

Cross-Network Directory Service (CNDS) Enabling Meaningful Collaboration Across Organizations

Jessica Malenfant, MPH

Department of Population Medicine of Harvard Medical School & Harvard
Pilgrim Health Care Institute

Health Care Systems Research Network Annual Conference
April 12, 2018



Outline

- Why a Cross-Network Directory Service
- Design: Flexible and Extensible Architecture
- CNDS Pilot
- Key Features and Functionality
- What's Next

Outline

- **Why a Cross-Network Directory Service**
- Design: Flexible and Extensible Architecture
- CNDS Pilot
- Key Features and Functionality
- What's Next

Why Create a Cross-Network Directory Service (CNDS)?

- Distributed health data research and surveillance networks are proliferating



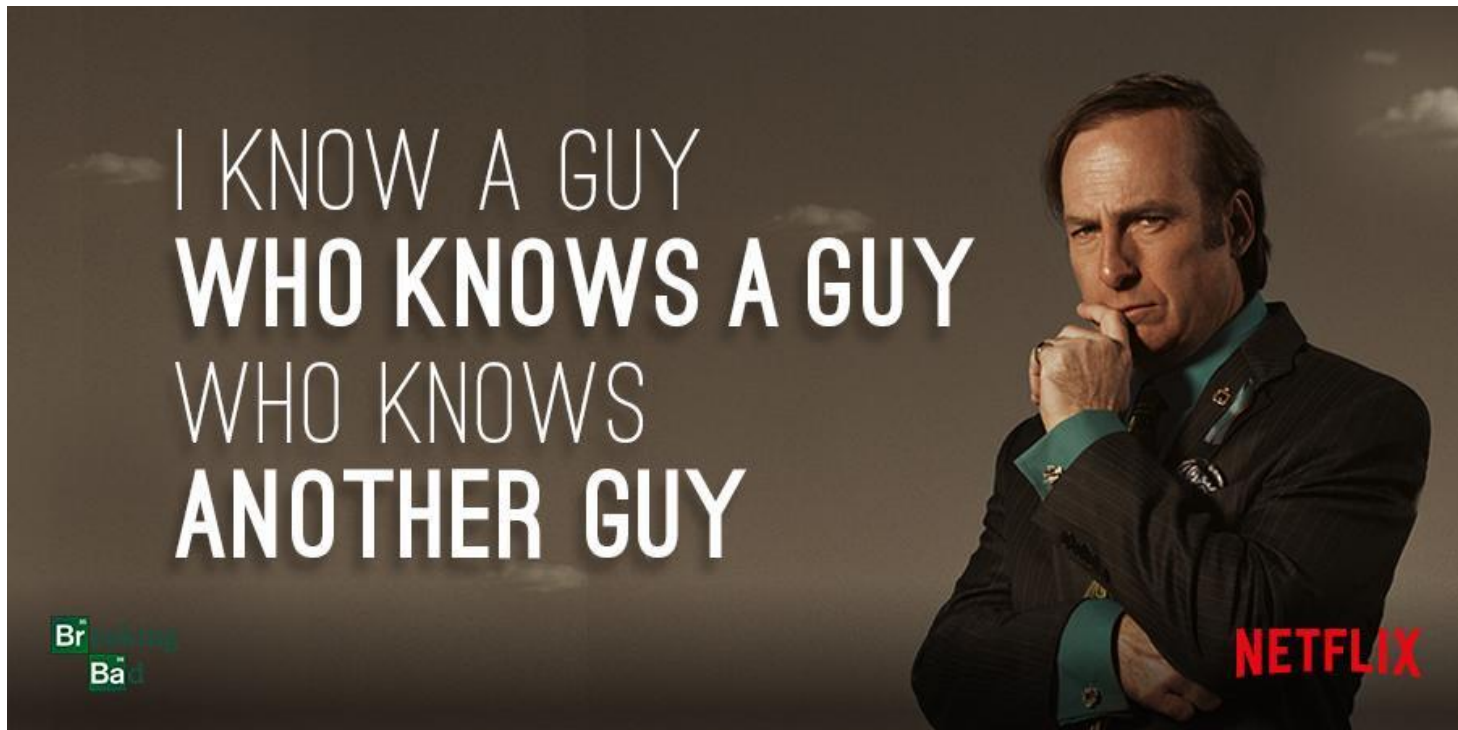
- It is challenging to find the right data resources and collaborators



CNDS Purpose

CNDS aims to facilitate collaboration across the networks by enabling users to . . .

