Informatics for Integrating Biology and the Bedside

i2b2 Design Document

Identity Management (IM) Cell

Document Version: *1.7.08-002*

i2b2 Software Version: *1.7.08*

TABLE OF CONTENTS

DOCUMENT MANAGEMENT **4**

ABOUT THIS GUIDE **5**

1 IDENTITY MANAGEMENT CONCEPTS **6**

1.1 OBJECTIVES OF IDENTITY MANAGEMENT CELL AND VIEWS 6
1.2 IDENTIFICATION OF USERS 6
1.2.1 Clinical Researcher *6*
1.2.2 Manager of Clinical Researcher *6*
1.2.3 Administrator *6*

2 REQUIREMENTS **8**

2.1 DESIGN REQUIREMENTS 8
2.1.1 Patient Mapping View *8*
2.1.2 Admin Tool View *9*
2.1.2.1 View User Audit 9
2.1.2.2 View Patient Audit 10
2.1.2.3 View All Audit Records for the Project 10
2.2 TABLE USAGE 11
2.3 SECURITY 11

3 IDENTITY MANAGEMENT TABLES **13**

3.1 GENERAL INFORMATION 13
3.2 PROJECT TABLES 13
3.2.1 Project Sites Table *13*
3.2.1.1 Requirements for Project Sites Table 14
3.2.1.2 Column Definitions 14
3.2.2 Project Patients Table *15*
3.2.2.1 Requirements for Project Patients Table 15
3.2.2.2 Column Definitions 15
3.3 MAPPING TABLES 16
3.3.1 Demographics Table *16*
3.3.1.1 Requirements for Demographics Table 16
3.3.1.2 Column Definitions 17
3.3.2 MPI Mapping Table *18*
3.3.2.1 Requirements for MPI Mapping Table 18
3.3.2.2 Column Definitions 18
3.4 SECURITY TABLES 19
3.4.1 Audit Table *19*
3.4.1.1 Requirements for the Audit Table 20
3.4.1.2 Column Definitions 20
3.5 JOINING COLUMNS 21

4 OBJECTS **22**

4.1 PATIENT DATA OBJECT 22
4.2 IM DATA OBJECT 22

5 PATIENT MAPPING SCENARIOS **23**

5.1 SELF-MAPPING 23

6 GLOSSARY 24

6.1 GENERAL TERMS 24

END NOTES 25

DOCUMENT MANAGEMENT

Revision Number	Date	Author	Description of change
1.7.00	03/26/13	Janice Donahoe	Created initial version of document (1.7 software)
1.7.00-001	08/11/2015	Janice Donahoe	Fixed the following: • Spelling and grammar issues. • Table data has been updated to reflect the correct table name and data types. • Updated images to reflect current Admin Tool view for displaying an audit trail.
1.7.08-002	10/04/2016	Janice Donahoe	Fixed revision number.

ABOUT THIS GUIDE

The i2b2 design document describes the requirements, overview of the technical functionality, and the intended capabilities of the **Identity Management (IM) Cell**. This document is to be used as a guideline and continuing reference point as developers write the code and quality assurance writes the test plans.

1 IDENTITY MANAGEMENT CONCEPTS

1.1 Objectives of Identity Management Cell and Views

Information in the Identity Management cell is related to the setup, maintenance and security of patients from heterogeneous sources. This data may be encrypted and is restricted by project and user.

1.2 Identification of Users

1.2.1 Clinical Researcher

- Member of the research team who is setup with access to the project in i2b2.
- Their access role is **USER**.
- They can edit their user profile.

1.2.2 Manager of Clinical Researcher

- Manager of the research team.
- Their access role is **MANAGER**.
- They can create and edit users associated to their project.
- The can create and edit project related information.

1.2.3 Administrator

- They may or may not be part of the research team.
- They are responsible for the administrative tasks related to the i2b2.
- Their access role is **ADMIN** and the project id is @.
- They can create and edit users associated to any project.
- They can create and edit all projects.

- They can create and edit hive information.

