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**

**Website**

**Step 1**
Login

- username/password

**Step 2**
Build Query

- specify search criteria

**Step 3**
View Results

- return aggregate counts

**Web Services**

eCommons Authentication

- SHRINE Users Table (Authorization)

Query Aggregator

- (aggregates the results from each source institution)

**Source Institutions**

(only one shown)

**Web Services**

- send query and security token via SSL

**Database**

- if valid query

Clinical Data Sources

- one-time data load

i2b2 Data Mart

- return # of patients
Demo
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