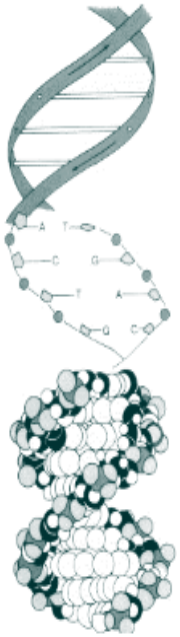


Introduction to i2b2 Software Platform

Shawn Murphy MD, Ph.D.
Griffin Weber MD, Ph.D.
Michael Mendis
Vivian Gainer MS
Andrew McMurry MS
Lori Phillips MS
Rajesh Kuttan
Wensong Pan MS
Nick Benik
Janice Donahue
Susanne Churchill Ph.D.
John Glaser Ph.D.
Isaac Kohane MD, Ph.D.

Agenda

8:30 i2b2 Overview (Zak Kohane)

8:35 Introduction to i2b2 Software Platform
(Shawn Murphy, Mike Mendis, Vivian Gainer, Griffin Weber)

9:35 SHRINE Regulatory Issues ("How One Multipart System Agreed to Share Data") (Susanne Churchill)

10:05 Break

10:35 Introduction to SHRINE (Andy McMurry)

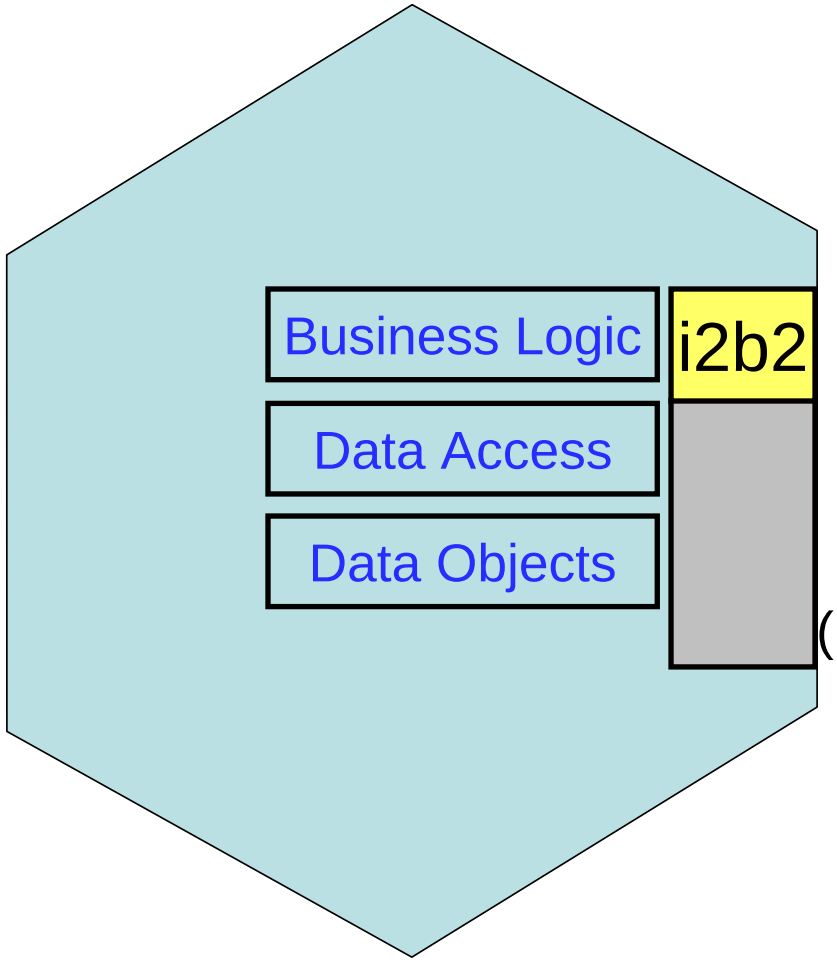
11:05 i2b2 Planned Applications for Cool Science (Zak Kohane)

11:30 i2b2 Plans for New Software Enhancements and Incorporating Contributions from the Community (Shawn Murphy)

The National Center for Biomedical Computing entitled Informatics for Integrating Biology and the Bedside (i2b2) Clinical Research Chart, what is it?