2 REQUIREMENTS

2.1 Design Requirements

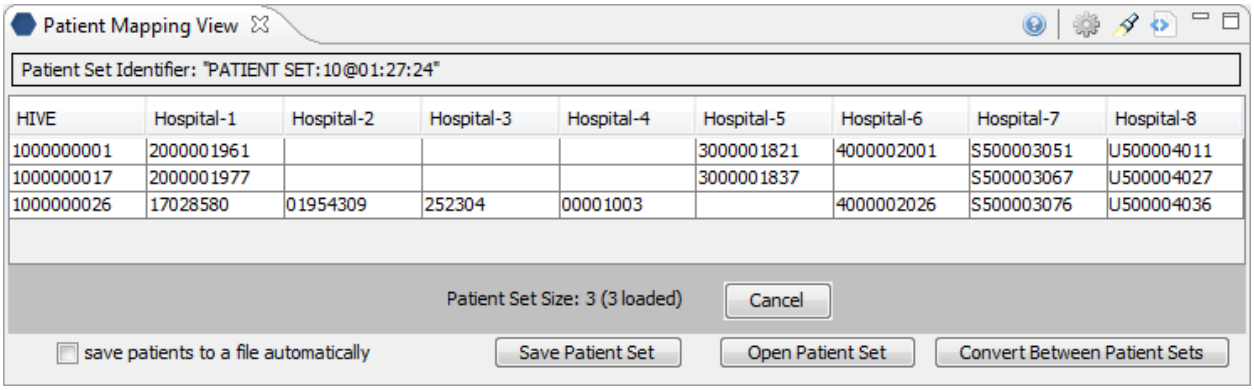
The following section outlines some of the basic design requirements and new views in the i2b2 Workbench to support the new Identity Management (IM) cell.

2.1.1 Patient Mapping View

A new plug-in called *edu.harvard.i2b2.eclipse.plugins.patientMapping* has been provided with the source code for the i2b2 Workbench. The view for this plug-in is called *Patient Mapping* and is available in the i2b2 Workbench.

The Patient Mapping view allows project managers and administrators to see all site IDs for a set of patients or a single patient.

These patient IDs can be dragged from the Patient Mapping view and dropped into other i2b2 views such as the Query Tool and Workplace views. All site IDs are able to be dragged; it is not limited to the HIVE number.



Note

The patient IDs shown in the above screen print are from the i2b2 demo data and are strictly for demonstration purposes. It does NOT contain real patient information.

2.1.2 Admin Tool View

A new plug-in called *edu.harvard.i2b2.eclipse.plugins.adminTool* has been provided with the source code for the i2b2 Workbench. The view for this plug-in is called *Managers Tool* and is available in the i2b2 Workbench.

