comments (ref: IMAG\_Services\_Description\_20100610\_Draft.xlsx):

“The IDs Services reuse caGrid ID Services/management and allow creation, readonly, update, and delete (CRUD) operations as well as establishment of mapping and/or reference to external IDs for the sync of image related data residing in different and distributed locations. A set of ID could be managed and handled as a group/cohesively set. The IDs services require extreme high performance, reliability/availability under security. Other needed specific considerations include the federation of multiple instances, deployment of Imaging ID Services over Cloud Computing, and the IDs Services load balance and standby for the redundancy. The IDs Services can be a collaborative work with caTissue ID Services.”

## Use Cases:

### High Level Use Case Narrative:

1. A study coordinator enrolls subjects for an imaging study and orders imaging for the subject.
2. An imaging technician acquires images for the subject using specified imaging protocols and stores images in clinical or research PACS for clinical and research use.
3. An imaging researcher identifies and retrieves a set of images from a PACS server (research PACS or institutional PACS) and stores the images in sSOA-based research image archive.
4. The imaging researcher performs a search on the image archive for a subset of images, filtering on image metadata.
5. The imaging researcher retrieves search results for human and CAD analysis, the outputs of which are image annotations.
6. A reviewer performs a federated search for images and annotations matching specified image metadata and annotation attributes, and retrieves both images and annotations for display and review.
7. A biostatistician performs a federated query across multiple study sites for image metadata and annotations for a study and conduct statistical analysis on the data.

### Low Level Use Cases:

#### Imaging Identifier Creation

|  |  |  |  |
| --- | --- | --- | --- |
| Use Case Name | Create Imaging Identifier for existing data | Use Case Unique ID | **1.1** |
| Brief Description | IDS creates an imaging identifier from existing data or reference to data | | |
| Actors | Data Service, IDS | | |
| Trigger | A record (trial subject, imaging order, image, annotation) is submitted/created/registered in a Data Service | | |
| Preconditions | * Imaging data has been submitted to the Data Service | | |
| Main Flow | 1. Data Service gathers relevant attributes of the imaging data 2. Data Service submits request for new identifier to IDS with imaging data attributes as parameter 3. IDS creates new identifier and stores imaging data attributes 4. IDS returns new identifier to Data Service 5. Data Service stores identifier with the data | | |
| Alternate Flow | 2a. Data Service sends existing reference such as unique identifier in external system to IDS.  5a. Data Service may choose to not store identifier | | |
| Postconditions | * Identifier exists in IDS with mapping to metadata. * Data Service has received new identifier | | |
| Notes | * Example of existing identifier may be one assigned by DICOM modality systems. * A User may act in place of a Data Service by carrying out the flow using IDS web interface. * Authorization policy on the newly created identifier needs to be investigated * When an identifier exists with the same metadata, IDS behavior is unspecified | | |

|  |  |  |  |
| --- | --- | --- | --- |
| Use Case Name | Create Imaging Identifier for New Data | Use Case Unique ID | **1.2** |
| Brief Description | IDS creates an image identifier as a place holder for anticipated new data. | | |
| Actors | Application, IDS | | |
| Trigger | The creation of a new record (trial subject, imaging order, image, annotation) is starting in an application | | |
| Preconditions |  | | |
| Main Flow | 1. Application submits request for new identifier to IDS with no parameter 2. IDS creates new identifier with “minimal information” attribute. 3. IDS returns new identifier to Application 4. Application associates identifier with data and uses identifier during data updates. | | |
| Alternate Flow |  | | |
| Postconditions | * Identifier exists in IDS with attribute that indicates it’s an Identifier with minimal information. * Application has received new identifier | | |
| Notes | * Example of application may be an imaging workstation during secondary image capture. * A User may act in place of an Application by carrying out the flow using IDS web interface. * Subsequent Update operations add to the metadata * Authorization policy on the newly created identifier needs to be investigated | | |

