CROSS-NETWORK DIRECTORY SERVICE PROJECT

DESIGN AND TECHNICAL DOCUMENTATION

Prepared by the Sentinel Operations Center
January 31, 2018
Version: 1.0

The Sentinel System is sponsored by the U.S. Food and Drug Administration (FDA) to proactively monitor the safety of FDA-regulated medical products and complements other existing FDA safety surveillance capabilities. The Sentinel System is one piece of FDA's Sentinel Initiative, a long-term, multi-faceted effort to develop a national electronic system. Sentinel Collaborators include Data and Academic Partners that provide access to healthcare data and ongoing scientific, technical, methodological, and organizational expertise. The Sentinel Coordinating Center is funded by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223201400030I. This project was funded by the FDA through HHS Mini-Sentinel contract number HHSF223200910006I.

Table of Contents

Modification History

Version	Date	Modification	By
1.0	01/31/2018	• Initial version	Sentinel Operations Center

I. BACKGROUND

The growing adoption of distributed health data networks to facilitate large-scale evidence generation studies (e.g., comparative safety and effectiveness), as well as other public health activities, provides an opportunity to leverage those investments to create a national resource that enables a true Learning Health System. FDA, PCORI, NIH, ONC, CDC and others are supporting various forms of distributed health data networks. Together, these networking infrastructure investments can be integrated to support needs across funding agencies and the broader public health community.

The PopMedNet™ (PMN) software application currently enables creation, operation, and governance of distributed health data networks. It supports distributed within-network querying for Sentinel, PCORnet, MDPHnet, HCSRN, HMO Cancer Research Network, and NIH Health Care Systems Research Collaboratory.

The Cross-Network Directory Service (CNDS) extends PMN's existing functionality to enable cross-network discovery of potential collaborators and data sources and querying of those sources while enforcing governance rules.

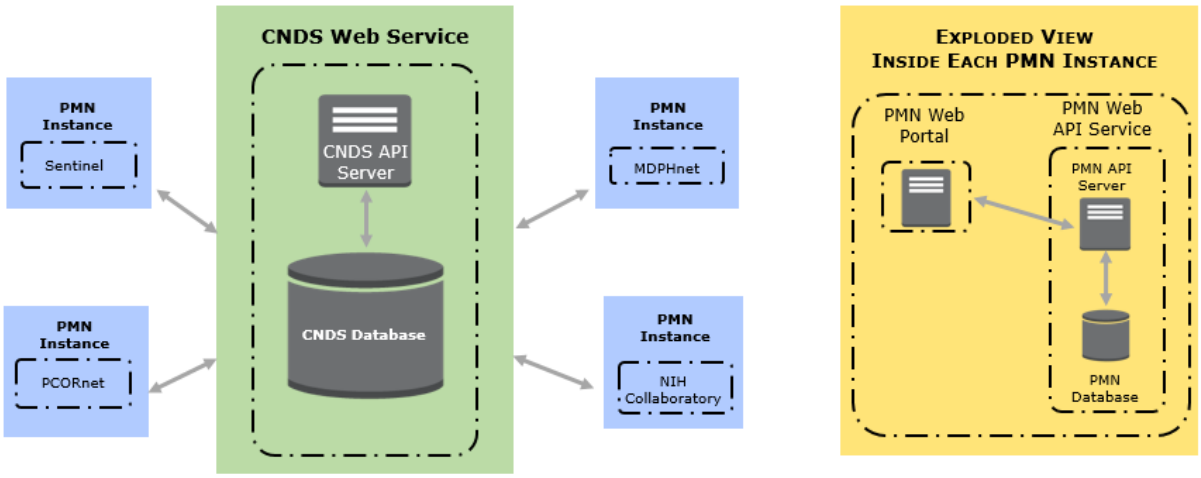
This report is a final deliverable and updates earlier technical documentation versions. **Section II** of this report is a brief architectural overview intended as a refresher about the objectives and approach of this project. The overview describes the core concepts of system entities and metadata and each of the four user-facing functions provided by CNDS: Registration, Discovery, Communication, and Governance. It also describes the administration of the system. **Section III** documents the technical requirements and results of software quality assurance testing. The accompanying deliverable—**ASPE CNDS User Documentation**—is aimed at a more general audience.