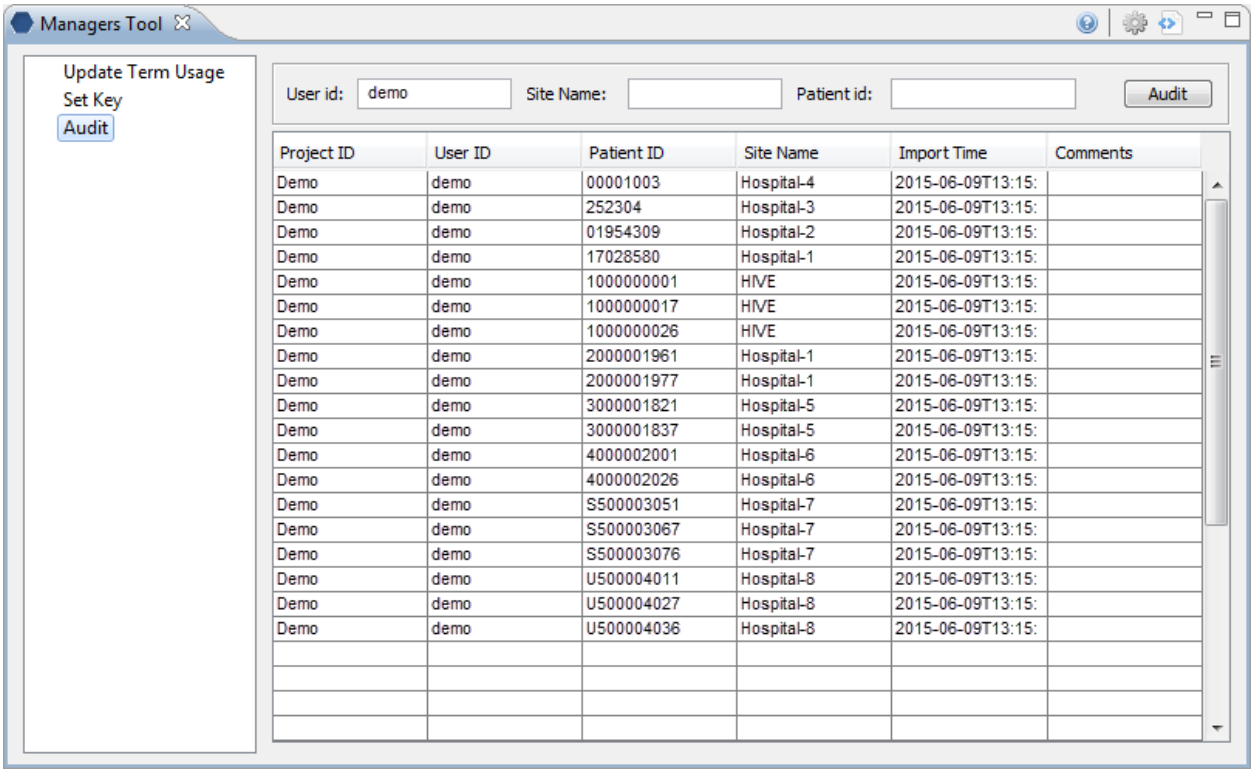
The Admin Tool view allows project managers and administrators to set or validate the AES encryption key for the IM cell. This same view also allows project managers and administrators to see the audit information stored in the **IM_AUDIT** table. This information can be viewed by user, by patient, or all audit records for the project.

2.1.2.1 View User Audit

A manager or administrator can view the audit trail for a particular user.

Example:

In the example shown below an audit report was run for the user “demo”. The IM cell returned all the patients that were accessed by the user “demo”.



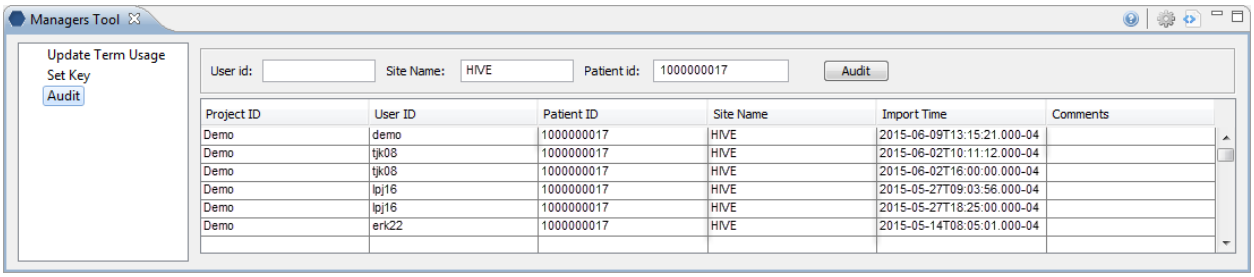
Project ID	User ID	Patient ID	Site Name	Import Time	Comments
Demo	demo	00001003	Hospital-4	2015-06-09T13:15:	
Demo	demo	252304	Hospital-3	2015-06-09T13:15:	
Demo	demo	01954309	Hospital-2	2015-06-09T13:15:	
Demo	demo	17028580	Hospital-1	2015-06-09T13:15:	
Demo	demo	1000000001	HIVE	2015-06-09T13:15:	
Demo	demo	1000000017	HIVE	2015-06-09T13:15:	
Demo	demo	1000000026	HIVE	2015-06-09T13:15:	
Demo	demo	2000001961	Hospital-1	2015-06-09T13:15:	
Demo	demo	2000001977	Hospital-1	2015-06-09T13:15:	
Demo	demo	3000001821	Hospital-5	2015-06-09T13:15:	
Demo	demo	3000001837	Hospital-5	2015-06-09T13:15:	
Demo	demo	4000002001	Hospital-6	2015-06-09T13:15:	
Demo	demo	4000002026	Hospital-6	2015-06-09T13:15:	
Demo	demo	S500003051	Hospital-7	2015-06-09T13:15:	
Demo	demo	S500003067	Hospital-7	2015-06-09T13:15:	
Demo	demo	S500003076	Hospital-7	2015-06-09T13:15:	
Demo	demo	U500004011	Hospital-8	2015-06-09T13:15:	
Demo	demo	U500004027	Hospital-8	2015-06-09T13:15:	
Demo	demo	U500004036	Hospital-8	2015-06-09T13:15:	