Find Collaborators



Discover and Query Data Sources

My study requires inpatient medication and transfusion data for at least 200,000 inpatient visits in the past 3 years.

~

Where can I find data?



Outline

- Why a Cross-Network Directory Service
- **Design: Flexible and Extensible Architecture**
- CNDS Pilot
- Key Features and Functionality
- What's Next

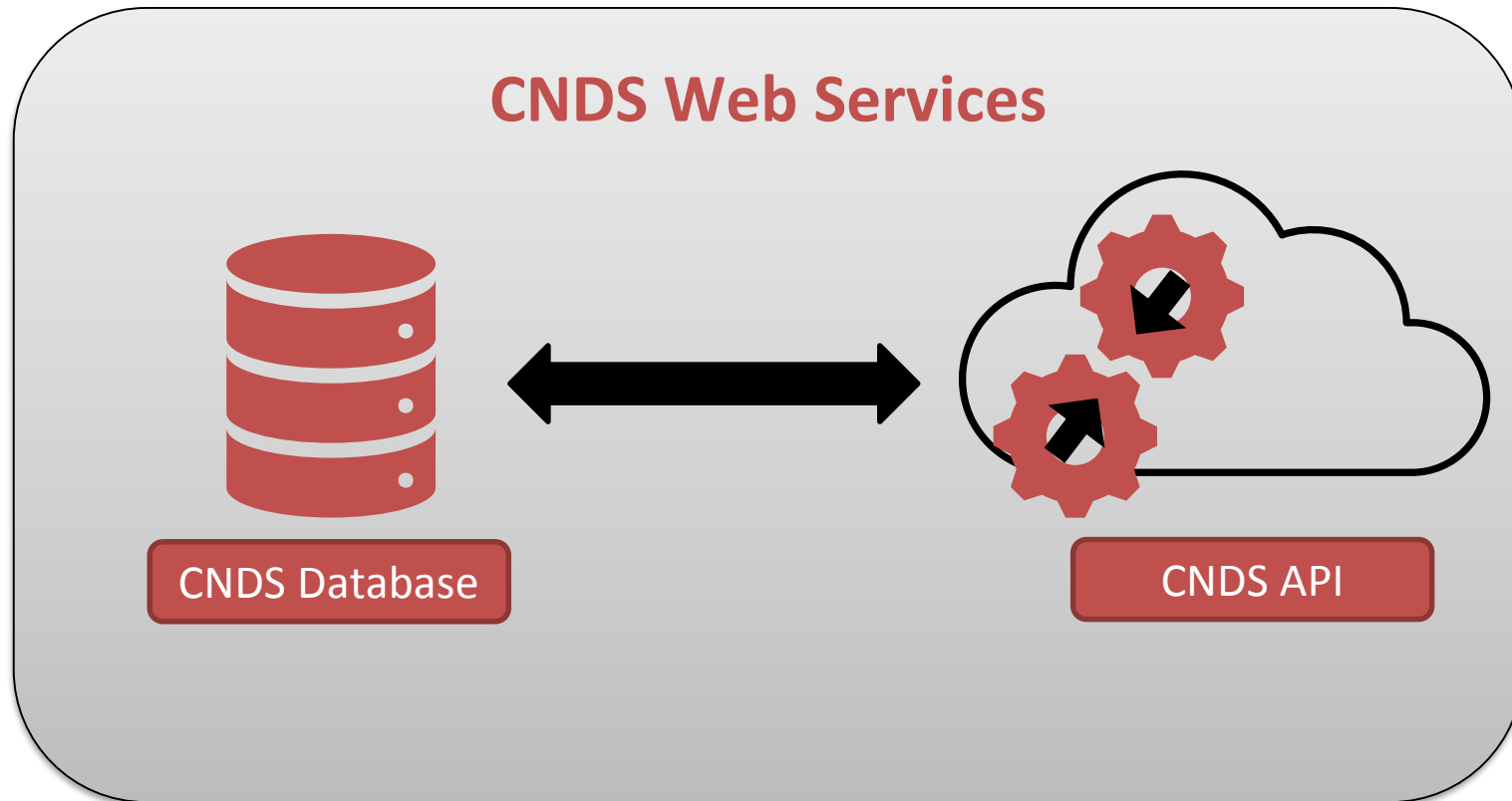
Design: Flexible and Extensible

- PopMedNet™ (PMN) is an open-source software platform used by Sentinel, PCORnet, HCSRN, and other health data networks



- CNDS, built lightly on PMN uses a modular design that minimizes the impact on existing networks
- The CNDS metadata data model enables changes to metadata elements without software redesign or programming