II. ARCHITECTURAL OVERVIEW

To minimize the impact on existing networks, CNDS is built gently on top of PMN. CNDS leverages the existing PMN application. This is achieved by implementing CNDS as a set of services that can be invoked by PMN instances, rather than by wholesale modification of PMN. In slightly more technical terms, CNDS provides a standard set of functions that PMN can call upon through application programming interfaces (APIs).

Figure 1 shows the high-level architecture of CNDS in relationship to PMN. Each PMN network has its own instance of PMN with the same functionality. The orange box shows an exploded view of what is in each PMN instance (an instance is a stand-alone installation of the application on a server). PMN networks can easily be made part of CNDS by loading the PMN user, organization, and data source information into the CNDS database and configuring the network's instance to have access to the CNDS API.

Figure 1. PMN - CNDS Integrated Design



Through its APIs, CNDS offers functionality to:

- Process registration requests
- Capture metadata describing users, organizations, registries/research data sets, and queryable data sources
- Enable users to search the metadata in order to explore characteristics of electronic healthcare data sources and identify potential collaborators across and outside of existing PMN distributed data networks
- Route requests and responses across networks

A. LOCATION OF CNDS PILOT

The pilot CNDS implementation is currently hosted in a test environment. Two mock websites representing the Sentinel and the PCORnet networks participating in CNDS are accessible from the following URLS and represent how CNDS would work in production:

- **Sentinel CNDS site:** https://cndsedge-sentinel.popmednet.org
- **PCORnet CNDS site:** https://cndsedge-pcornet.popmednet.org

B. SYSTEM CONCEPTS

Underlying the CNDS are two basic concepts. "System entities" are the participants in the system and are discoverable through it. "Metadata" refers to standardized data elements about entities that make it possible to discover and identify organizations and data sources.

1. System Entities

The following entities, which exist in PMN, also exist in CNDS:

- Users (currently, CNDS members must be PMN members)
- Organizations
- Data Sources (e.g., PMN DataMart[1], cancer registry, clinical research database)

2. Metadata

CNDS is powered by metadata—data elements that describe organizations and data sources. CNDS provides flexible management, storage, and retrieval of metadata about organizations and data sources. Metadata is also used to determine what organization and data source metadata is visible to whom. The **Physical Metadata Model** was designed to enable changes to metadata elements without software redesign or programming. All data elements are stored in a single column with a second column capturing hierarchical relationships between elements. CNDS **Manage Metadata** functionality allows system administrators to quickly add, delete, or modify metadata elements. For example, if CNDS stakeholders decide to start capturing information about laboratory result data not previous captured, new fields can be added to the database and available in CNDS Registration and Discovery on demand.

C. KEY FUNCTIONAL COMPONENTS

CNDS provides functionality for Registration, Discovery, Communication, Governance, and Administration as described below.

1. Registration

The Registration component of CNDS will enable users to access the system and enter information about themselves, their organizations and their electronic healthcare data resources. PMN is enhanced with a front-end user interface (UI) that facilitates secure user access and data entry into the CNDS database. The UI is designed to capture this information in an easily retrievable manner.

CNDS registration extends PMN beyond a platform for single networks by enabling users to allow themselves and their data sources to be discovered and their metadata queried by users outside their networks.

CNDS provides user interfaces for registration and entry and update of metadata and governance rules about the visibility of those metadata.

[1]A DataMart is the piece of software that receives a request from a PMN network and translates it into a query that can be run against local data.