2.1.2.2 View Patient Audit

A manager or administrator can view the audit trail for a particular patient.

Example:

In the example shown below an audit report was run for the patient whose ID is 1000000017. The data returned from the IM cell shows all the users who accessed this patient.

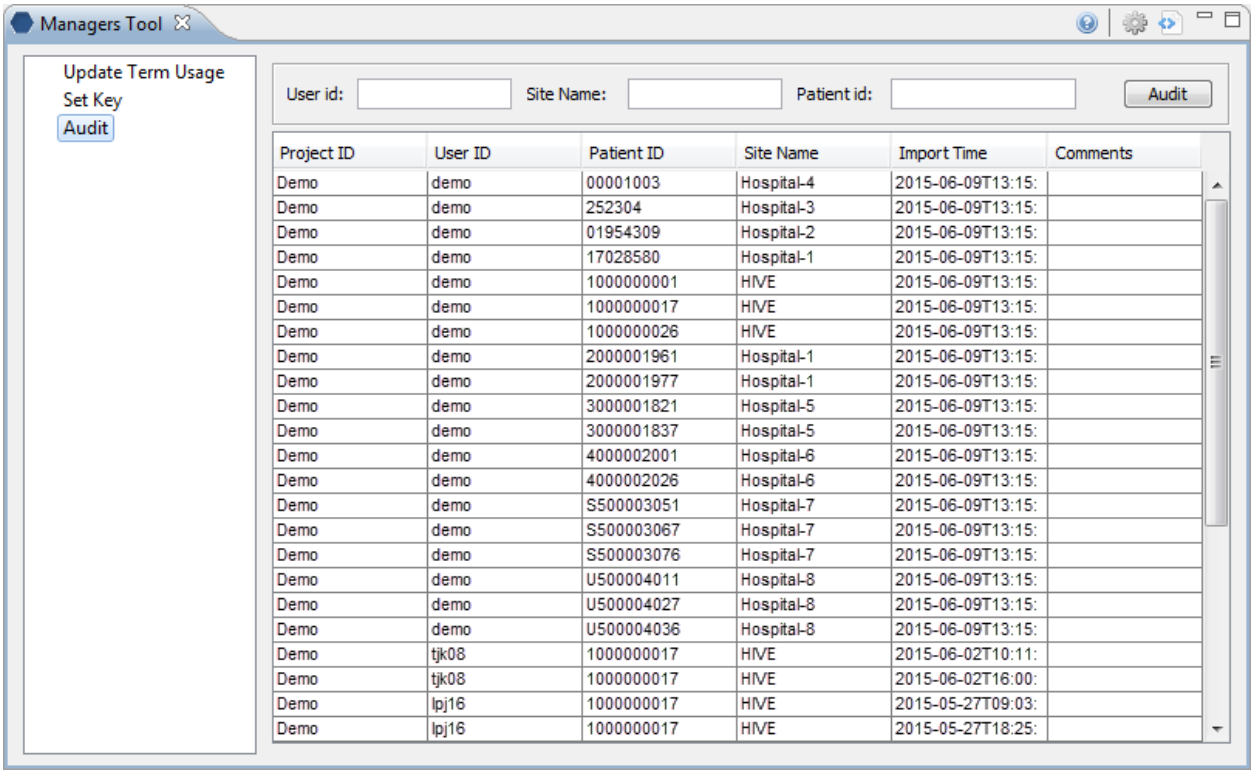


2.1.2.3 View All Audit Records for the Project

A manager or administrator can view the audit trail for all users and patients in the project.

