

# **i2b2 Clinical Research Chart**

***Shawn Murphy MD, Ph.D.***

***Griffin Weber MD, Ph.D.***

***Michael Mendis***

***Vivian Gainer MS***

***Lori Phillips MS***

***Rajesh Kuttan***

***Wensong Pan MS***

***Henry Chueh MD***

***Susanne Churchill Ph.D.***

***John Glaser Ph.D.***

***Isaac Kohane MD, Ph.D.***

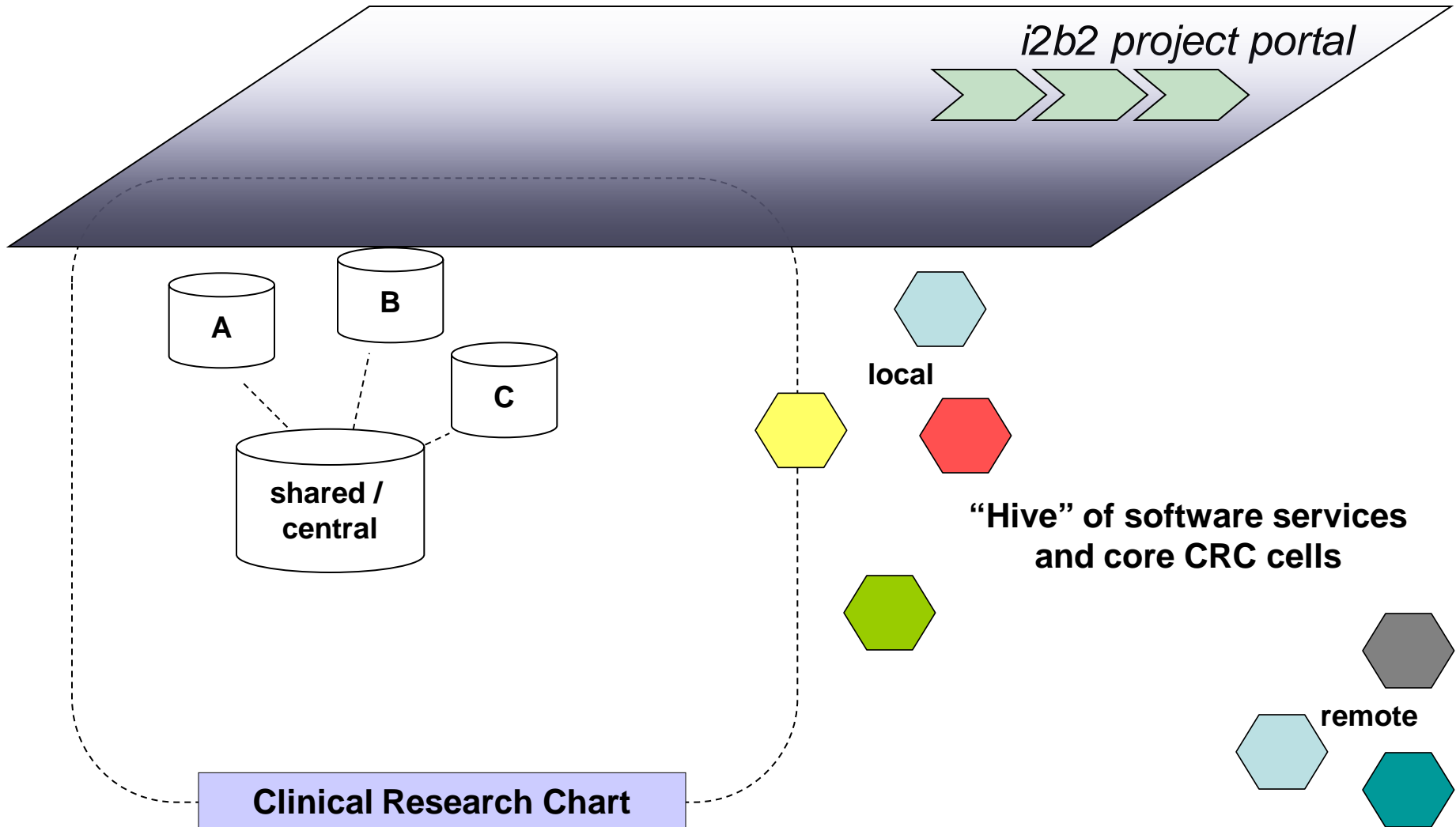
# Use Cases

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

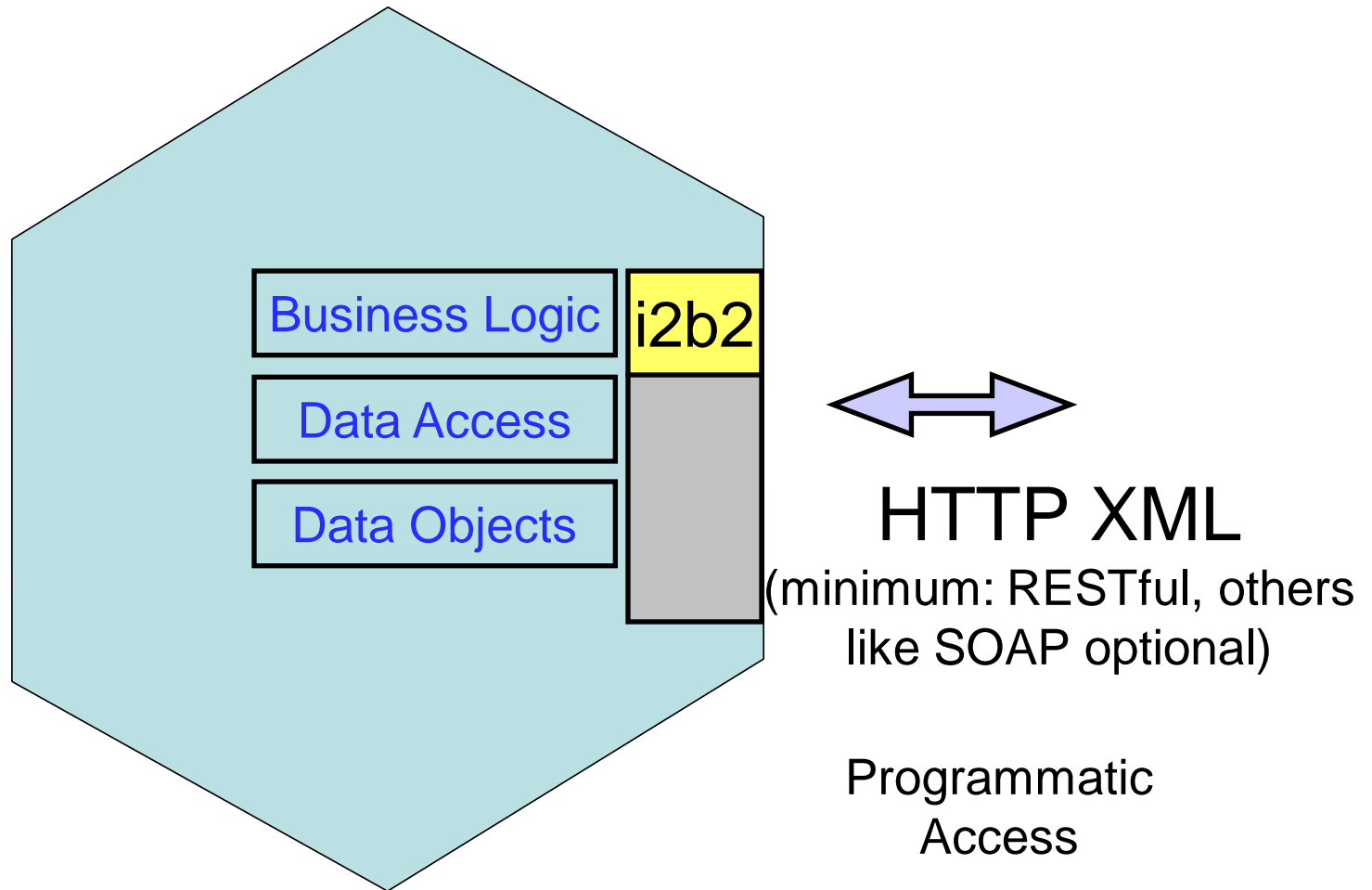
# Technical Overview

- Formed as a collection of interoperable services provided by i2b2 Cells
- Loosely coupled
- Makes no assumptions about proximity
- Connected by Web services
- Activity can be directed manually or automatically

# i2b2 Environment



# i2b2 Cell: Canonical Hive Unit

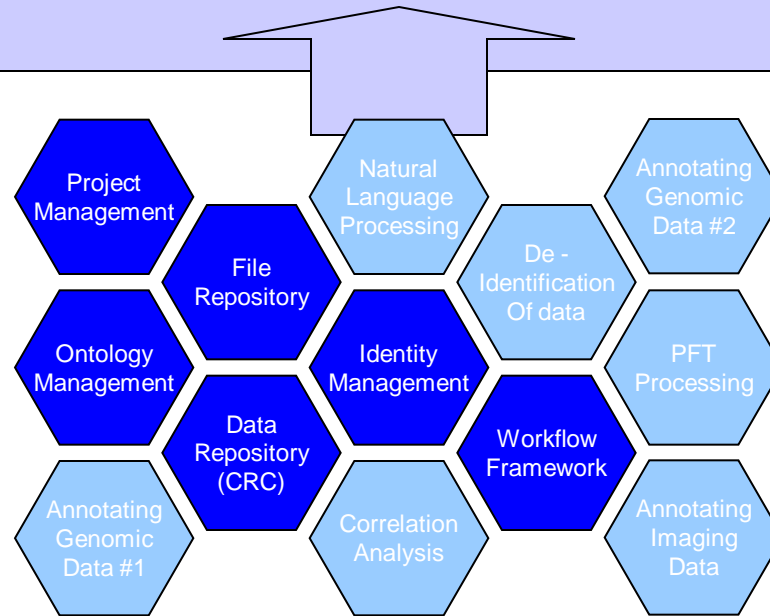


# Use Case Development of i2b2 technology

Enterprise-wide repurposing and distribution of medical record data for research

- Hospital-wide availability of clinician-driven queries to find research cohorts
- Suggestion of novel relationships between genes and disease through mutual information theory methods
- Display of large collections of odds ratios as calculated following environmental exposures (Pharmacovigilance)

# Enterprise-wide repurposing and distribution of medical record data for research



- Enable high performance collection of medical record data for querying and distribution
  - Enterprise web client
  - Create patient cohorts for further investigation
- Enable discovery within data on enterprise wide scale
  - Relevance networks
  - Pharmacovigilance

# Enterprise web client

<http://services.i2b2.org/webclient/>

The screenshot shows the i2b2 Query & Analysis Tool web client interface. The browser window title is "i2b2 Web Client - Microsoft Internet Explorer" and the address bar shows "http://services.i2b2.org/webclient/#". The interface includes a navigation menu with "Find Patients", "Analysis Tools", "Message Log", and "Help | Logout".

The main interface is divided into several sections:

- Navigate Terms / Find Terms:** Search by Names (contains) and Search by Codes (multiple sclerosis). A "Find" button is present. Below the search results, "Multiple sclerosis" is listed.