2. Discovery (aka Search)

The Discovery component enables users to search the metadata, entered as part of registration, to find contact information for potential new collaborators and data sources. Like the CNDS **Physical Metadata Model**, it is designed flexibly so that the application does not require re-programming as the metadata change. That is, the list of elements that can be searched is automatically generated from the metadata stored in the database. As metadata fields are added, changed, or deleted using **Manage Metadata** functionality, the metadata available in Discovery change the next time the user navigates to or refreshes the screen. The result sets returned from a search will be constrained by the **Visibility** level set by the metadata owner. For example, if an owner of data source X indicates inpatient diagnosis data is available but they will not share it out-of-network, out-of-network user Y will not discover that data source X contains those data.

3. Communication

PMN currently has functionality for creating, distributing, responding to, and viewing a variety of request types, and it sends related email notifications within a single PMN network. CNDS extends this capability across networks by mapping common request types used by multiple networks.

Using CNDS, users, regardless of network affiliation, can send and receive requests according to the rules of the recipients. Users can export the result set from Discovery and, when searching for data sources that can be queried through PMN (i.e., DataMarts), the user will be able to distribute a data request. Recipients will receive email notification of the request and can decide whether to execute it and return the results. When results are returned the requester will receive email notification.

Due to the complexities of other request types and the differences in Common Data Models, this first version of CNDS provides functionality for sending file transfer requests only.

4. Governance

a) Visibility

The underpinning of CNDS governance is the ability to encode visibility rules in metadata (Registration) and enforce those rules (Discovery). Visibility rules identify "who" is authorized to see each organization and data source metadata element. Users can indicate metadata elements as being visible to:

- No one (i.e., just myself and CNDS system administrators)
- Registrants in my PMN network
- Registrants in any PMN network
- All CNDS Registrants

b) Access Controls

PMN provides an extensive set of access controls which are also available to CNDS. They control every aspect of use of the application, for example: adding, editing, deleting, and viewing users, organizations, and DataMarts; responding to, rejecting, and uploading results; managing security; running audit reports and permission to set other users permissions. Additional access controls implemented in CNDS are described in **Table 1**.

Table 1. CNDS Access Controls

Access Control	Description
Discovery	
Search CNDS	Governs whether the user sees the "search" menu item used to access CNDS search and therefore whether the user can access CNDS search functionality. No additional levels of governance are applied for accessing search. Users without this permission cannot see the "Search" option in the CNDS menu.
Communication	
Create CNDS Request	Governs the ability to create a request that will be sent to DataMarts in and out of network. Users who have this permission can create a request from the results of a Discovery search. Existing PMN permissions govern all other request creation functionality (e.g., edit, copy, and distribute requests).
Map Request Type	Governs the ability to associate a request type in one network with a request type in another network. Users without this permission cannot see the "Manage Request Type Mappings" option in the CNDS menu.
Administration	
Manage Metadata	Governs the ability to perform all functions related to metadata management including adding, editing, deleting domains, and assigning domains to organization and/or data sources. Users without this permission cannot see the "Manage Metadata" option in the CNDS menu.
Manage CNDS Access & Permissions	Governs the ability to set CNDS permissions for security groups and assign users to CNDS security groups. Users without this permission cannot see the "Permissions" option in the CNDS menu.
Create CNDS Security Group	Governs the ability to create a CNDS security group
Edit CNDS Security Group	Governs the ability to edit the description/name of a CNDS security group. (Note: It does not govern the ability to assign permissions to the security group. This is covered by the access control “Manage CNDS Access & Permissions”).
Delete CNDS Security Group	Governs the ability to delete a CNDS security group. Deleting is performed by clicking "remove" in associated row of the security group table. Deleting will remove the group from the CNDS database and all profiles to which it is assigned.