Example:

In the example shown below the data returned for the IM includes all the audit records for the entire project.



Project ID	User ID	Patient ID	Site Name	Import Time	Comments
Demo	demo	00001003	Hospital-4	2015-06-09T13:15:	
Demo	demo	252304	Hospital-3	2015-06-09T13:15:	
Demo	demo	01954309	Hospital-2	2015-06-09T13:15:	
Demo	demo	17028580	Hospital-1	2015-06-09T13:15:	
Demo	demo	1000000001	HIVE	2015-06-09T13:15:	
Demo	demo	1000000017	HIVE	2015-06-09T13:15:	
Demo	demo	1000000026	HIVE	2015-06-09T13:15:	
Demo	demo	2000001961	Hospital-1	2015-06-09T13:15:	
Demo	demo	2000001977	Hospital-1	2015-06-09T13:15:	
Demo	demo	3000001821	Hospital-5	2015-06-09T13:15:	
Demo	demo	3000001837	Hospital-5	2015-06-09T13:15:	
Demo	demo	4000002001	Hospital-6	2015-06-09T13:15:	
Demo	demo	4000002026	Hospital-6	2015-06-09T13:15:	
Demo	demo	S500003051	Hospital-7	2015-06-09T13:15:	
Demo	demo	S500003067	Hospital-7	2015-06-09T13:15:	
Demo	demo	S500003076	Hospital-7	2015-06-09T13:15:	
Demo	demo	U500004011	Hospital-8	2015-06-09T13:15:	
Demo	demo	U500004027	Hospital-8	2015-06-09T13:15:	
Demo	demo	U500004036	Hospital-8	2015-06-09T13:15:	
Demo	tjk08	1000000017	HIVE	2015-06-02T10:11:	
Demo	tjk08	1000000017	HIVE	2015-06-02T16:00:	
Demo	lpj16	1000000017	HIVE	2015-05-27T09:03:	
Demo	lpj16	1000000017	HIVE	2015-05-27T18:25:	

2.2 Table Usage

There are several tables in the Identity Management (IM) Cell that have a unique role in the process of storing information about sites, projects and patients. This section outlines the various processes of accessing and using the IM tables.

Additional information about each table can be found in the *Identity Management Tables* section.

2.3 Security

The Identity Management (IM) Cell contains Protected Health Information (PHI), which can be used to identify patients. The HIPAA privacy rules state;

> **Audit Controls.** A covered entity must implement hardware, software, and/or procedural mechanisms to record and examine access and other activity in information systems that contain or use e-PHI.[1]

In order to comply with this regulation the IM Cell contains an Audit table that stores information about patient access. For more information about this table please see the section called *Audit Table* in the *Security Tables* section of this document.

3 IDENTITY MANAGEMENT TABLES

In the Identity Management (IM) cell there are several tables designed to store information from various sources. The information stored in these tables is listed below.

1. Projects associated with the different sites.
2. Patients who are included in a project.
3. Patient demographic information.
4. Mapping of the patient's site MRN and their i2b2 patient number.
5. Audit trail for patient access via the IM cell.

3.1 General Information

All the IM tables have the following five technically-oriented or administrative columns.

Column Name	Datatype	Allow Nulls	Column Definition
UPDATE_DATE	DATETIME	Y	Date the row was update by the source system The date is obtained from the source system
DOWNLOAD_DATE	DATETIME	Y	Date the data was downloaded from the source system
IMPORT_DATE	DATETIME	Y	Date the data was imported into the IM table
SOURCESYSTEM_CD	VARCHAR(50)	Y	A coded value for the data source system
UPLOAD_ID	INT	Y	A numeric id given to the upload

3.2 Project Tables

3.2.1 Project Sites Table

The **IM_PROJECT_SITES** table contains information that links a project and local site.