- Workplace:** A file explorer showing "SHARED" and "demo" folders.
- Query Tool:** Query Name: "Moderat-Multipl@02:14:47". It displays a query structure with three groups:

Group 1	Group 2	Group 3
Moderate Alzheimers dise...	Multiple sclerosis	

Below the groups are logical operators: "none of these" (red), "AND", "one or more of these" (green), "AND", and "drop a term on here" (yellow). Buttons for "Run Query", "New Query", and "New Group" are visible.
- Previous Queries:** A list of recent queries including "Moderat-Multipl@02:14:47 [2-24-2009] [demo]", "BetaBlockerStudy [2-24-2009] [demo]", and "Age-Circu-WBC (@21:36:01 [2-24-2009] [demo)".
- Query Status:** Shows "Executing query...", "Elapsed time (seconds): 2.8", "Query Finished..", and "Matching patients: 2".



i2b2 Web Client - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites RSS Mail Print Writable Yellow Folder Mailbox People

Address http://services.i2b2.org/webclient/# Go

Google Go RS Settings **lenovo**

**i2b2 Query & Analysis Tool** Find Patients | **Analysis Tools** | Message Log | Help | Logout

**Navigate Terms** **Find Terms**

Search by Names Search by Codes

contains multiple sclerosis

Find Any Category

Multiple sclerosis

**Workplace**

SHARED

demo

**Previous Queries**

- Moderat-Multipl@02:16:36 [2-24-2009] [de]
- Results of Moderat-Multipl@02:16:36 [2-24-2009] [de]
  - Patient Set - 2 patients
  - Patient Count - 2 patients
- Moderat-Multipl@02:14:47 [2-24-2009] [de]
- Moderat-Multipl@02:14:24 [2-24-2009] [de]

**Timeline**

Specify Data View Results Plugin Help

<<< start: 1 size: 10 go >>> zoom: - + pan: < >

3/14/1998 11/16/2001 7/23/2005

**Person\_#100000016\_Female\_23yroid\_Indian**  
Multiple sclerosis | | | | |

**Person\_#100000078\_Female\_27yroid\_White**  
Multiple sclerosis | | | | | | | | | | | |

**Plugins**

Detailed List View Category: ALL

**Demographics (2 Patient Sets) - Simple Counts**  
This plugin compares the demographic break-down of two Patient Sets.

**Timeline**  
This plugin creates a visual representation of when selected observations occur within a given patient set.