5. Administration

a) Manage Metadata

Users with sufficient rights can manage the metadata. They can add, edit, or delete metadata elements by selecting "CNDS – Manage Metadata" from the CNDS menu.
The available metadata types are text, whole number, true|false, reference, and Boolean group.[2] References can be single or multi-select. Most of the data types are conventional and self-explanatory except Boolean group, which both holds other data types and allows for the creation of hierarchy among metadata elements. This functionality allows for data elements to be organized, and thereby searched for hierarchically. An example of a hierarchy of metadata elements is:
Types of Data Collected—Inpatient Encounters—Inpatient Diagnosis Codes—ICD-9, ICD-10, SNOMED
Metadata fields can be associated with organizations and/or data sources. For example:

- "Willingness to accept data requests" would be associated with data sources, but not organizations
- "Clinical Trial Expertise" would be associated with organizations, but not data sources
- "Data Models" would be associated with both data sources and organizations

b) Manage Request Type Mappings

In PopMedNet request types are defined to express questions investigators wish to ask. Questions are sent to selected DataMarts via a chosen request type (e.g., file distribution). Request types are subject to local governance controls and security policies at both the network and project levels. A project is an entity within a network that allows for users and DataMarts to be grouped according to investigator questions, request types, security policies, and governance. For example, a group within a network that is working on obesity research can be set up as a "project" which includes a subset of the larger network's DataMarts and request types. One DataMart can be a part of multiple projects.

Traditionally in PMN, the combination of a project, request type, and DataMart is defined as a route. Requests can only be sent via routes to DataMarts within the same project. CNDS expands this by enabling questions to be sent to DataMarts across projects and networks (i.e., "external" routes). To accomplish this, a CNDS system administrator creates mappings that define allowed external routes. An external route is defined as a combination of a network, project, request type, and DataMart.

Since a request type in one network is defined independently from a request type in another, CNDS depends on the CNDS administrator to correctly identify the external route that can service a request type created in the network initiating the request. Discovery may return DataMarts that have and are willing to share the data of interest, but the necessary route must be in place for CNDS Communication to handle the request.

[2] Note that another metadata type called "Container" exists in Manage Metadata, but it is not fully implemented in this version of the software and should not be used.

III. REQUIREMENTS AND TESTING

CNDS software was developed collaboratively using the Agile software development method[3]. The Agile method is routinely used in the management of software development. Atlassian's JIRA software[4] was used to manage and track the development requirements, which were expressed in JIRA "issues". The functionality required for each of the five CNDS components is described in a set of issues. **Table 2** summarizes these issues, organized by CNDS component. Quality control and user acceptance testing of all CNDS components passed.

Table 2. Software Requirements

Issue Title	Issue ID	Requirement Summary
Registration and Metadata		
Registration and Landing Page for Users in CNDS	216	Users can register for an account in an instance of CNDS. A CNDS system administrator has the ability to approve the accounts.
De-Register PMN Users	221	A CNDS system administrator has the ability to deactivate or delete someone from the CNDS registry. "Deactivate" means the user cannot log in, but is still displayed in the User Profile section. "Delete" also deactivates the user, but also causes the user to not display in the User Profile section.
Update Metadata Values	244	Enable users to update their own metadata, as well as organization and data source metadata; this ability is governed by access controls and permissions. The user interface for metadata management must read the structure of the CNDS repository and programmatically generate a web page to display all relevant metadata elements.
Auto-Registration of New Users in CNDS	220	New users that register for an account in one of the two CNDS instances (PCORnet or Sentinel), are automatically registered in CNDS.
Manage Metadata Elements	185	Create the ability for a CNDS system administrator to add or modify the set of metadata elements in the metadata repository; create a user interface to access this capability; govern this capability by access controls.
Accommodate Hierarchical Check List Using the Domain Data Model	490	The CNDS metadata repository supports a multi-level hierarchical Boolean metadata points stored in the Domain tables. In parent-child relationships, the value of a metadata element is either "Yes we have it" or "No we do not have it." Enable those values to be set independently for a parent and all its children.
Discovery		
CNDS Discovery: Search Metadata for Data Sources and Organizations	621	Implement Discovery solely within CNDS. Search metadata to identify data sources, and their associated organizations, with