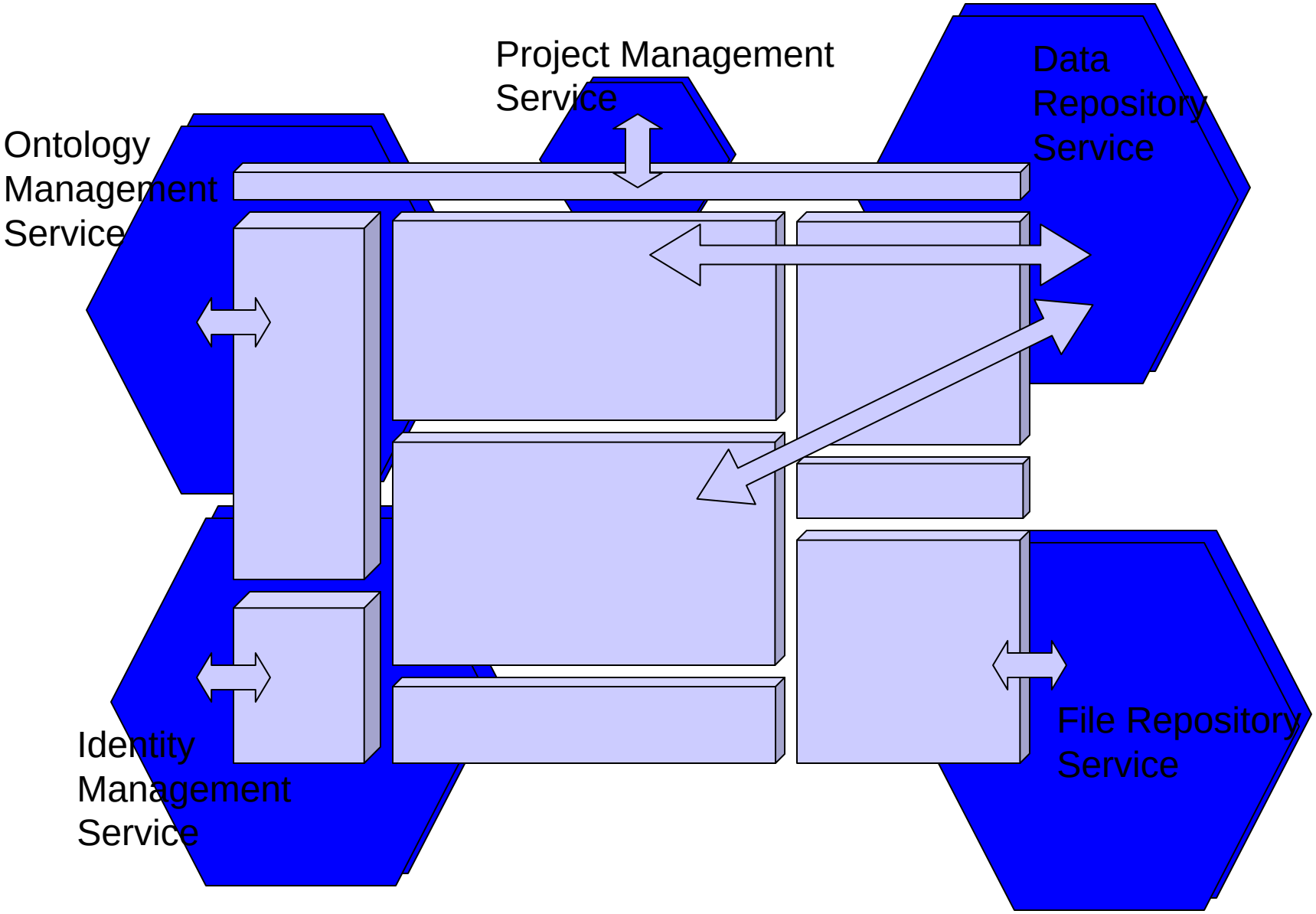
- Explicitly organized and transformed person-oriented clinical data optimized for clinical genomics research
- An architecture that allows different studies to come together seamlessly
- An integration of clinical data, trials data, genotypic data, and knowledge annotation
- A portable and extensible application framework