Internet

# Workbench display of Mutual Inform. Correlations

i2b2 Workbench

File Window Help Search Run

i2b2 Workbench for Demo Data Vladimir Valtchinov Status: ● Wiki

Correlation Analysis Cell

Concepts Selection >> Calculate Valid Intervals >> Calculate Correlations >> Results | About

Pairs Results | Compare Graphs

Search And Display

Metric: MIC

TYPE 1 or TYPE 2 is: (shows only pairs with this type)

Threshold: (sets threshold value on selected metric)

Top: (most correlated pairs on selected metric)

Calculation method: NORMAL

Show: All results

type 2	Description 2	MIC	Pearson	Vect
97_s_at	29100	0.8636	0.8831	1390
00_s_at	28971	0.8581	0.8891	1390
96_s_at	84060	0.795	0.765	1390
00_s_at	80233	0.7393	0.8119	1390
97_s_at	29100	0.6629	0.7993	1390
96_s_at	84060	0.6348	0.5554	1390
97_s_at	29100	0.6099	0.5387	1390
221608_at	28971	0.5928	0.5447	1390
00_s_at	28971	0.5615	0.5783	1390
02_s_at	9214	0.5565	0.9631	1390

yED Graph Options

Show metric

Negative metric colour

221597\_s\_at

Overlap in time: Scatter plot

Row No.	Vector 1	Vector 2
1	392.10287	1089.9792
2	737.32404	1621.6381
3	586.41644	863.0268
4	589.6359	1308.3345
5	444.2756	1392.4344
6	790.24536	1030.8126
7	525.5525	890.79297
8	572.63	1724.3481
9	545.0431	1450.0469
10	575.0798	1538.5139
11	646.5513	1274.1931
12	301.8552	1118.3007
13	536.53143	1036.4288
14	635.1086	1126.2897
15	898.6878	1378.0665
16	284.69394	1368.532

Patient Set: All

Close

# Odds Ratios for Diseases expressed in comparing Rosiglitazone vs. Pioglitazone

**i2b2 Workbench for Demo Group** | Demo User | Status: ● | i2b2

**Query Tool**

Query Name: \_\_\_\_\_ | Reset Groups

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude

Thiazolidinediones

The terms of this group are joined then intersected with other groups | Drag terms from Navigate, Find and Workplace into this group | Drag terms from N Find and Workplace int

Run Query | Patient(s) returned: \_\_\_\_\_

**Pharmacovigilance Analysis** | Timeline View | Text Analyzer

Set up analysis | **Treemap** | Breakdown | Timeline | Export

Analysis Title: Avandia vs. Pioglitazone Calculation for SNM0 on 3-04-2009

Population is Defined by Query: All Diabetics with ICD9 250.XX coded in the Diabetes clinic

Make Treemap |  Make Breakdown |  Make Timeline

Run Calculation | Cancel Calculation

**Outcome to be measured**

Dates	Occurs > 0x	Exclude
<input checked="" type="checkbox"/>		
<input type="checkbox"/>		
<input type="checkbox"/>		

**Outcome to be compared**

Dates	Occurs > 0x	Exclude
<input checked="" type="checkbox"/>		
<input type="checkbox"/>		
<input type="checkbox"/>		

**Exposed group definition**

Dates	Occurs > 0x	Exclude
<input checked="" type="checkbox"/>		
<input type="checkbox"/>		
<input type="checkbox"/>		

**Unexposed group definition**

Dates	Occurs > 0x	Exclude
<input checked="" type="checkbox"/>		
<input type="checkbox"/>		
<input type="checkbox"/>		

**Calculation Results**

Running calculation...

Matrix:

4454	54673
6579	61275

Chi Square: 1.34  
Error range: 1.20 - 1.46

Making Treemap...

**Observations should occur:**

Any time in patient history

In same patient visit

In same patient period

With exposure first

**Unexposed group should:**

Match ages

Match Genders

Match Race / Ethnicities

Match co-morbidity levels

Open File .. | Save File .. | Open in Excel

**Workplace**

- SHARED
- demo
  - Cardiac Tests
  - Cardiovascular agents
  - Definitions (Queries)
  - Definitions (Query Groups)
  - Demogra-Diagnos@02:31:36 [1-11-
  - Diabetes mellit@22:57:00 [1-15-200
  - Diagnoses
  - IRON (LOINC:2498-4)
  - Iron (G-ALT (5G@09:08:27 [12-15-2
  - Observations
  - PATIENT:1000000025
  - PATIENTSET\_8971\_Ischemic
  - Patient Sets

**Navigate Terms**

- Laboratory Tests
- Medications
  - Alternative medicines
  - Anti-infectives
  - Antihyperlipidemic agents
  - Antineoplastics
  - Biologicals
  - Cardiovascular agents
  - Central nervous system ager
  - Coagulation modifiers
  - Gastrointestinal agents
  - Hormones
    - Adrenal cortical steroids
    - Antidiabetic agents
      - Alpha-glucosidase int
      - Insulin
      - Meglitinides
      - Non-sulfonylureas
      - Sulfonylureas
      - Thiazolidinediones
      - Bisphosphonates