[3] See, for example, https://www.scrumalliance.org/

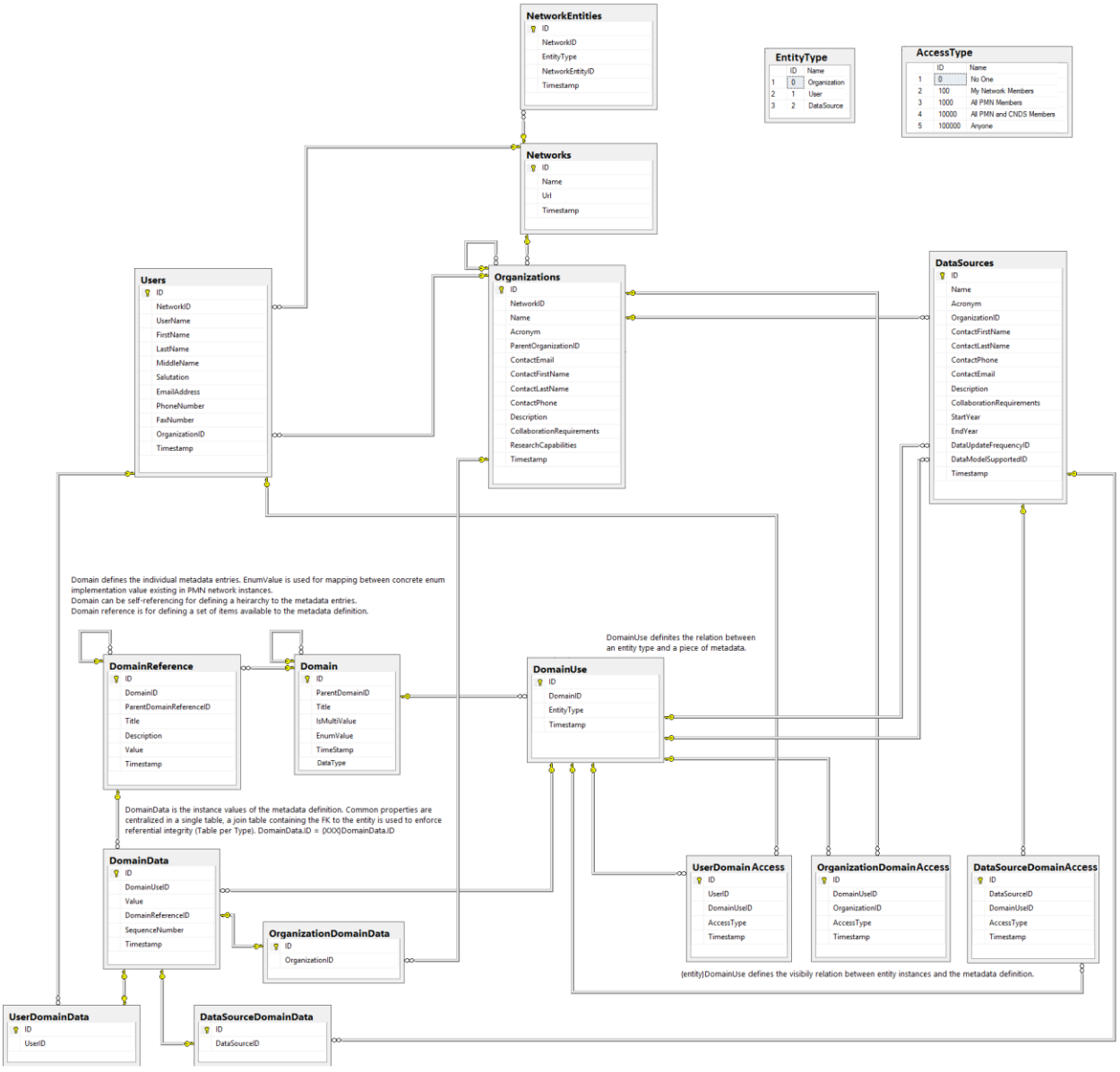
[4] Atlassian's JIRA: https://www.atlassian.com/software/jira