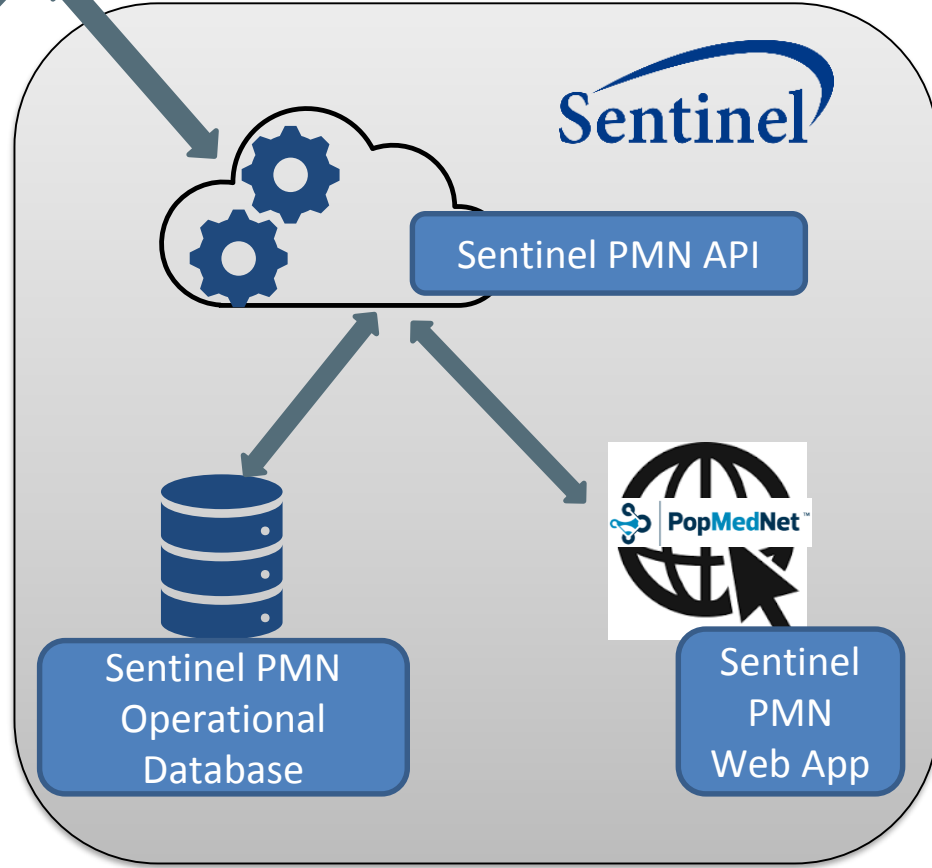
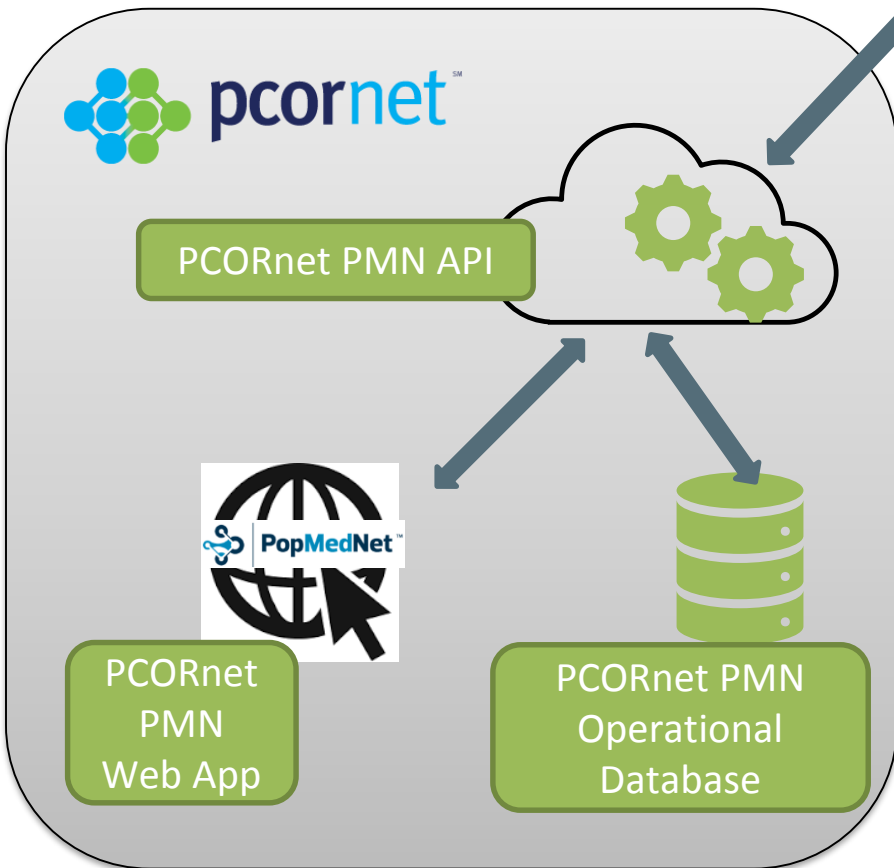
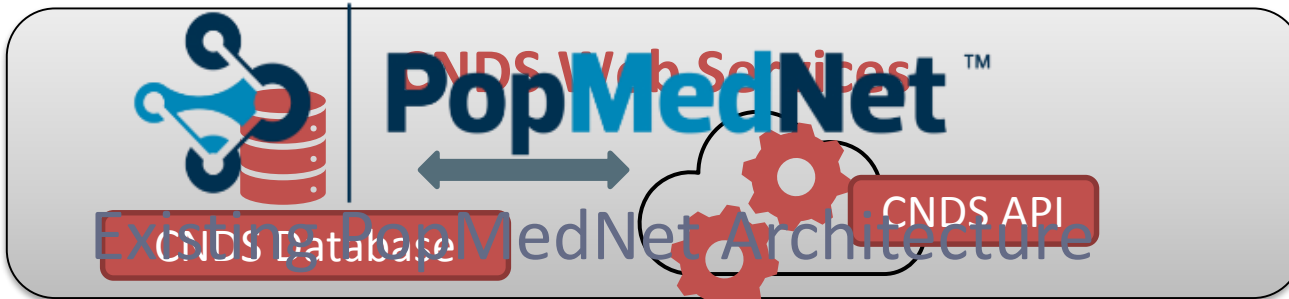
Software Overview: Architecture