|  |  |  |  |
| --- | --- | --- | --- |
| Use Case Name | Create Imaging Identifier from Existing Imaging Identifiers | Use Case Unique ID | **1.3** |
| Brief Description | IDS creates an imaging identifier from existing IDS identifiers, to represent relationships between identifiers and their underlying data. | | |
| Actors | Application, IDS | | |
| Trigger | * A user requests that an application create a group identifier for one or more IDS identifiers, or * An application creates secondary data and needs to indicate the derivative relationship | | |
| Preconditions | * IDS has the target IDS identifiers | | |
| Main Flow | 1. Application submits request for new identifier to IDS with existing IDS identifiers and the relationship type (part-of, derivative) as parameters 2. IDS creates new identifier and stores mapping and relationship types to the IDS identifiers 3. IDS returns new identifier to Application 4. Application associates identifier with data SET in subsequent operations | | |
| Alternate Flow | 1a. Application sends query statement instead of existing IDS identifiers.  4a. Application may choose to not store identifier | | |
| Postconditions | * Identifier exists in IDS with mapping to other IDS identifiers. * Application has new group identifier | | |
| Notes | * New group Identifier may have its own metadata as well, e.g. for secondary image captures. * A User may act in place of an Application by carrying out the flow using IDS web interface. * If query statement (e.g. CQL) is used, it is expected that the data the IDS group identifier refers to is a mutable set. * Authorization policy on the newly created identifier needs to be investigated * When an identifier exists with the same child identifiers, IDS behavior is unspecified | | |

#### Identifier Query

|  |  |  |  |
| --- | --- | --- | --- |
| Use Case Name | Query for Imaging Identifier using metadata | Use Case Unique ID | **2.1** |
| Brief Description | The IDS service is queried for imaging identifiers based on matching metadata | | |
| Actors | Application, IDS | | |
| Trigger | * User initiates a query request in the Application | | |
| Preconditions | * User has logged in to the Application | | |
| Main Flow | 1. Application sends query request to IDS, containing the user’s credential and the metadata attribute values to match 2. IDS checks authorization of the request 3. IDS performs search for the matching identifier 4. IDS returns matching identifiers to the Application 5. Application receives the identifiers | | |
| Alternate Flow | 1a. Application does not have user credential. Application generates exception and warn User  2a. IDS does not authorize User. Return “unauthorized” exception  3a. IDS does not find results. Return “no results” indication. | | |
| Postconditions | Application has identifiers | | |
| Notes | * IDS should perform authorization on the query operation to prevent brute force attack on the metadata * There may be more than 1 matching identifiers * A User may act in place of a Application by carrying out the flow using IDS web interface. * A Service may act in place of an Application in the workflow, using delegated user credentials | | |

|  |  |  |  |
| --- | --- | --- | --- |
| Use Case Name | Query for metadata using Identifier | Use Case Unique ID | **2.2** |
| Brief Description | The IDS service is queried for imaging metadata associated with an identifier | | |
| Actors | Application, IDS | | |
| Trigger | * User initiates a query request in the Application | | |
| Preconditions | * User has logged in to the Application | | |
| Main Flow | 1. Application sends query request to IDS, containing the user’s credential and the identifier 2. IDS checks authorization of the request 3. IDS performs search for the corresponding metadata 4. IDS returns metadata to the Application 5. Application receives the metadata | | |
| Alternate Flow | 1a. Application does not have user credential. Application generates exception and warn User  2a. IDS does not authorize User. Return “unauthorized” exception  3a. IDS does not manage this identifier. Return “no such identifier” indication.  3b. IDS does not have metadata for this identifier. Return “no metadata” indication | | |
| Postconditions | Application has metadata | | |
| Notes | * IDS should perform authorization on the query operation to prevent brute force attack on the metadata * A User may act in place of an Application by carrying out the flow using IDS web interface. * A Service may act in place of an Application in the workflow, using delegated user credentials | | |

|  |  |  |  |
| --- | --- | --- | --- |
| Use Case Name | Query for group Identifiers using member Identifiers | Use Case Unique ID | **2.3** |
| Brief Description | The IDS service is queried for group identifiers whose corresponding groups contain the specified member identifiers with the specified relationship | | |
| Actors | Application, IDS | | |
| Trigger | * User initiates a query request in the Application | | |
| Preconditions | * User has logged in to the Application * Application has member Identifier | | |
| Main Flow | 1. Application sends query request to IDS, containing the user’s credential, one or more member Identifiers, and the relationship type 2. IDS checks authorization of the request 3. IDS performs search for group identifiers 4. IDS returns matching identifiers to the Application 5. Application receives the identifiers | | |
| Alternate Flow | 1a. Application does not have user credential. Application generates exception and warn User  2a. IDS does not authorize User. Return “unauthorized” exception  3a. IDS does not find results. Return “no results” indication.  3b. IDS finds circular reference. Return “circular reference” indication  3c. IDS does not manage this identifier. Return “no such identifier” indication. | | |
| Postconditions | Application has identifiers | | |
| Notes | * There may be more than 1 matching identifiers * Relationship types may include “part-of”, “derived from”, etc. * A User may act in place of an Application by carrying out the flow using IDS web interface. * A Service may act in place of an Application in the workflow, using delegated user credentials * IDS should return identifiers for nested groups to which the specified identifiers are member. | | |