3.2.1.1 Requirements for Project Sites Table

The IM_PROJECT_SITES table has two **REQUIRED columns**.

1. PROJECT_ID
    - A unique id for the project.
    - This column is part of the primary key.

2. LCL_SITE
    - The local site that is associated to this project.
    - This column is part of the primary key.

3.2.1.2 Column Definitions

IM_PROJECT_SITES				
Key	**Column Name**	**Datatype**	**Allow Nulls**	**Column Definition**
PK	**PROJECT_ID**	VARCHAR(50)	N	A unique ID for the project. This column is equivalent to the **PROJECT_ID** column in the *PM_PROJECT_DATA* table in the i2b2 pmdata.
PK	**LCL_SITE**	VARCHAR(50)	N	The local site (data source) that is included in the project defined at PROJECT_ID. Example: MGH, BWH This column is equivalent to the **PATIENT_IDE_SOURCE** column in the *PATIENT_MAPPING* and *ENCOUNTER_MAPPING* tables in the i2b2 datamart.
	PROJECT_STATUS	VARCHAR(50)	Y	The status of the project. A = Active I = Inactive

	UPDATE_DATE	DATETIME	Y	As defined in the above section ("*General Information*")
	DOWNLOAD_DATE	DATETIME	Y	As defined in the above section ("*General Information*")
	IMPORT_DATE	DATETIME	Y	As defined in the above section ("*General Information*")
	SOURCESYSTEM_CD	VARCHAR(50)	Y	As defined in the above section ("*General Information*")
	UPLOAD_ID	INT	Y	As defined in the above section ("*General Information*")

3.2.2 Project Patients Table

The **IM_PROJECT_PATIENTS** table stores the patients that are part of a project.

3.2.2.1 Requirements for Project Patients Table

The IM_PROJECT_PATIENTS table has two **REQUIRED columns**.

1. PROJECT_ID
   - A unique id for the project.
   - This column is part of the primary key.

2. GLOBAL_ID
   - An internal ID for the patient.
   - This column is part of the primary key.

3.2.2.2 Column Definitions

IM_PROJECT_PATIENTS				
Key	**Column Name**	**Datatype**	**Allow Nulls**	**Column Definition**
PK	**PROJECT_ID**	VARCHAR(50)	N	A unique ID for the project. This column is equivalent to the **PROJECT_ID** column in the *PM_PROJECT_DATA* table in the i2b2 pmdata.
PK	**GLOBAL_ID**	VARCHAR(200)	N	This is a unique identifier for the patient. It is equivalent to a **Master Patient Index** (MPI) id. Note: some sites may refer to it as the Enterprise Master Patient Index (EMPI).
	PATIENT_PROJECT_STATUS	VARCHAR(50)	Y	The status of the patient in this project. A = Active I = Inactive
	UPDATE_DATE	DATETIME	Y	As defined in the above section ("*General Information*")
	DOWNLOAD_DATE	DATETIME	Y	As defined in the above section ("*General Information*")
	IMPORT_DATE	DATETIME	Y	As defined in the above section ("*General Information*")
	SOURCESYSTEM_CD	VARCHAR(50)	Y	As defined in the above section ("*General Information*")
	UPLOAD_ID	INT	Y	As defined in the above section ("*General Information*")

3.3 Mapping Tables

3.3.1 Demographics Table

The **IM_MPI_DEMOGRAPHICS** table contains general demographic information for the patients.

3.3.1.1 Requirements for Demographics Table

The IM_MPI_DEMOGRAPHICS table has one **REQUIRED column**.