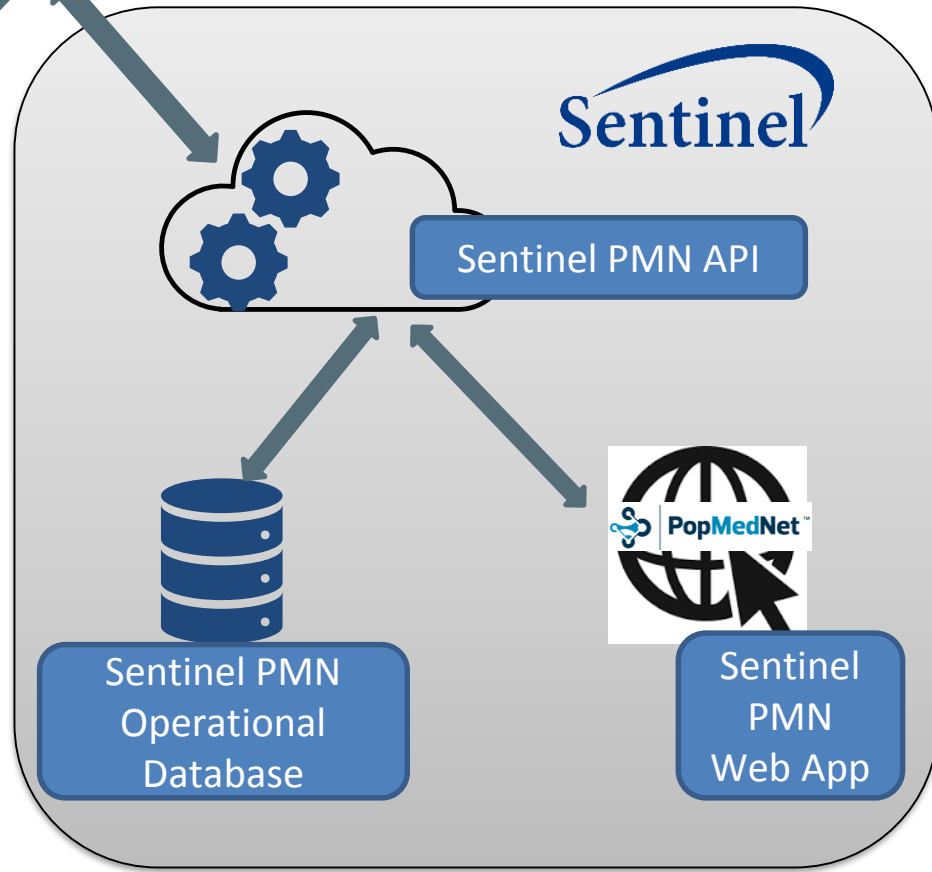
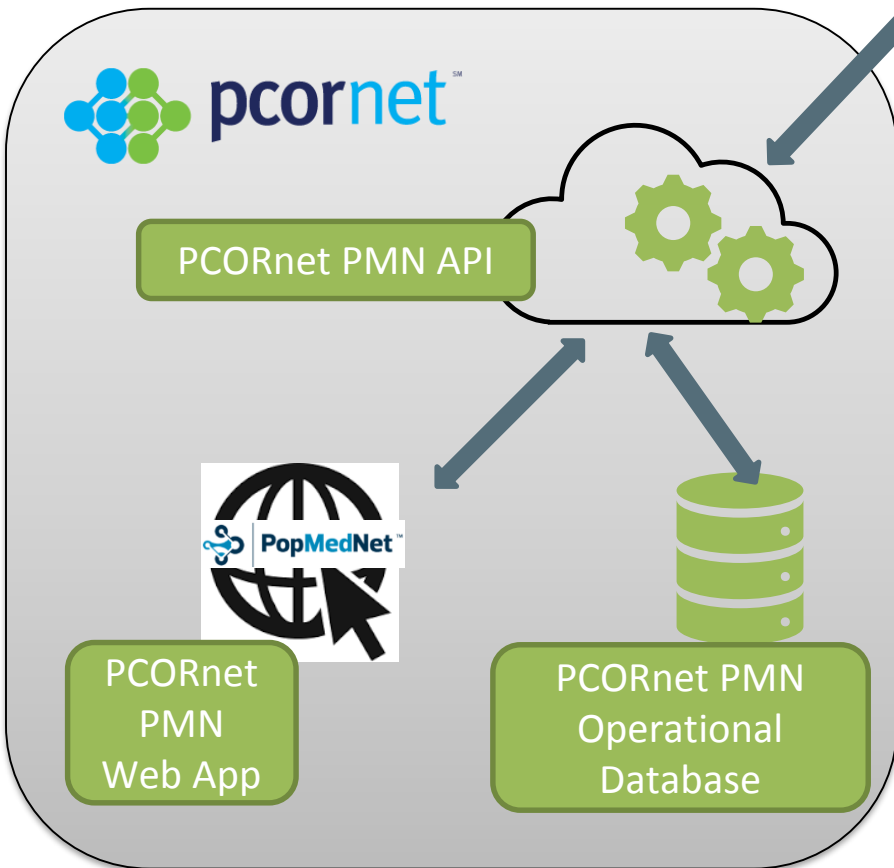
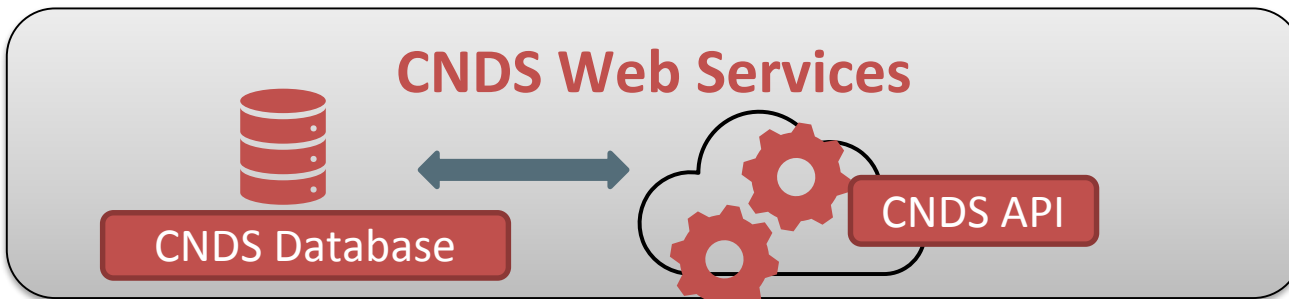