i2b2 Cell: Canonical Hive Unit



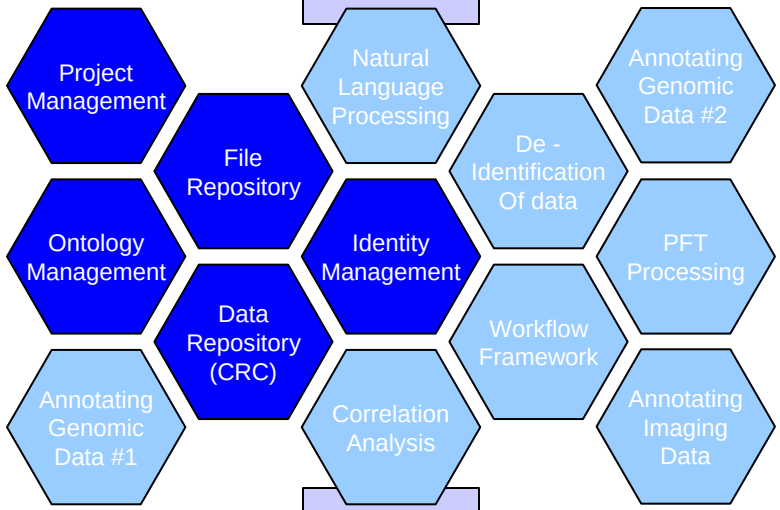
HTTP XML
(minimum: RESTful, others like SOAP optional)

Programmatic Access


Project Management Service
Data Repository Service
Ontology Management Service
Identity Management Service
File Repository Service

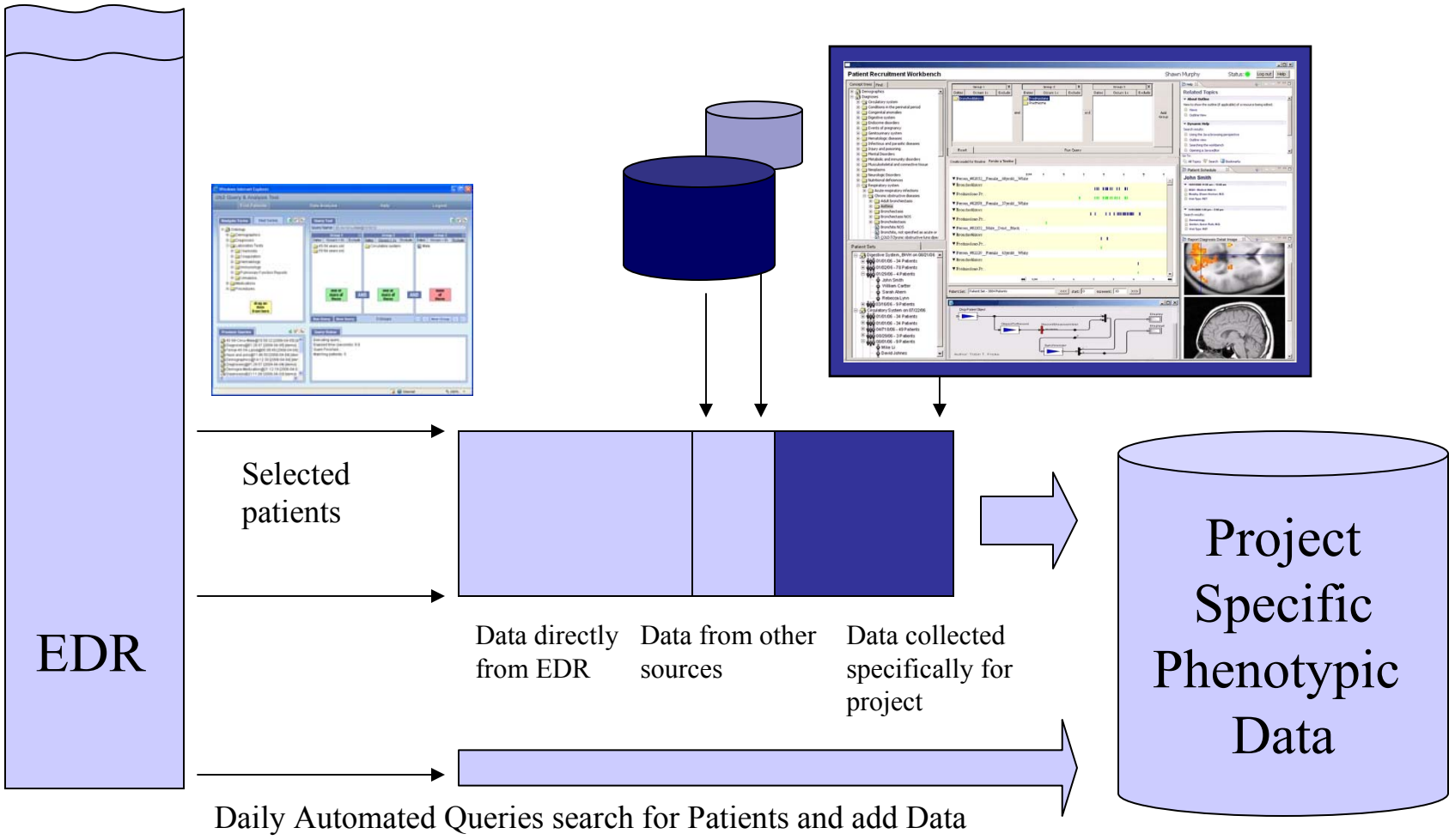
Enterprise-wide repurposing and distribution of medical record data for research