|  |  |  |  |
| --- | --- | --- | --- |
| Use Case Name | Query for child identifiers using group Identifier | Use Case Unique ID | **2.4** |
| Brief Description | The IDS service is queried for child identifiers associated with a group identifier | | |
| Actors | Application, IDS | | |
| Trigger | * User initiates a query request in the Application | | |
| Preconditions | * User has logged in to the Application | | |
| Main Flow | 1. Application sends query request to IDS, containing the user’s credential, the group identifier, and the relationship type 2. IDS checks authorization of the request 3. IDS performs search for the child identifiers 4. IDS returns identifiers to the Application 5. Application receives the identifiers | | |
| Alternate Flow | 1a. Application does not have user credential. Application generates exception and warn User  2a. IDS does not authorize User. Return “unauthorized” exception  3a. IDS does not manage this identifier. Return “no such identifier” indication.  3b. There are no children for this identifier. Return “no result” indication  3c. IDS finds circular reference. Return “circular reference” indication | | |
| Postconditions | Application has identifiers | | |
| Notes | * There may be more than 1 matching identifiers * Relationship types may include “part-of”, “derived from”, etc. * A User may act in place of an Application by carrying out the flow using IDS web interface. * A Service may act in place of an Application in the workflow, using delegated user credentials * IDS should return member identifiers for nested groups that are descendants of the requested group identifier. * Returned value may be a query statement instead of identifiers | | |

|  |  |  |  |
| --- | --- | --- | --- |
| Use Case Name | Query for child identifiers’ metadata using group Identifier | Use Case Unique ID | **2.5** |
| Brief Description | The IDS service is queried for metadata of the child identifiers associated with a group identifier | | |
| Actors | Application, IDS | | |
| Trigger | * User initiates a query request in the Application | | |
| Preconditions | * User has logged in to the Application | | |
| Main Flow | 1. Application sends query request to IDS, containing the user’s credential, the group identifier, and the relationship type 2. IDS checks authorization of the request 3. IDS performs search for the child identifiers 4. IDS performs search for the metadata of the child identifiers 5. IDS returns metadata to the Application 6. Application receives the metadata | | |
| Alternate Flow | 1a. Application does not have user credential. Application generates exception and warn User  2a. IDS does not authorize User. Return “unauthorized” exception  3a. IDS does not manage this identifier. Return “no such identifier” indication.  3b. There are no children for this identifier. Return “no result” indication  3c. IDS finds circular reference. Return “circular reference” indication  4a. IDS does not manage the child identifiers. Return “no such identifier” indication  4b. IDS does not have metadata for the child identifiers. Return “no result” indication | | |
| Postconditions | Application has metadata | | |
| Notes | * IDS should perform authorization on the query operation to prevent brute force attack on the metadata * There may be more than 1 matching identifiers * Relationship types may include “part-of”, “derived from”, etc. * A User may act in place of an Application by carrying out the flow using IDS web interface. * A Service may act in place of an Application in the workflow, using delegated user credentials * IDS should return member identifiers’ metadata for nested groups that are descendants of the requested group identifier. | | |