**Previous Queries**

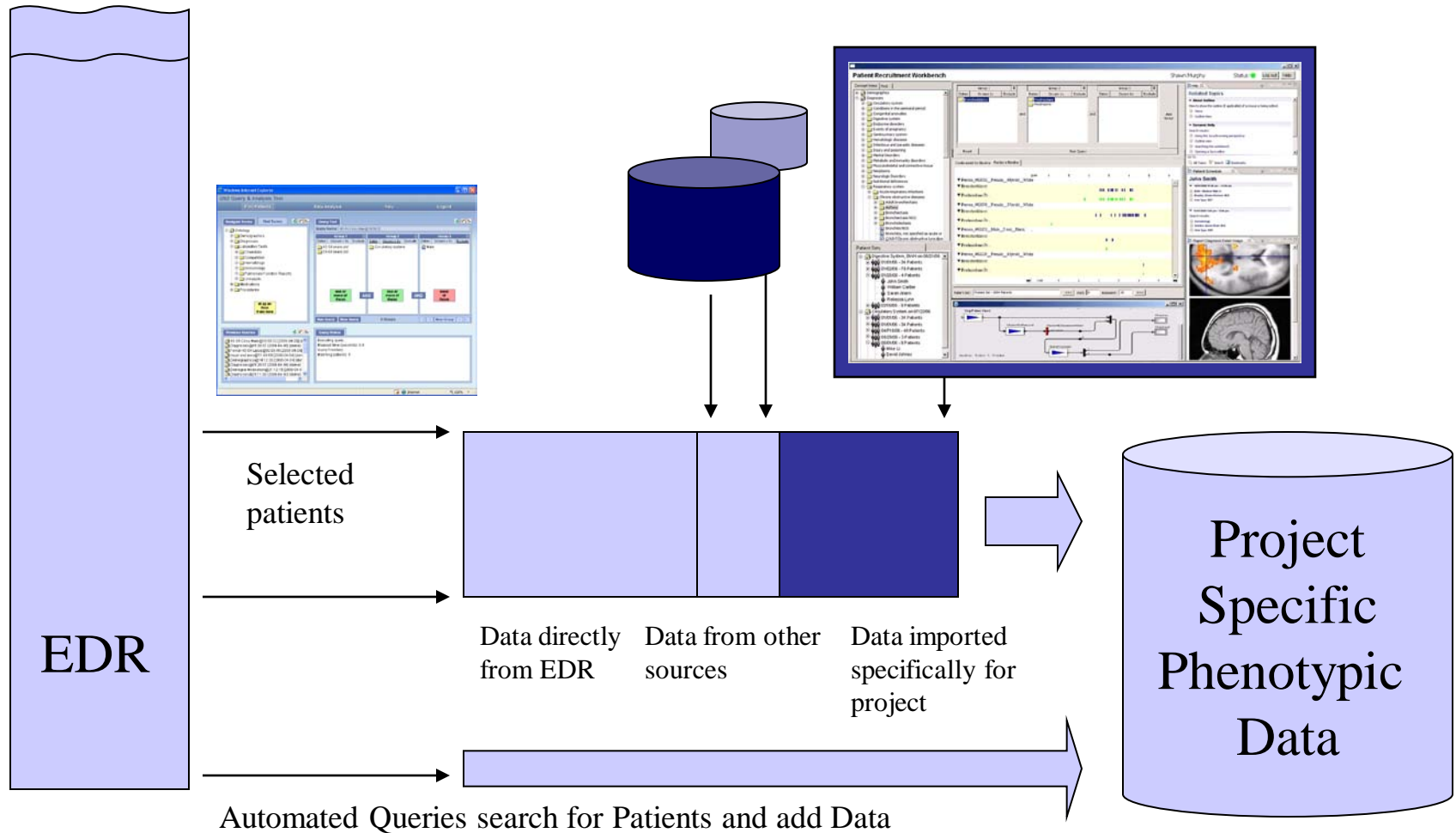
- Female-Asthma@04:54:07 [02-16-2009 ]
- Female-Asthma [02-16-2009 ] [demo]
- Immunology@01:38:28 [02-15-2009 ] [d
- Demographics@18:15:23 [02-14-2009 ] [
- 0-9 yea-Male@22:21:20 [02-14-2009 ] [c
- Demographics@18:15:23 [02-14-2009 ] [
- Demographics@18:15:23 [02-14-2009 ] [
- 0-9 yea-Male@22:21:20 [02-14-2009 ] [c
- Demographics@18:15:23 [02-14-2009 ] [
- Demographics@18:15:23 [02-14-2009 ] [
- 0-9 yea-Male@22:21:20 [02-14-2009 ] [c
- Demographics@18:15:23 [02-14-2009 ] [
- Demographics@18:15:23 [02-13-2009 ] [
- 0-9 yea-Male@22:21:20 [02-13-2009 ] [c
- Demographics@18:15:23 [02-13-2009 ] [
- Demographics@18:15:23 [02-13-2009 ] [
- 0-9 yea-Male@22:21:20 [02-13-2009 ] [c

# Use Case Development of i2b2 technology

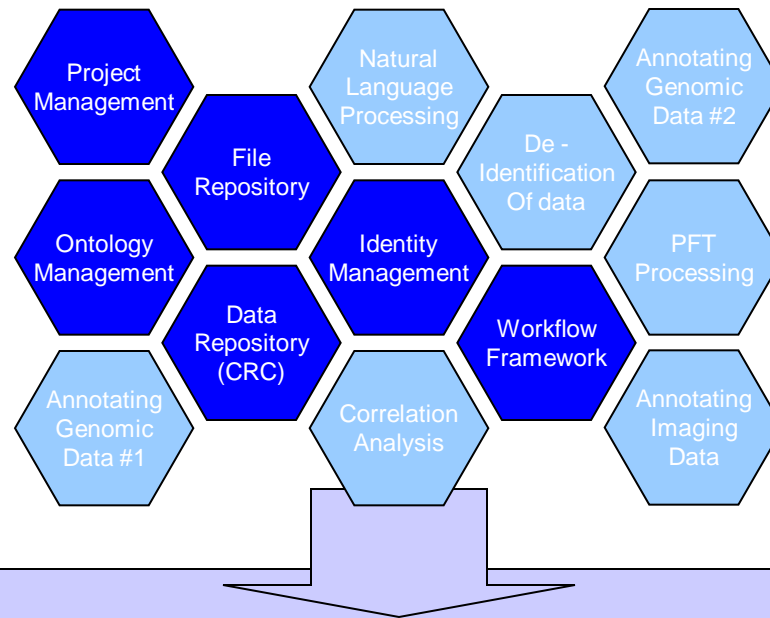
Provide distribution of medical records and tools for their manipulation in supporting specific research projects.

- Select patients for clinical trials
- Working out details for calculating odds ratios of pharmacovigilance
- Annotating images in clinical studies

# Set of patients is selected through Enterprise Repository and data is gathered into a data mart



- Repurpose medical record information for research studies
  - I2b2 Workbench
  - Natural language processing
- Enable genomic studies
  - Tissue/blood selection
  - Genetic data integration