i2b2 Hive

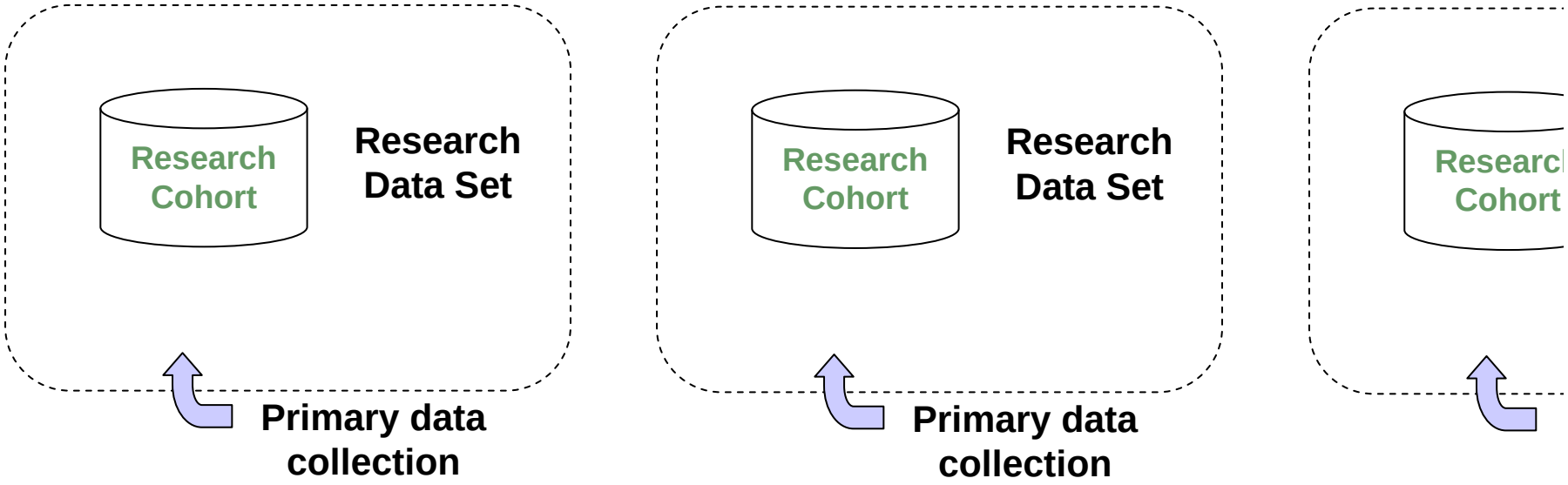


Use of medical record data in clinical studies focused upon genomics and pharmacology