#### Identifier Update

1. Update metadata
2. Update group membership

|  |  |  |  |
| --- | --- | --- | --- |
| Use Case Name | Update Imaging Identifier with metadata | Use Case Unique ID | **3.1** |
| Brief Description | IDS updates the metadata associated with an imaging identifier | | |
| Actors | Data Service, IDS | | |
| Trigger | A record (trial subject, imaging order, image, annotation) is updated in a Data Service | | |
| Preconditions | User has logged in to the Data Service | | |
| Main Flow | 1. Data Service gathers updated attributes of the imaging data 2. Data Service submits update request with Identifier and updated metadata attributes as parameter 3. IDS checks authorization of the request 4. IDS updates identifier’s metadata 5. IDS returns acknowledgement to Data Service | | |
| Alternate Flow | 2a. Data Service updates external references such as unique identifier in external system to IDS.  3a. IDS does not authorize user. Return “unauthorized” exception  4a. IDS does not manage the identifier. Returns “no such identifier” exception | | |
| Postconditions | * Identifier updated in IDS with new metadata. | | |
| Notes | * Example of existing identifier may be one assigned by DICOM modality systems. * A User may act in place of a Data Service by carrying out the flow using IDS web interface. * When an identifier exists with the same metadata, IDS behavior is unspecified * Versioning policy in IDS is unspecified. | | |

|  |  |  |  |
| --- | --- | --- | --- |
| Use Case Name | Update group Identifier with new or replace child identifiers | Use Case Unique ID | **3.2** |
| Brief Description | IDS updates the child identifiers associated with an group identifier | | |
| Actors | Application, IDS | | |
| Trigger | User requests update to the group identifier with new child identifier or query statement in the application | | |
| Preconditions | User has logged in to the Application | | |
| Main Flow | 1. Application submits update child identifiers with group Identifier and relationship type as parameter 2. IDS checks authorization of the request 3. IDS updates identifier’s child membership and relationship 4. IDS returns acknowledgement to Application | | |
| Alternate Flow | 1a. Application may submit query statement instead of child identifiers to IDS.  2a. IDS does not authorize user. Return “unauthorized” exception  3a. IDS does not manage the group identifier. Returns “no such identifier” exception | | |
| Postconditions | * Identifier updated in IDS with new child identifiers. | | |
| Notes | * A User may act in place of a Data Service by carrying out the flow using IDS web interface. * When an identifier exists with the same child identifiers, IDS behavior is unspecified * Versioning policy in IDS is unspecified. | | |

#### Identifier Deletion

|  |  |  |  |
| --- | --- | --- | --- |
| Use Case Name | Delete identifier | Use Case Unique ID | **4.1** |
| Brief Description | IDS deletes the identifier and the associated metadata and child identifiers | | |
| Actors | Application, IDS | | |
| Trigger | User requests an identifier and the associated mapping to metadata and child identifiers be delete from IDS. | | |
| Preconditions | User has logged in to the Application | | |
| Main Flow | 1. Application submits delete request with identifier. 2. IDS checks authorization of the request 3. IDS deletes the metadata and mappings to child identifiers for the identifier, and the identifier itself. 4. IDS returns acknowledgement to Application | | |
| Alternate Flow | 2a. IDS does not authorize user. Return “unauthorized” exception  3a. IDS does not manage the group identifier. Returns “no such identifier” exception | | |
| Postconditions | * Identifier deleted from IDS, and its associated mapping to child identifiers, and metadata and query statement * Child identifiers are not deleted. | | |
| Notes | * A User may act in place of an Application by carrying out the flow using IDS web interface. * Versioning policy in IDS is unspecified. | | |

## Requirements:

1. Provide mapping between grid identifier and external identifier

Data stored in external systems may be referenced in caBIG environment. IDS manages external identifier to grid identifier mapping so data in external systems can be explicitly located (access mechanism is out of scope here). External identifier includes the system unique identifier as well as the system URL. External identifier may be of any format, including UUID, OID, URI, LSID, etc. Identifiers themselves may need to use OIDs to provide DICOM compatibility.

1. Provide mapping between metadata and grid identifier

Data managed in caBIG environment should be locatable by metadata. Metadata attributes and value domains are defined by application and drawn from LS-DAM and BRIDG. Identifier to metadata mapping provide 2 way discovery of the identifier and corresponding data. Identifiers themselves may need to use OIDs to provide DICOM compatibility. Batched identifier creation may be provide a convenience for external systems.

1. Label identifiers with category information/namespacing

Identifier namespacing provides a mechanism for categorization of the resources referenced by the identifiers. It facilitates scalability and federation of IDS. Data may be organized into categories, each with its own identifier namespacing. Examples include images, annotations, patients, orders.