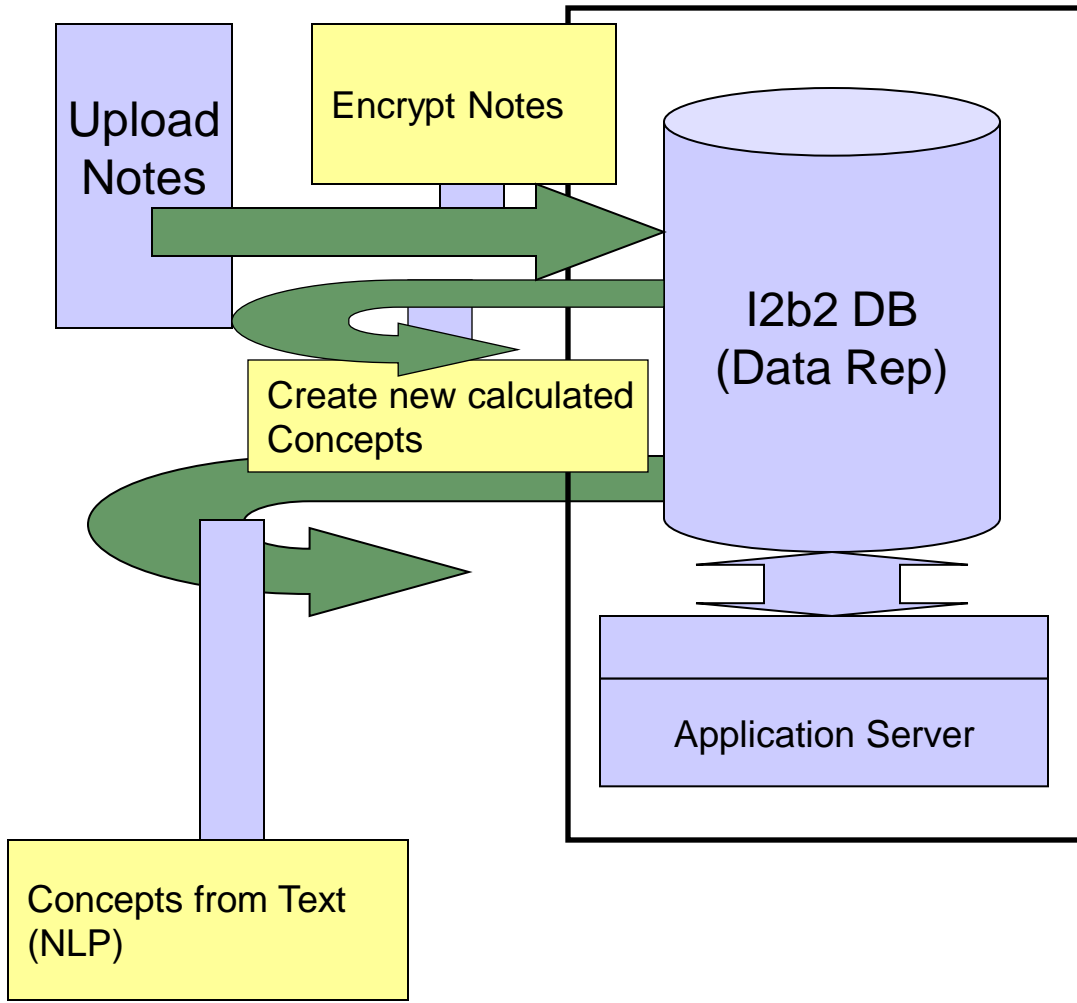
Use of medical record data in clinical studies  
focused upon genomics and pharmacology

# Select patients for clinical trials

The screenshot displays the BIRN Workbench interface, which is used for managing and viewing medical data. The main window is titled "BIRN Workbench for Demo Group" and shows a "Query Tool" with a "Timeline View" tab. The "Select Patients" section displays a table of patient data:

Set #	Decision	PSet #	Patient ID	Patient Name	Gender	Race	Date of Birth	Age	MRNs
1	Yes	1	1000000001	xxxxxx, xxxxxx	F	black		74	
2	Yes	1	1000000002	xxxxxx, xxxxxx	F	white		55	
3	Yes	1	1000000003	xxxxxx, xxxxxx	M	asian		73	
4	UnD	1	1000000004	xxxxxx, xxxxxx	M	black		28	

The "Previous Queries" panel on the left shows a list of image sets, with "Patient Set - 416 Patients" highlighted. The "Image View" panel on the right shows a list of events for a specific subject (ID: 1000000001) and a corresponding MRI image. The image is labeled "OAS1\_0001\_MR1/RAW/OAS1\_0001\_MR1\_mpr-1\_anon.img" and shows a sagittal view of a brain. The interface also includes a "Wensong Pan" status bar and a "Wiki" link.





**Navigate Terms**

- Avandia\_MLP
  - PV
    - Avandia
      - Cardiovascular Event
        - Enrolled
          - No Rosi And CV
          - No Rosi And No CV
          - Rosi And CV
          - Rosi And No CV
- RPDR
  - Demographics
  - Diagnoses
  - DRG
  - Health History
  - Laboratory Tests
  - Medications
    - Alternative medicines
    - Anti-infectives
    - Antihyperlipidemic agents
    - Antineoplastics
    - Biologicals
    - Cardiovascular agents
    - Central nervous system agents
    - Coagulation modifiers
    - CPT Devices
    - CPT Medications
    - Devices
    - Gastrointestinal agents
    - Hormones
      - Adrenal cortical steroids
      - Antidiabetic agents
        - Alpha-glucosidase inhibitors
          - Byetta (exenatide) - LMR 5176
        - Insulin
        - Meglitinides
        - Non-sulfonylureas
        - Sulfonylureas
        - Thiazolidinediones
          - Metformin-rosiglitazone
          - Pioglitazone
          - Rosiglitazone

**Previous Queries**

- Avandia-Cardiov@09:20:31 [11-24-2008] [snm0]
- Enrolled@10:55:15 [11-13-2008] [snm0]
- Actos@09:39:18 [11-03-2008] [snm0]
- Actos@09:37:00 [11-03-2008] [snm0]
- Actos@09:36:48 [11-03-2008] [snm0]
- Acute Rheumatic@11:07:49 [07-02-2008] [snm0]
- Circulatory sys@11:05:19 [07-02-2008] [snm0]
- Circulatory sys@11:00:50 [07-02-2008] [snm0]
- Medications@11:00:31 [07-02-2008] [snm0]

**Query Tool**

Query Name:  Reset Groups

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	:clude	Dates	Occurs > 0x	:clude	Dates	Occurs > 0x	:clude

Add Group

Timeline  
 Patient Count

Run Query Patient(s) returned:

**Timeline View** Export Data KTreeMap View

Create model for Timeline Render a Timeline

Query Name:  Avandia-Cardiov@09:20:31 [11-24-2008] [snm0]

Row #	Name of Terms	Value	Value Text	Height	Color
1	Enrolled	N/A	N/A	Medium	
2	Rosiglitazone event	N/A	N/A	Medium	
3	Cardiovascular event	N/A	N/A	Medium	
4	No Rosi And CV	N/A	N/A	Medium	
4	No Rosi And No CV	N/A	N/A	Medium	
4	Rosi And CV	N/A	N/A	Medium	
4	Rosi And No CV	N/A	N/A	Medium	
5	Avandia - MLP	N/A	N/A	Medium	
5	rosiglitazone - MLP	N/A	N/A	Medium	

Delete From List Delete All Put In Order Move Up Move Down

Display concepts with no data  Display patient demographics

Patient Set:  Patient Set: 954 patients 
 start:  increment:

**Workplace**

- SHARED
  - snm0
    - Avandia-Cardiov@09:20:31 [11-24-2008] [snm0]