Issue Title	Issue ID	Requirement Summary
		specific provenance and/or data types or adhering to specific common data models.
Implement CNDS Search Data Source Function with Ability to Create a Request from Results	627/633	Implement UI, Kendo, Knockout, and API changes that would enable the results of a Data Source search to initiate a new request to an external network.
Communication		
Network Request Type Mappings	524	Enable a CNDS system administrator to maintain a table of correspondences between pairs of request types, one from each of two PopMedNet networks. This table of mappings will enable requests to be sent between PopMedNet networks.
Send Requests Cross-Network Pt. 1 & 2	206/525	Enable the ability of a user to directly create a new request following a Data Sources search (i.e. Discovery). The selected request type populates the target recipients with the compatible Data Sources from the search result set. Users can fill out request metadata and save the request. The PopMedNet portal must distinguish between data requests sent within the local PMN network and those sent to external PMN instances.
Receive and Process Cross-Network Requests	439	Data requests to be sent cross-network are processed within CNDS according to the mapping table of request types (see **Manage Request Type Mappings**). Enable the ability for CNDS to submit queries to cross-network recipients, as well as the ability for the recipients to receive and process the request.
Receive and Process Request Responses	443	Once recipient DataMarts have processed a request, the DataMart Client transmits responses within its own network as normal; responses destined for an external submitter are routed back through CNDS.
Governance		
Metadata Visibility	252/621	Provide a user interface so that each metadata element can be labeled according to its visibility to other parties. The web page enabling visibility setting must be programmatically generated from the metadata repository.
Access Control Lists for CNDS Functions	253	Implement access controls for functions specific to CNDS and not covered by standard PopMedNet access controls. Access controls are needed for Discovery, Communication, and Governance; no access controls are anticipated for Registration. The CNDS-Specific Controls are: • Search CNDS • Create CNDS Requests • Map Request Types • Manage Metadata • Manage CNDS Access • Manage CNDS Security Groups
Administration		

Issue Title	Issue ID	Requirement Summary
Activate PopMedNet Notifications for CNDS	659	Configure the SMTP server to send out notification emails from CNDS instances. The following notifications are activated: 1) DataMart administrator receives notification of submitted data request 2) Requester of data receives notification of responses from DataMart 3) Requester of data receives notification of request completion
Update CNDS to New Data Model	273	The CNDS instance is created as a copy of PopMedNet release 6.0. The **Physical Metadata Model** is more sophisticated and extensive than the original PMN metadata model. Hence the CNDS instance's metadata model must be upgraded to the new model.
PMN-CNDS Infrastructure Setup	309	Create the development, test, and edge servers needed for CNDS software development. Create URLs enabling access to these servers. Create a separate branch for CNDS in the PopMedNet source code management system.

IV. TECHNICAL APPENDICES

A. PHYSICAL METADATA MODEL



B. DATA DICTIONARY

In the **Physical Metadata Model** above, the user-friendly table and field names are displayed. The data dictionary below, maps the user-friendly table and field names to the actual names in the database. Table names are highlighted in light blue.

The tables in the data dictionary match the tables in the metadata model reading first top to bottom and then left to right, except the two lookup tables EntityType and AccessType.

Physical Model Name	Database Field or Table Name
NetworkEntities	**NTWRK_ENTY_T**
ID	ID
NetworkID	NTWRK_ID
EntityType	ENTY_TYP_CD
NetworkEntityID	NTWRK_ENTY_ID
Timestamp	CNDS_UPDT_TS
Networks	**NTWRK_T**
ID	ID
Name	NTWRK_NM
URL	URL_TXT
Timestamp	CNDS_UPDT_TS
Users	**USR_T**
ID	ID
NetworkID	NET_ID
UserName	USR_NM
FirstName	FRST_NM
LastName	LAST_NM
MiddleName	MID_NM
Salutation	PREFX_NM
EmailAddress	EMAIL_NM
PhoneNumber	PHONE_NBR
FaxNumber	FAX_NBR
OrganizationID	ORG_ID
Timestamp	CNDS_UPDT_TS
Organizations	**ORG_T**
ID	ID
NetworkID	NTWRK_ID
Name	ORG_NM
Acronym	ORG_ACRYN_TXT
ParentOrganizationID	PARNT_ORG_ID
ContactEmail	CONTC_EMAIL_NM
ContactFirstName	CONTC_FRST_NM
ContactLastName	CONTC_LAST_NM
ContactPhone	CONTC_PHONE_NBR
Description	ORG_DSC
CollaborationRequirements	COLLB_REQMT_TXT
ResearchCapabilities	RSRCH_CAP_TXT