CNDS is a web service with a database and an Application Programming Interface (API)

Architecture

- APIs enable communication between web applications
- Implementing CNDS using API calls between PMN and CNDS:
 - Makes CNDS feel like part of PMN while insulating PMN and CNDS from each other
 - Enables changes to either system without affecting the other





CNDS Purpose

CNDS aims to facilitate collaboration across the networks by enabling users to find collaborators and discover and query data sources

Design: Flexible and Extensible Architecture

- CNDS connects distinct research networks
- The data model captures information about expertise, data sources, and technology resources
- Users can explore the data and connect to others

Outline

- Why a Cross-Network Directory Service
- Design: Flexible and Extensible Architecture
- **CNDS Pilot**
- Key Features and Functionality
- What's Next

CNDS Pilot

- Workgroup convened (representatives from Harvard Pilgrim Health Care Institute, FDA, a Sentinel data partner – Humana, a PCORnet data partner – University of Michigan, and the software developer)
- Larger stakeholder meetings held
- Create demonstration versions of PMN for Sentinel and PCORnet with new CNDS interfaces and functionality
- Populated user, organization, and data source information into CNDS database
- We did **not** focus on what to collect; instead we built a system that can collect nearly anything
- Tested, tested, tested

Outline

- Why a Cross-Network Directory Service
- Design: Flexible and Extensible Architecture
- CNDS Pilot
- **Key Features and Functionality**
- What's Next

Metadata

- At the core of CNDS is metadata
- Metadata about organizations and their data sources are captured in a “Profiles” screen

Cross Network Directory Service

Welcome, A Person

Home My Profile Resources Profiles CNDS Contact Us Log Off

Home / DataMarts / A Datamart

DataMart Name* **Acronym*** **Organization***

Adapter Supported
None

Data Domains [Installed Models](#)