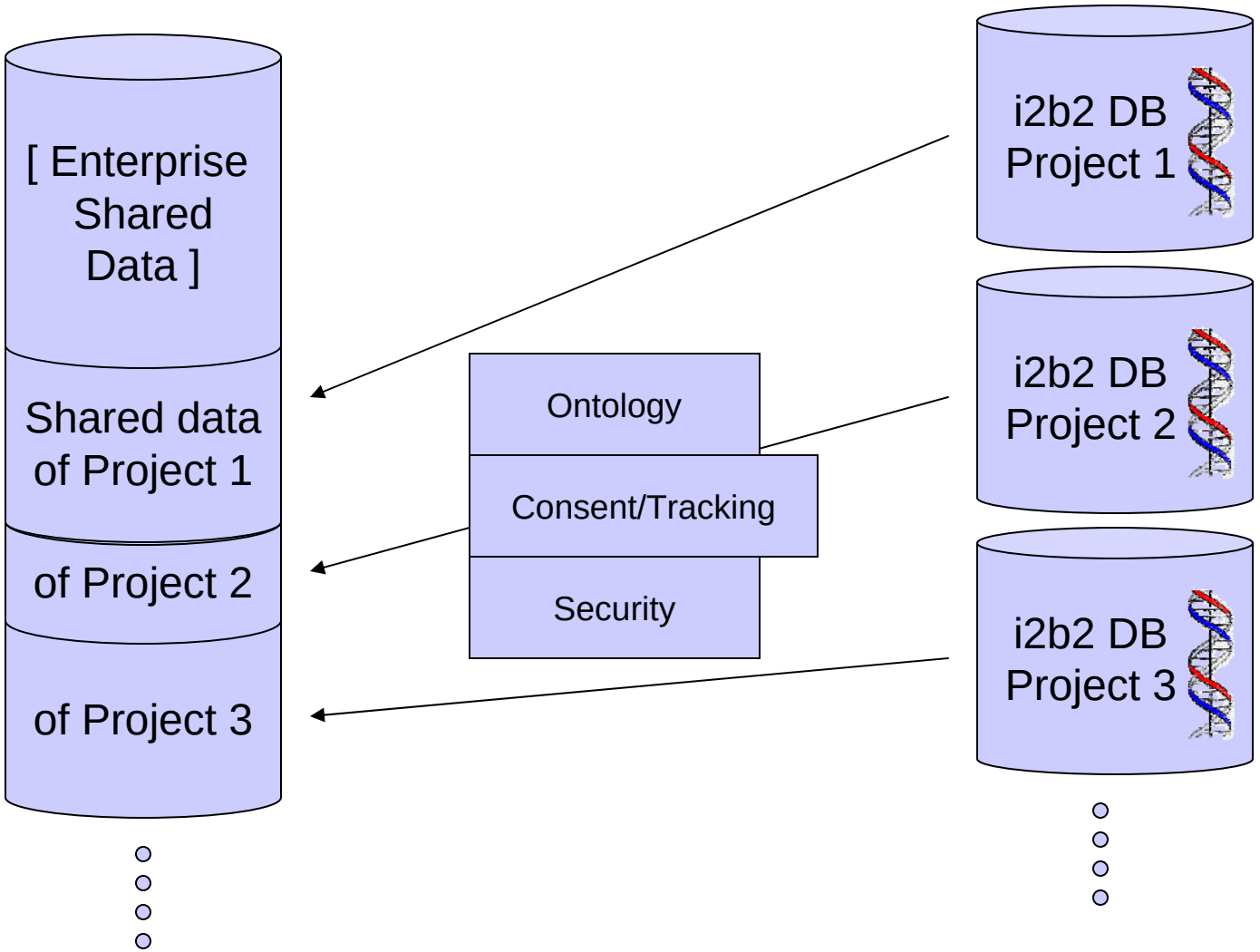
Set of patients is selected through Enterprise Repository and data is gathered into a data mart