**Navigate Terms**

- Avandia\_MLP
  - PV
    - Avandia
    - Cardiovascular Event
    - Enrolled
    - No Rosi And CV
    - No Rosi And No CV
    - Rosi And CV
    - Rosi And No CV
  - RPDR
    - Demographics
    - Diagnoses
    - DRG
    - Health History
    - Laboratory Tests
    - Medications
      - Alternative medicines
      - Anti-infectives
      - Anthyperlipidemic agents
      - Antineoplastics
      - Biologicals
      - Cardiovascular agents
      - Central nervous system agents
      - Coagulation modifiers
      - CPT Devices
      - CPT Medications
      - Devices
      - Gastrointestinal agents
      - Hormones
        - Adrenal cortical steroids
        - Antidiabetic agents
          - Alpha-glucosidase inhibitors
            - Byetta (exenatide) - LMR 5176
          - Insulin
          - Meglitinides
          - Non-sulfonylureas
          - Sulfonylureas
          - Thiazolidinediones
            - Metformin-rosiglitazone
            - Pioglitazone
            - Rosiglitazone

**Query Tool**

Query Name:  Reset Groups

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	!clude	Dates	Occurs > 0x	!clude	Dates	Occurs > 0x	!clude

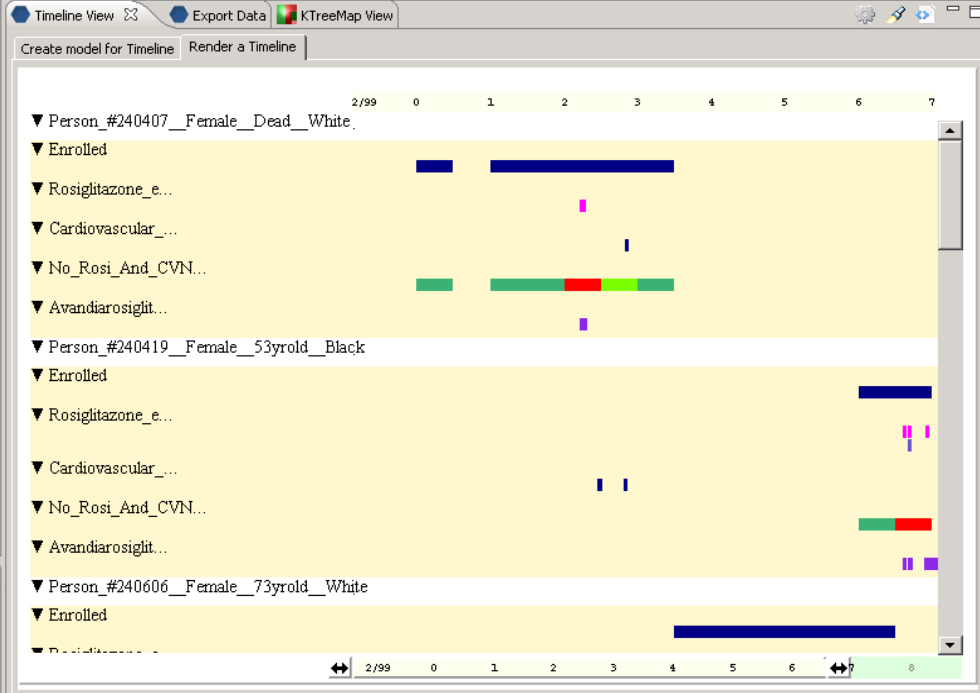
Add Group

Timeline  
 Patient Count

Run Query Patient(s) returned:

**Workplace**

- SHARED
  - snm0
    - Avandia-Cardiov@09:20:31 [11-24-2008] [snm0]