Data Model

PCORnet CDM Sentinel CDM
 Other Internal model

Source of Data

EHR Patient-reported
 Claims Registry
 Biorepository Survey Data

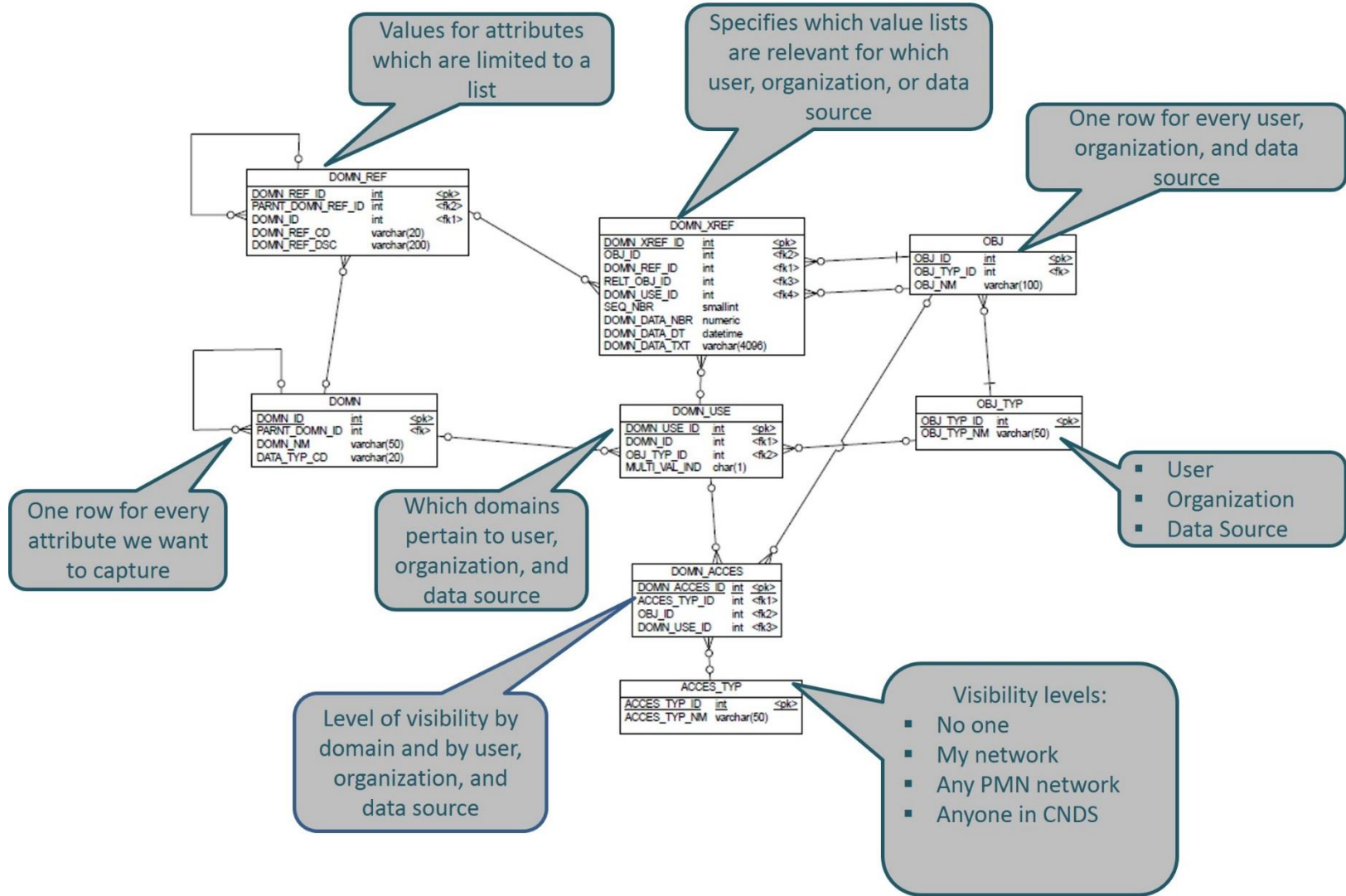
Data linkages

Birth Registry Immunization Registry
 Cancer Registry Death Registry

Data Model and User Interface

- The metadata inventory will change and grow over time
- The data model and software were built to accommodate change
 - The **data model** can be quickly updated with minimal training (no programming)
 - The **software user interface** (“profile” and “discovery” pages) are generated automatically from the metadata to simplify expansion and modifications
 - CNDS services include a robust metadata model that is extensible to **accommodate a virtually unlimited inventory of metadata fields**, without requiring further software development, and user interfaces that are programmatically generated based on the contents of the metadata model.

Data Model



Key Features and Functionality

- The CNDS software application is made up of four components:
 - Governance
 - Registration
 - Discovery
 - Communication

Governance

- Governance is central to CNDS
- All access controls currently available in PMN are supported in CNDS, for example:
 - Create, edit, delete networks, data sources, organizations, users
 - Respond to or reject requests
 - See Request Queue
- Governance is also incorporated into all CNDS components



Registration

- Much of governance is encoded in “Registration” metadata
 - Users can indicate “What” data they collect and “Who” they will share information with



Registration – What is collected

Cross Network Directory Service Welcome, A Person

Home My Profile Resources Profiles ▾ CNDS ▾ Contact Us Log Off

Home / DataMarts / A Datamart

DataMart Name* Acronym* Organization*

Data Domains [CNDS Visibility](#) [DataMart Permissions](#) [Request Type Permissions](#) [Installed Models](#)

Data Model

- PCORnet CDM
- Other
- Sentinel CDM
- Internal model

Source of Data

- EHR
- Claims
- Biorepository
- Patient-reported
- Registry
- Survey Data

Data linkages

- Birth Registry
- Cancer Registry
- Immunization Registry
- Death Registry

Types of Data Collected

- Rx Dispensing
- Demographics
- Outpatient Pharmacy Dispensing
- Longitudinal Capture
- Geocoded Data
- Vital Signs
- Cause of Death
- Biorepositories
- Patient Reported Health Behavior
- Death
- Other
- Cost Data
- Zipcode

Profile page where data source owners indicate data collected

