

SHRINE

Shared Health Research Information Network

A Federated Query Tool for i2b2 Databases

Use Cases

- Identifying cohorts for clinical trials
- Access to human specimens
- Population health surveillance
- Observational studies of genetic variants

SHRINE Prototype

- Goals
 - Query three i2b2 databases (4 hospitals)
 - Get approval from IRBs and hospital SVPs
 - Six month deadline
- Challenges
 - Protect the Developers
 - Protect the Patients
 - Protect the Hospitals

Protect the Developers

Limit the Scope

- Start with BIDMC, CHB, and PHS
- Simplify IRB process by limiting to aggregate queries and only seven users
- Limit ontology to demographics and diagnoses (ICD-9)
- Quick and dirty technical architecture

Protect the Patients

Obtain IRB Approval

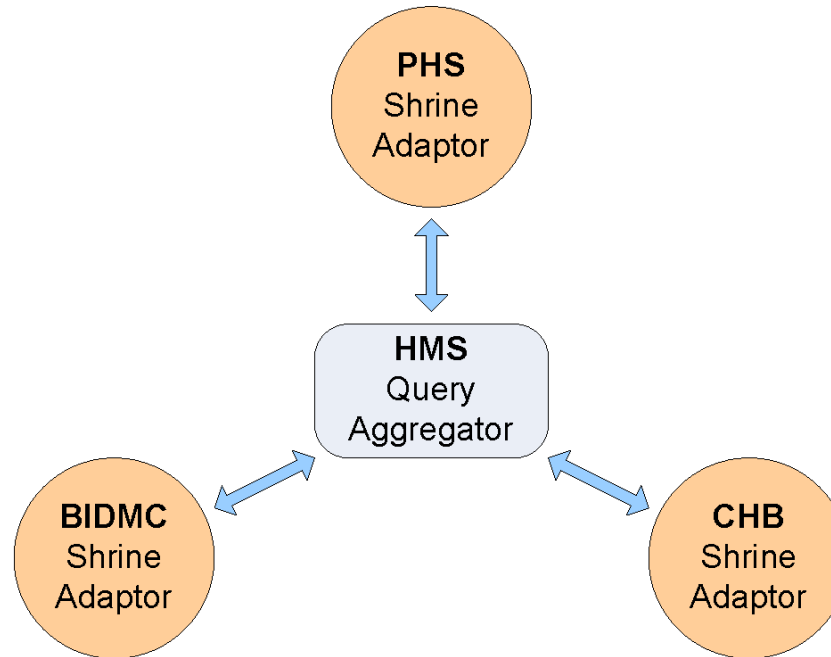
- No central database
- Only data from 2006
- Destroy prototype after limited time
- Patients signed HIPAA notice allowing personal health information to be used for research
- Aggregate counts only; obfuscate by adding small random number; display “<10” for small counts
- Log all queries
- Local PI responsible for data

Protect the Hospitals

Get Approval from Senior Leadership

- Hospitals can back out at any time
- Hospitals not identified by name (e.g., Hospital #1, #2, #3)
- List counts in random order
- Scale counts based on size of hospital
- Reveal all counts simultaneously so the speed doesn't reveal hospital

Prototype Architecture



Central “aggregator” broadcasts query to local hospital “adaptors”, which return aggregate counts only