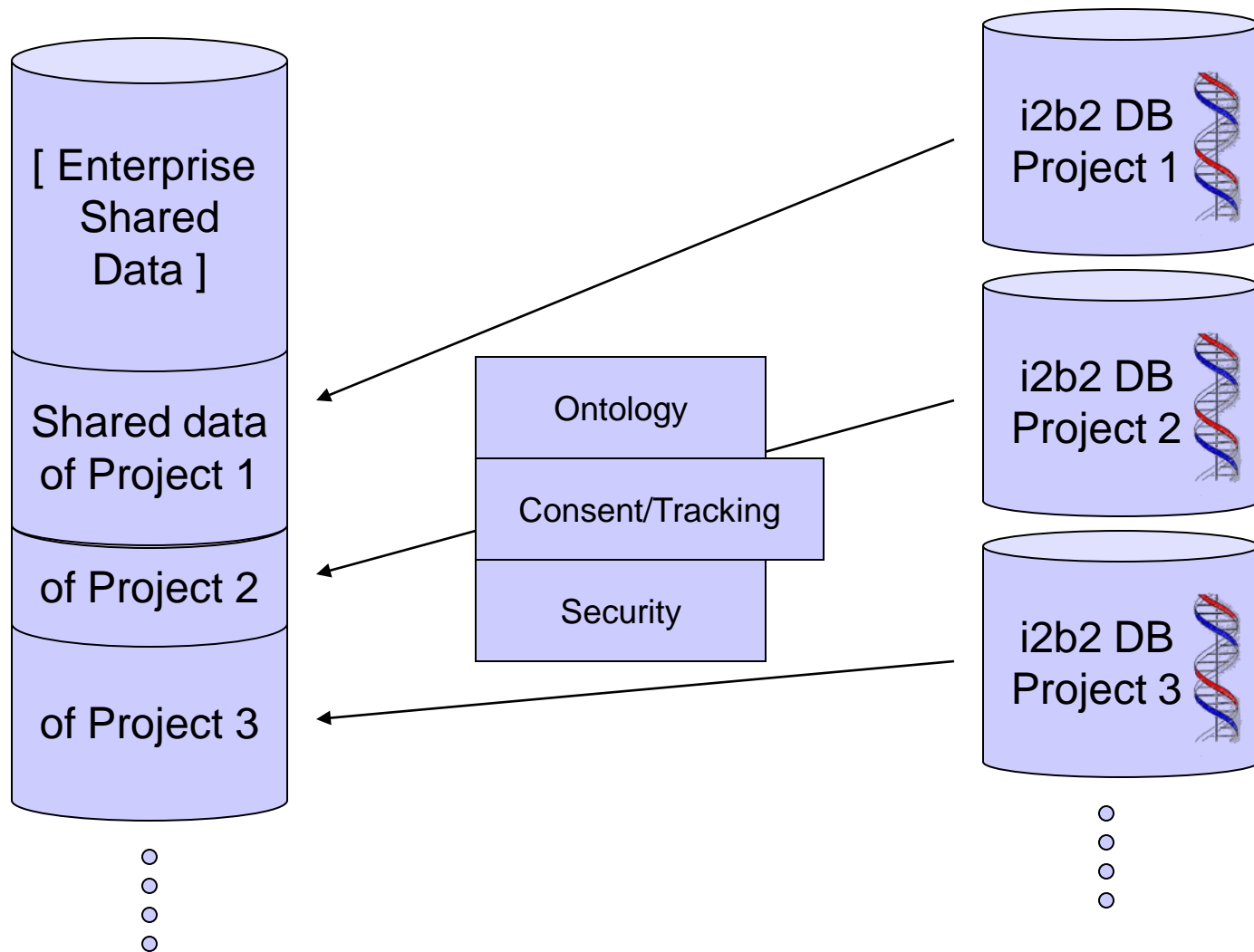


**Previous Queries**

- Avandia-Cardiov@09:20:31 [11-24-2008] [snm0]
- Enrolled@10:55:15 [11-13-2008] [snm0]
- Actos@09:39:18 [11-03-2008] [snm0]
- Actos@09:37:00 [11-03-2008] [snm0]
- Actos@09:36:48 [11-03-2008] [snm0]
- Acute Rheumatic@11:07:49 [07-02-2008] [snm0]
- Circulatory sys@11:05:19 [07-02-2008] [snm0]
- Circulatory sys@11:00:50 [07-02-2008] [snm0]
- Medications@11:00:31 [07-02-2008] [snm0]

Patient Set:  <<< start: 1 increment: 10 >>>

# Project data can be added back to Enterprise Repository



# Enterprise Querytool allows:

- Selection of patient populations through complex queries
- Specifically targeted population analysis
- Privacy Management

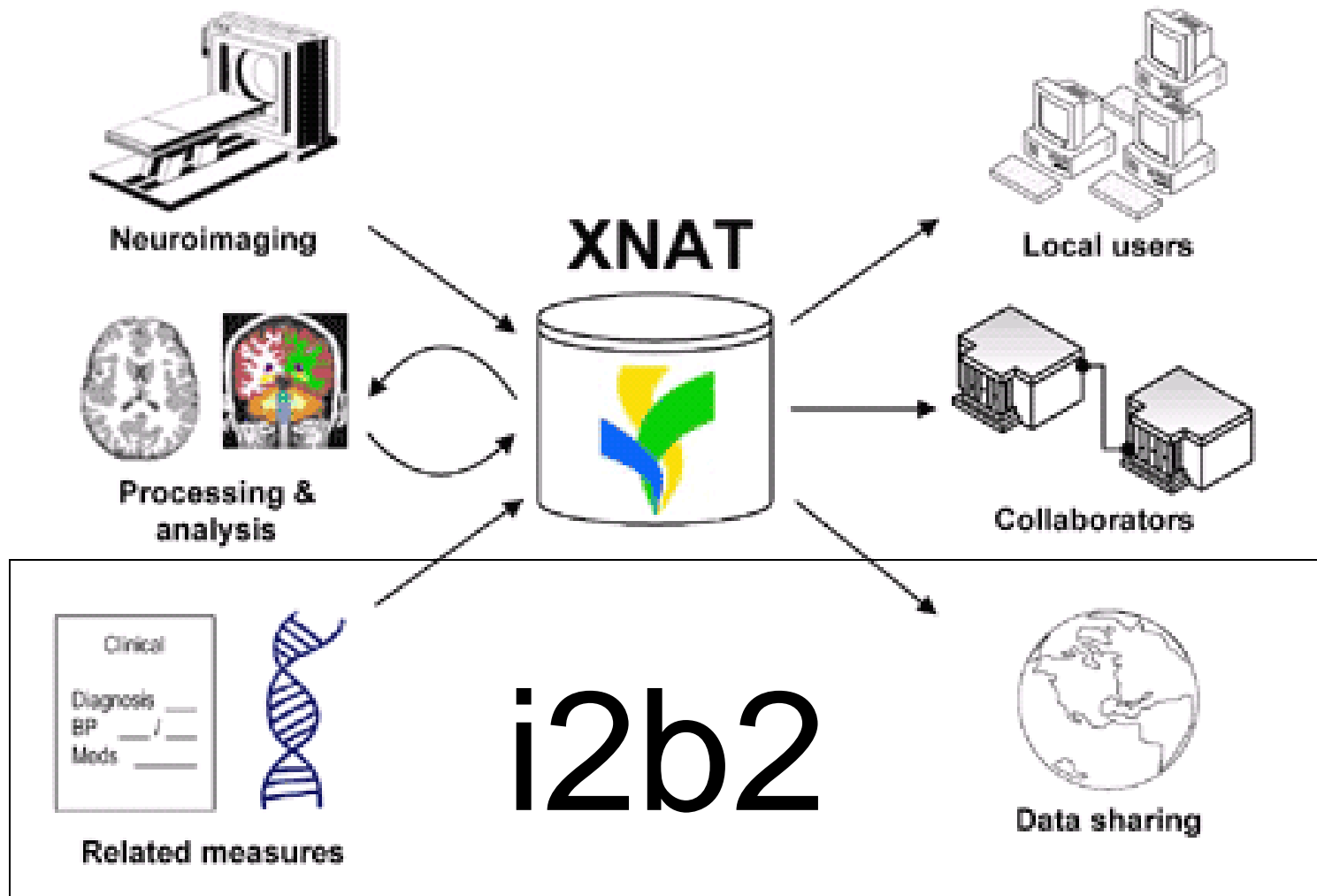
## Project Workbench allows:

- Patient population medical record review and exploration of data in a targeted population
- Use of Natural Language Processing
- Import of batch data sets
- Workflow support to simulate sequences of operations
- Export of batch outputs for specific applications
- Privacy management