Note that these are examples only, the system can capture anything deemed important to stakeholders



Registration – Who its Visible to

Data Domains **CNDS Visibility** DataMart Permissions Request Type Permissions Installed Models

| Domain | No One | My Network Members | All PMN Members | All PMN and CNDS Members |
|---------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|
| Types of Data Collected | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input checked="" type="radio"/> |
| Inpatient Encounters | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Inpatient Diagnosis Codes | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Inpatient Procedure Codes | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Laboratory Tests | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input checked="" type="radio"/> |
| Bilirubin | <input checked="" type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Glucose | <input type="radio"/> | <input checked="" type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Pregnancy | <input type="radio"/> | <input type="radio"/> | <input checked="" type="radio"/> | <input type="radio"/> |
| Genetic Testing | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input checked="" type="radio"/> |
| Zipcode | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input checked="" type="radio"/> |
| Data Model | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input checked="" type="radio"/> |

- Data source owners indicate what information is visible and discoverable
- Currently, there are 4 levels of visibility
- Visibility can be set per metadata element or by group

Discovery



The screenshot displays the 'Search Data Sources' interface. On the left, under 'Search Criteria', there are two sections: 'Types of Data Collected' and a list of categories. The 'Types of Data Collected' section is highlighted with a blue box and includes 'Inpatient Encounters' (with sub-items 'Inpatient Diagnosis Codes' and 'Inpatient Procedure Codes', each with 'ICD-9' and 'ICD-10' options) and 'Patient Reported Health Behavior' (with sub-items 'Vital Signs', 'Death', 'Cause of Death', and 'Other'). The second section is a tree view where 'Inpatient Encounters' and its sub-items are selected with checkboxes. On the right, the 'Data Source Found' section shows a 'New Request' button and 'Export to Excel' and 'Export to PDF' buttons. Below these is a table with columns: Network, Data Source, Organization, and Contact Infor... The table is currently empty.