Harvard Medical School

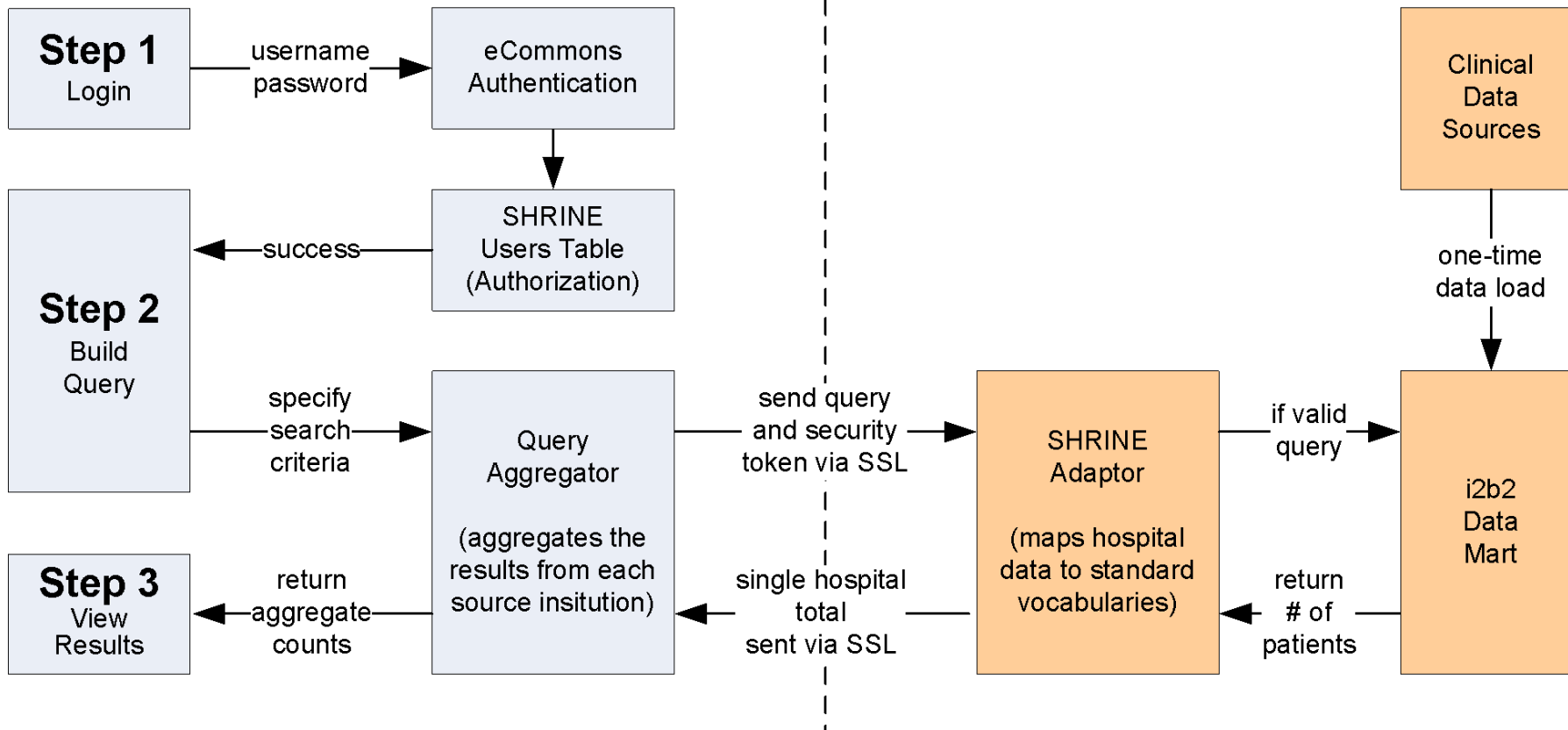
Source Institutions (only one shown)

Website

Web Services

Web Services

Database



Demo

i2b2 Query & Analysis Tool Logout

Navigate Terms **Find Terms**

- Demographics
 - Age
 - Gender
 - Female
 - Male
 - Unknown
 - Language
 - Marital-Status
 - Race
 - Vital-Status
- Diagnoses

drag an item from here

Query Tool

Query Name: _____

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude
			Regional enteritis Ulcerative colitis	Male		Neoplasms		