1. GLOBAL_ID
    - An internal ID for the patient.
    - This column is part of the primary key.

3.3.1.2 Column Definitions

IM_MPI_DEMOGRAPHICS				
Key	**Column Name**	**Datatype**	**Allow Nulls**	**Column Definition**
PK	**GLOBAL_ID**	VARCHAR(200)	N	This is a unique identifier for the patient. It is equivalent to a **Master Patient Index** (MPI) id. Note: some sites may refer to it as the Enterprise Master Patient Index (EMPI).
	GLOBAL_STATUS	VARCHAR(50)	Y	The status of the ID for this patient. A = Active I = Inactive M = Merged
	DEMOGRAPHICS	VARCHAR(400)	Y	This is an optional column.
	UPDATE_DATE	DATETIME	Y	As defined in the above section ("*General Information*")
	DOWNLOAD_DATE	DATETIME	Y	As defined in the above section ("*General Information*")
	IMPORT_DATE	DATETIME	Y	As defined in the above section ("*General Information*")
	SOURCESYSTEM_CD	VARCHAR(50)	Y	As defined in the above section ("*General Information*")
	UPLOAD_ID	INT	Y	As defined in the above section ("*General Information*")

3.3.2 MPI Mapping Table

The **IM_MPI_MAPPING** table maps the patient's *i2b2 number* and their *local ID* from the source system.

3.3.2.1 Requirements for MPI Mapping Table

The IM_MPI_MAPPING table has three **REQUIRED columns**.

1. GLOBAL_ID
   - An internal ID for the patient.
   - This column is part of the primary key.

2. LCL_SITE
   - The local site.
   - This column is part of the primary key.

3. LCL_ID
   - The patient's local id (site MRN).
   - This column is part of the primary key.

3.3.2.2 Column Definitions

IM_MPI_MAPPING				
Key	**Column Name**	**Datatype**	**Allow Nulls**	**Column Definition**
PK	**GLOBAL_ID**	VARCHAR(200)	N	This is a unique identifier for the patient. It is equivalent to a **Master Patient Index** (MPI)

				id. Note: some sites may refer to it as the Enterprise Master Patient Index (EMPI).
PK	**LCL_SITE**	VARCHAR(50)	N	The local site (data source). Example: MGH, BWH This column is equivalent to the **PATIENT_IDE_SOURCE** column in the *PATIENT_MAPPING* and ENCOUNTER_*MAPPING* tables in the i2b2 Datamart.
PK	**LCL_ID**	VARCHAR(200)	N	The local ID is the patient's MRN (Medical Record Number). The MRN is a unique identifier at a site (institution) that represents a patient. This column is equivalent to the **PATIENT_IDE** column in the *PATIENT_MAPPING* and *ENCOUNTER_MAPPING* tables in the i2b2 Datamart.
	LCL_STATUS	VARCHAR(50)	Y	The status of this local ID in the source system. A = Active I = Inactive
	UPDATE_DATE	DATETIME	Y	As defined in the above section ("*General Information*")
	DOWNLOAD_DATE	DATETIME	Y	As defined in the above section ("*General Information*")
	IMPORT_DATE	DATETIME	Y	As defined in the above section ("*General Information*")
	SOURCESYSTEM_CD	VARCHAR(50)	Y	As defined in the above section ("*General Information*")
	UPLOAD_ID	INT	Y	As defined in the above section ("*General Information*")

3.4 Security Tables

3.4.1 Audit Table

As stated previously in the *Security* section of the *Requirements*, the IM cell contains PHI which can be used to identify patients. The audit table will store the following information when a patient is accessed via the IM cell.

1. The date the information was accessed.

2. The site that requested access to the patient.
3. The ID of the patient that was accessed.
4. The user who accessed the patient.
5. The project that the user was logged into when the patient was accessed.

3.4.1.1 Requirements for the Audit Table

All of the columns in the IM_AUDIT table are required. The only exception is the COMMENTS column.

3.4.1.2 Column Definitions

IM_AUDIT				
Key	**Column Name**	**Datatype**	**Allow Nulls**	**Column Definition**
PK	**QUERY_DATE**	DATETIME	N	The date the request was received from the client and the response message was sent from the IM cell.
	LCL_SITE	VARCHAR(50)	N	The site that requested / accessed the information.
	LCL_ID	VARCHAR(200)	N	The local ID for the patient that was accessed.
	USER_ID	VARCHAR(50)	N	The unique ID for the user that accessed the patient. This ID is equivalent to the **USER_ID** in the *PM_USER_DATA* table in the PM Cell.
	PROJECT_ID	VARCHAR(50)	N	The project ID that the user was logged into when accessing the patient. This ID is equivalent to the **PROJECT_ID** in the *PM_PROJECT_DATA* table in the PM Cell.
	COMMENTS	TEXT	N	

3.5 Joining Columns

All of the IM tables can be linked together using SQL joins to obtain more data. The IM tables can also be linked to the PATIENT_MAPPING table in the crcdata and the PM_PROJECT_DATA table in the pmdata.