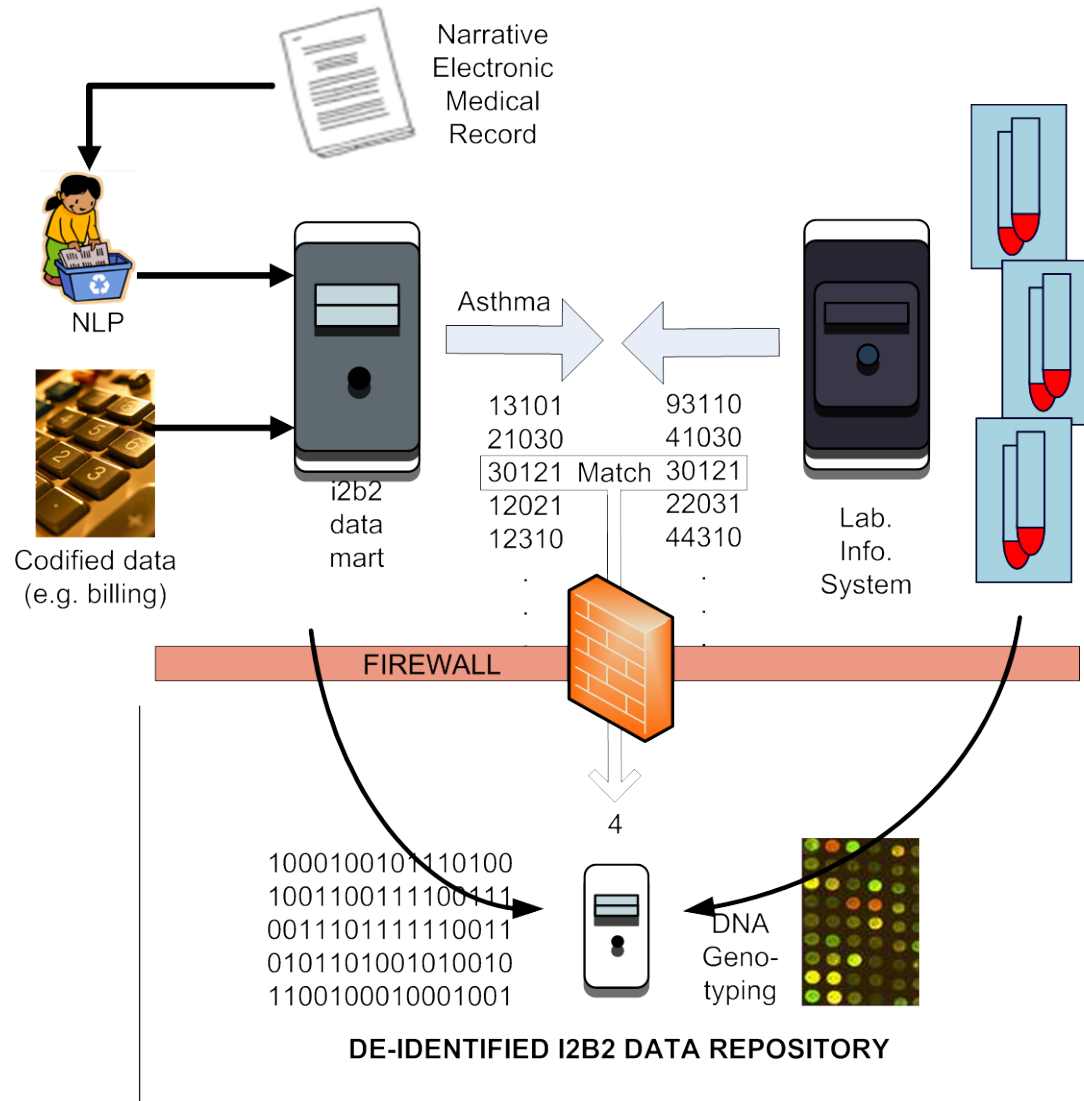
# Help outside of i2b2

- Ontology management
  - NCBO
- Imaging Research
  - BIRN / XNAT
- Tissue management
  - Crimson
- Cross Platform Integration
  - caBIG / SHRINE

# Extensible Neuroimaging Archive Toolkit (XNAT) used to integrate image management into i2b2

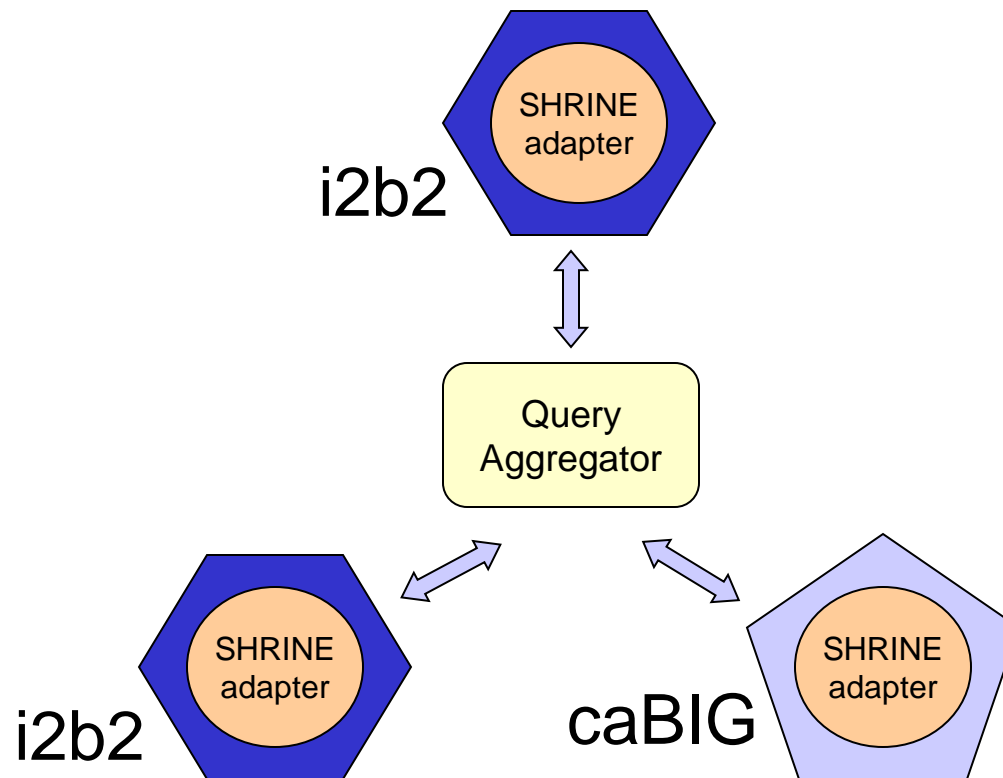


# Genotype samples and compare to controls





# SHRINE (Shared Research Informatics Network) = Distributed Queries



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

# Prototype Screenshot

**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**

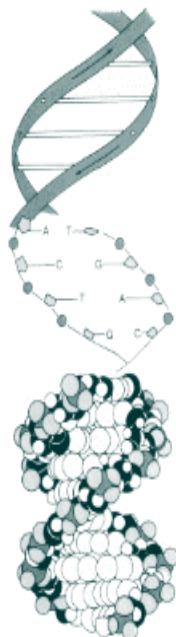
**Run Query**   **New Query**   3 Groups   **New Group**

**Previous Queries**

- Nonife-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)



# **i2b2 Clinical Research Chart**

***Shawn Murphy MD, Ph.D.***

***Griffin Weber MD, Ph.D.***

***Michael Mendis***

***Vivian Gainer MS***

***Lori Phillips MS***

***Rajesh Kuttan***

***Wensong Pan MS***

***Henry Chueh MD***

***Susanne Churchill Ph.D.***

***John Glaser Ph.D.***

***Isaac Kohane MD, Ph.D.***