one or more of these **AND** **none of these** **AND** **one or more of these**

Run Query **New Query** 3 Groups **New Group**

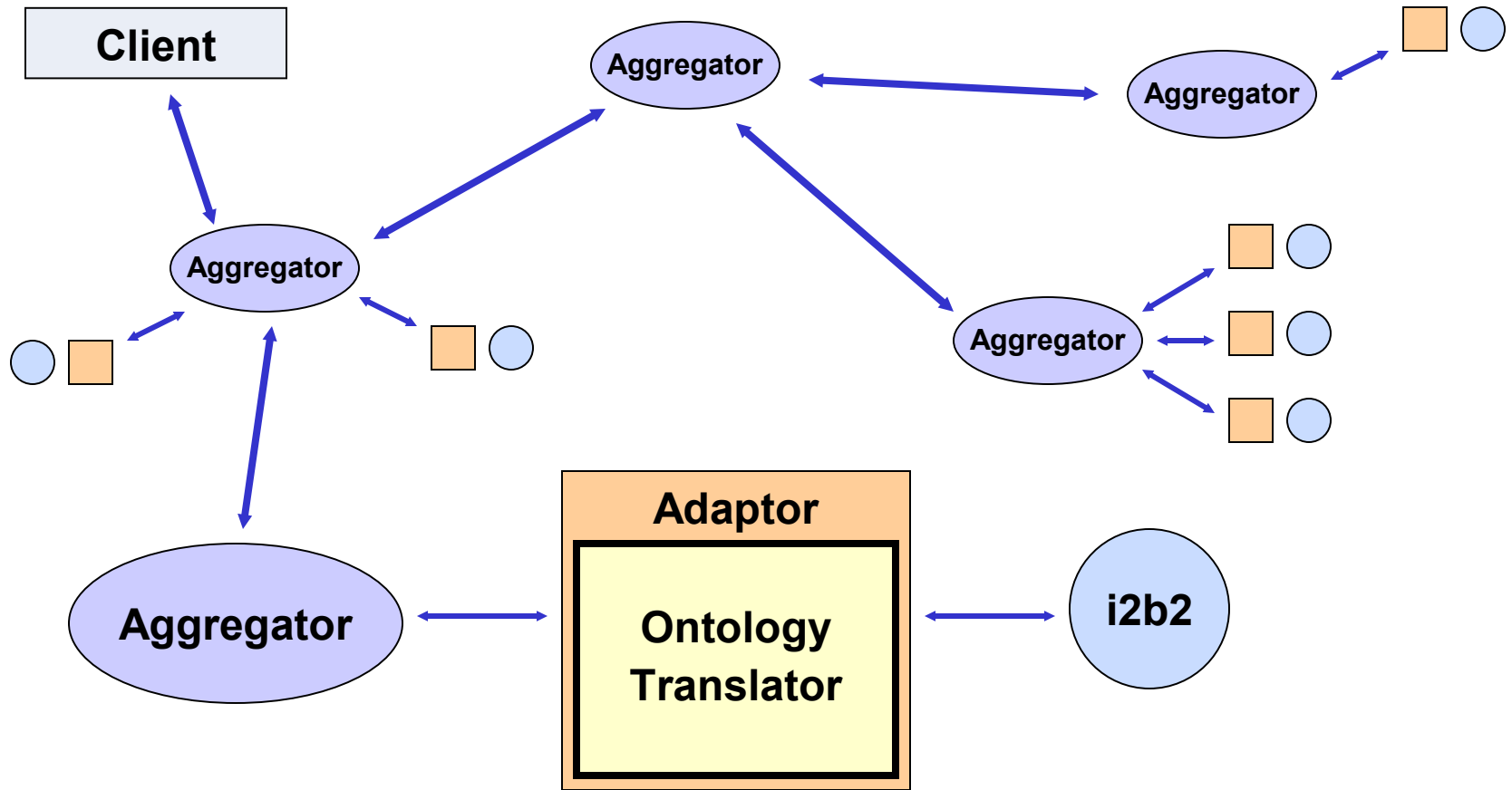
Previous Queries

- Noninfe-Female@16:39:03
- Nonin-Femal-Neopl@16:39:11
- Circulatory sys@16:39:18
- Circulatory sys@17:01:03
- Circula-Events @17:04:11
- Demographics@17:04:21
- 10-19 years old@17:04:39
- Albanian@02:31:25

Query Status

Executing query...
Elapsed time (seconds): 14.0
Query Finished...
Matching patients (hospital 1): 332 (+/-3)
Matching patients (hospital 2): 16 (+/-3)
Matching patients (hospital 3): 151 (+/-3)

Peer-to-Peer Architecture



Broadcast-Aggregation based on SPIN

We need your help creating the ontology translator

Timeline

Year 1 (Regulatory 12/1/2008, Go-Live 6/30/2009)

1. Harvard-wide query capability for aggregate data
2. Limited data sets – view one patient at a time

Year 2 (Regulatory 12/1/2009, Go-Live 6/30/2010)

1. Limited data sets – full download capability

If capacity is available and institutions agree

1. Limited data sets with NLP processed text notes
2. Limited data sets with tissue sharing