The following are some examples of columns that can be used to join IM Tables and the i2b2 tables

IM_MPI_DEMOGRAPHICS		
GLOBAL_ID in IM_MPI_DEMOGRAPHICS	can be joined to	GLOBAL_ID in IM_MPI_MAPPING
GLOBAL_ID in IM_MPI_DEMOGRAPHICS	can be joined to	GLOBAL_ID in IM_PROJECT_PATIENT

IM_MPI_MAPPING		
LCL_ID in IM_MPI_MAPPING	can be joined to	PATIENT_IDE in PATIENT_MAPPING
LCL_SITE in IM_MPI_MAPPING	can be joined to	PATIENT_IDE_SOURCE in PATIENT_MAPPING
LCL_STATUS in IM_MPI_MAPPING	can be joined to	PATIENT_IDE_STATUS in PATIENT_MAPPING

IM_PROJECT_SITES		
PROJECT_ID in IM_PROJECT_SITES	can be joined to	PROJECT_ID in IM_PROJECT_PATIENTS
PROJECT_ID in IM_PROJECT_SITES	can be joined to	PROJECT_ID in PM_PROJECT_DATA
LCL_SITE in IM_PROJECT_SITES	Can be joined to	LCL_SITE in IM_MPI_MAPPING

IM_PROJECT_PATIENTS		
PROJECT_ID in IM_PROJECT_PATIENTS	can be joined to	PROJECT_ID in PM_PROJECT_DATA

4 OBJECTS

4.1 Patient Data Object

The Patient Data Object (PDO) is the XML representation of patient data. This data corresponds to the values in the star schema tables in the i2b2 database.

The PDO object is maintained by the CRC cell. Additional information regarding the PDO can be found in the *CRC Design Document*.

4.2 IM Data Object

The IM Data Object (IMDO) holds patient and site information and performs auditing.

5 PATIENT MAPPING SCENARIOS

A patient may have more than one identifier in different source systems and will be given a single unique global identifier in the IM tables. All of these identifiers are grouped together in the **IM_MPI_MAPPING** table in the database.

The IM Mapping tables link the values used in the IM database to their counterparts in the source systems from which the identifiers came. The CRC PATIENT_MAPPING table links the ids in the IM_MPI_MAPPING table to their i2b2 PATIENT_NUM.

In the IM_MPI_MAPPING table the patient's local id is stored in the LCL_ID column. This number can be stored as either encrypted or unencrypted. Encrypted identifiers are indicated by appending '_e' to the name of the source system that is stored in the LCL_SITE column. For example, if the identifier is an encrypted number from Massachusetts General Hospital, the LCL_SITE will be 'MGH_e'.

Below is a generic <pid_set> from the XML Patient Data Object (PDO).

```
<pid_set>
    <pid>
        <patient_id source="source">value</patient_id>
        <patient_map_id source="source" status="A">value</patient_map_id>
        <patient_map_id source="source" status="A">value</patient_map_id >
        …
    </pid>
</pid_set>
```

5.1 Self-Mapping

Self-mapping in the IM cell is similar to the self-mapping that occurs in the CRC. Each patient will have one record in which their id is mapped to itself.

6 GLOSSARY

6.1 General Terms

The following table contains terms that are used throughout this document.

Term	Definition
EMPI	Enterprise Master Patient Index
IM	Identity Management
MPI	Master Patient Index
MRN	Medical Record Number

END NOTES

Audit Controls. A covered entity must implement hardware, software, and/or procedural mechanisms to record and examine access and other activity in information systems that contain or use e-PHI.[25]

[25] 45 C.F.R. § 164.312(b).