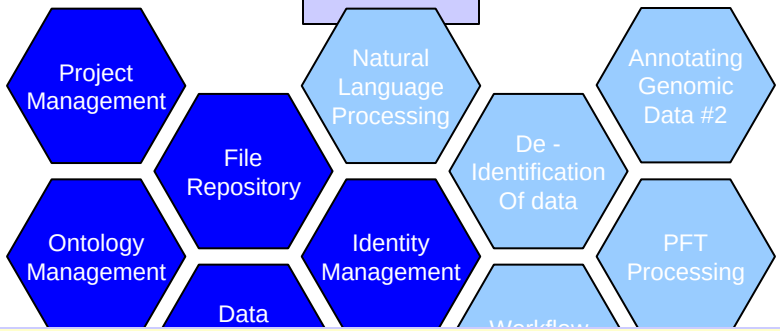
Research Silos



Project data can be added back to the ESD

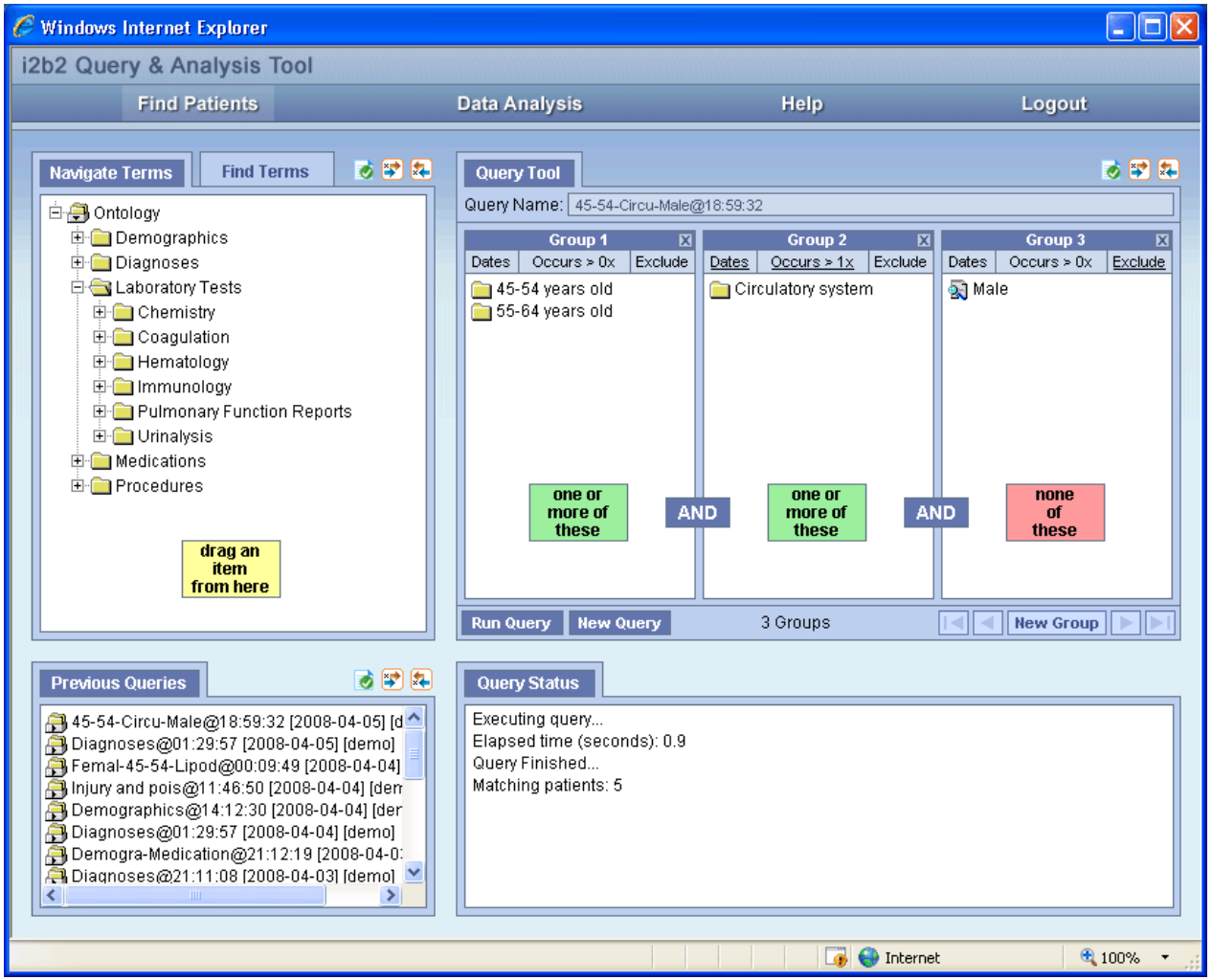


Enterprise-wide repurposing and distribution of medical record data for research

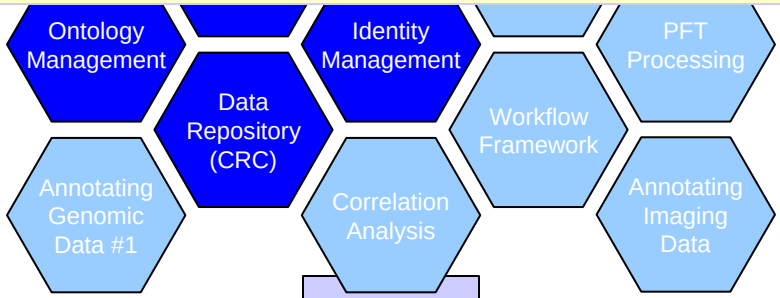


- Enable high performance collection of medical record data for querying and distribution