On the discovery page, the user selects metadata elements to search, sees a summary of the search . . .

Discovery



Cross Network Directory Service Welcome, A Person

Home My Profile Resources Profiles ▾ CNDS ▾ Contact Us Log Off

Home / CNDS / Search Data Sources

Search Criteria

- Types of Data Collected
 - Inpatient Encounters
 - Inpatient Diagnosis Codes
 - ICD-9
 - ICD-10
 - Inpatient Procedure Codes
 - ICD-9
 - ICD-10

Demographics

- Patient Reported Health Behavior
- Vital Signs
- Death
- Cause of Death
- Other
- Inpatient Encounters
 - Inpatient Diagnosis Codes
 - ICD-9
 - ICD-10
 - SNOMED
 - Inpatient Procedure Codes
 - ICD-9
 - ICD-10

Data Source Found New Request

| Network | Data Source | Organization | Contact Infor... |
|-------------------|-------------------------------|---------------------------|------------------|
| CNDS Edge PCORNet | Honeybee Health Plan DataMart | Honeybee Health Plan | |
| CNDS Edge PCORNet | Caterpillar Cancer Center DM | Caterpillar Cancer Center | |
| CNDS Edge PCORNet | Blue Community DM | Blue Community Hospital | |

... and gets a result set with data sources whose owners have indicated they collect the data of interest and want this to be discoverable



Communication

- After finding data sources in Discovery, a user can:
 - Send an information or data request (query) to the data sources in a secure environment
 - Data request handling builds on standard PMN functionality
 - Track all requests and responses within the system

Outline

- Why a Cross-Network Directory Service
- Design: Flexible and Extensible
- Software Components
- **What's next**

5 Factors Keeping Networks Disconnected

- 1 ✓ Networks have **different governance policies** and different requirements for participation.
- 2 ✓ There is **no mechanism for broadcasting research capabilities** — the types of data available and the research and clinical expertise of their staffs — in a way that facilitates discovering common research interests and gives network participants control over who sees what.
- 3 ✓ Between networks there is **no secure and reliable means of making data requests** and tracking response activity.
4. There are no **operational standards or metrics for describing data** at a level that enables researchers to judge fitness-for-use of others' data sources.
5. There is **no reliable mechanism for sending queries** that will execute correctly across networks with different common data models.

CNDS addresses factors 1-3 through its Registration, Discovery, Communication, and Governance capabilities

Project Highlights

- Diversity of electronic health data sources creates challenges to identify data resources and potential collaborators.
- We implemented an open source web services system that enables:
 - Extensible metadata data model
 - Discovery of data sources
 - Governance rules
 - Cross-network communication
- Future work includes engaging with stakeholders to promote metadata curation, establishing governance rules, and implementing the service in a production environment.

What's Next

- The CNDS project has demonstrated the feasibility of:
 - Discovering potential research collaborators and new data sources and seamlessly querying them across independent distributed research networks.
 - This was demonstrated in Sentinel and PCORnet

What's Next

- To fully unify these networks — and pave the way for others to join — we recommend establishing a cross-network coordinating center to:
 - Develop and implement governance rules between networks
 - Promote and manage network engagement with CNDS
 - Operate the CNDS infrastructure
- Future work includes: software development to integrate CNDS into the main line of the PopMedNet software code; enhancing the Discovery functionality; implementing in production; engaging with stakeholders to promote metadata curation; establishing a CNDS coordinating center and governance rules.
- The web service uses standard APIs and could be integrated with other software applications in addition to PopMedNet.
- Despite considerable overlap between their common data models (CDMs), PCORnet and Sentinel data partners cannot send requests to each other. CNDS delivers a framework that enables such cross-network communication.

Acknowledgements

- CNDS team at Harvard Pilgrim Health Care Institute
 - Jeffrey Brown
 - Chayim Herzig-Marx
 - Jenny Hochstadt
 - Bridget Nolan
 - Kimberly Barrett
 - Zachary Wyner
 - Adam Paczuski
- Software development partners
 - Daniel Dee
 - Dean Corriveau
- PCORnet Partner - University of Michigan
 - Principal Investigator: Charles Friedman
- Sentinel Partner – Humana
 - Principal Investigator: Vinit Nair
- FDA
- ASPE



Jessica_malenfant@harvardpilgrim.org



popmednet.org
populationmedicine.org



@DeptPopMed