1. Provide identifier grouping capability

A set of data may be related and it would be useful to be able to identify them as a group. An example of this may be the results of a query, or a set of annotations and their source images. Query results may be defined by the identifiers of the individual entries in the results, or by the query statement. When grouping a set of identifiers, two types of relationships have been identified – derivation and membership. Derivation indicated that one data record is derived from another data record. Membership indicates that one data record is a part of a group. The two relationships are not mutually exclusive.

1. Secure access to IDS

Secure access to IDS follows a complex set of rules with the objective of protecting the metadata directly and any unauthorized inferencing using the identifier relationships. Some of the issues to consider include:

* 1. Identifier creation and read – who are authorized to create, who are the owners, and who can subsequently read the content?
  2. Private information management – what is the minimum amount of metadata to have unique mapping to a unique identifier? What is the maximum amount of metadata that can still maintain anonymity? Should IDS only contain anonymous metadata? How to guarantee external references uses opaque identifier (to maintain data privacy, counter example: SSN)?

1. Scalable performance

As IDS provides a fundamental capability in the imaging services inventory, each data creation, access, update, deletion, etc in the imaging grid would necessarily require interaction with IDS. IDS performance is therefore critical and needs to be scalable to large number of identifiers as well as large number of requests with low latency and high throughput.

1. Utilize existing technology where possible.

It has been specified in the IVI middleware contract that IDS leverages existing technology where possible, specifically the caGrid 1.4 Identifiers Framework.

## Existing caBIG Work

### caGrid 1.4 Identifiers

caGrid 1.4 provides an Identifier design and implementation. It has the following properties:

1. Naming authority handles the identifier to metadata mapping. Metadata is used for query to the actual data source

2. Prefix authority for federating naming authority, specifically using PURL

3. Identity curator is used to generate the metadata for the data. The metadata is submitted to the naming authority for X.

4. Owner of the data identifier is a data source/service

The framework provides basic identifier to metadata mapping. However, there are certain challenges that will need to be addressed in adoption of this identifier framework for use with IDS. The following is a mapping of IDS requirements to caGrid 1.4 identifier capabilities

|  |  |
| --- | --- |
| IDS requirement ID | caGrid 1.4 Identifier Framework Capability mapping to IDS Requirements |
| 1 | Can represent external identifier as metadata. May require special metadata attributes |
| 2 | Satisfied. Generation of metadata requires explicit policies. Identifier itself is limited to UUID format |
| 3 | Require special identifier attributes |
| 4 | Derivation and membership are not supported, and there is no support of relationships between identifiers, including the identifier’s authorization and ownership information. |
| 5 | Actual security capability needs to be assessed based on further requirement specifications |
| 6 | Actual performance needs to be assessed. |
| 7 | satisfied |

### Global Specimens Identifier (GSID)

The TBPT workspace has designed and prototyped a global specimens identifier framework that is based on caGrid 1.4 Identifiers framework. The business problem it sees to address is the need to generate globally unique identifiers in a consistent format for a biospecimens and explicitly represent its relationships to the source and derived biospecimens. To manage this derivation relationship a hierarchical, directed acyclic graph can be constructed. In addition, GSID also enforce a common set of identifier attributes for biospecimens. As the data managed is in a constrained domain, the standardization of the metadata attributes is relatively straight forward. GSID service provides grid service interfaces as well as web based user interface for managing the identifiers. The service utilizes UUID for identifier. GSID is deployed as a secure service with user based authorization. Deployment is assumed to be centralized and is not namespaced as it is limited to the biospecimens domain. GSID provides capability to batch create identifiers for use with local biospecimens management systems.

|  |  |
| --- | --- |
| IDS requirement ID | GSID Capability mapping to IDS Requirements |
| 1 | Supports external identifiers. Metadata is constrained to common set needed for biospecimens. Image IDS would need more flexibility |
| 2 | Relies on underlying caGrid 1.4 identifiers. Identifier itself is limited to UUID format. |
| 3 | No namespacing support, or categorization capabilities |
| 4 | Derivation relationship is supported, and allows one more identifiers to be associated with a new identifier  Membership relationship is not supported. |
| 5 | Actual security capability needs to be assessed based on further requirement specifications |
| 6 | Actual performance needs to be assessed. Centralized deployment may not be sufficient for Imaging IDS where there is a diversity of data with multiple authoritative sources. |
| 7 | satisfied |