Physical Model Name	Database Field or Table Name
Timestamp	CNDS_UPDT_TS
DataSources	**DATA_SRC_T**
ID	ID
Name	DATA_SRC_NM
Acronym	DATA_SRC_ACRYN_TXT
OrganizationID	ORG_ID
ContactFirstName	CONTC_FRST_NM
ContactLastName	CONTC_LAST_NM
ContactPhone	CONTC_PHONE_NBR
ContactEmail	CONTC_EMAIL_TXT
Description	DATA_SRC_DSC
CollaborationRequirements	COLLB_REQMT_TXT
StartYear	BEG_YEAR_NBR
EndYear	END_YEAR_NBR
DataUpdateFrequencyID	DATA_UPDT_FREQ_ID
DataModelSupportedID	DATA_MODEL_SUPP_ID
Timestamp	CNDS_UPDT_TS
DomainReference	**DOMN_REF**
ID	ID
DomainID	DOMN_ID
ParentDomainReferenceID	PARNT_DOMN_REF_ID
Title	DOMN_NM
Description	DOMN_REF_DSC
Value	DOMN_REF_CD
Timestamp	CNDS_UPDT_TS
Domain	**DOMN**
ID	ID
ParentDomainID	PARNT_DOMN_ID
Title	DOMN_NM
IsMultiValue	MULTI_VAL_IND
EnumValue	PMN_DOMN_NM
TimeStamp	CNDS_UPDT_TS
DataType	DATA_TYP_CD
DomainUse	**DOMN_USE**
ID	ID
DomainID	DOMN_ID
EntityType	ENTY_TYP_ID
Timestamp	CNDS_UPDT_TS
DomainData	**DOMN_DATA**
ID	ID
DomainUseID	DOMN_USE_ID
Value	DOMN_DATA_TXT
DomainReferenceID	DOMN_REF_ID
SequenceNumber	SEQ_NBR

Physical Model Name	Database Field or Table Name
Timestamp	CNDS_UPDT_TS
OrganizationDomainData	**ORG_DOMN_XREF**
ID	ID
OrganizationID	ORG_ID
UserDomainAccess	**USR_DOMN_ACCESS**
ID	ID
UserID	USR_ID
DomainUseID	DOMN_USE_ID
AccessType	ACCES_TYP_ID
Timestamp	CNDS_UPDT_TS
OrganizationDomainAccess	**ORG_DOMN_ACCESS**
ID	ID
DomainUseID	DOMN_USE_ID
OrganizationID	ORG_ID
AccessType	ACCES_TYP_ID
Timestamp	CNDS_UPDT_TS
DataSourceDomainAccess	**DATA_SRC_DOMN_ACCESS**
ID	ID
DataSourceID	DATA_SRC_ID
DomainUseID	DOMN_USE_ID
AccessType	ACCES_TYP_ID
Timestamp	CNDS_UPDT_TS
UserDomainData	**USER_DOMN_XREF**
ID	ID
UserID	USR_ID
DataSourceDomainData	**DATA_SRC_DOMN_XREF**
ID	ID
DataSourceID	DATA_SRC_ID
EntityType	**ENTY_TYP**
ID	ID
Name	ENTY_TYP_NM
AccessType	**ACCES_TYP**
ID	ID
Name	ACCES_TYP_NM