  - Enterprise web client
- Enable discovery within data on enterprise wide scale
  - Relationship networks
  - Pharmacovigilance

Enterprise web client

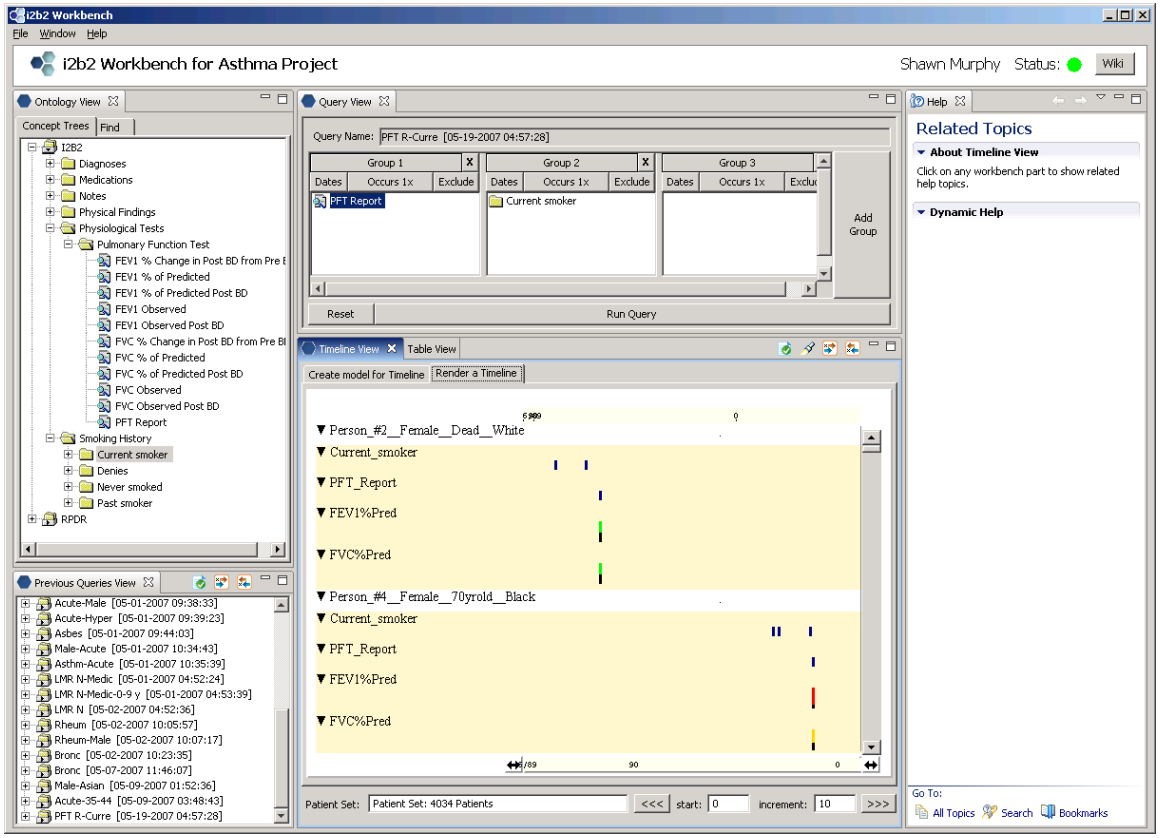


- Repurpose medical record information for research studies
  - I2b2 Workbench
  - Natural language processing
- Enable genomic studies
  - Tissue/blood selection
  - Data integration

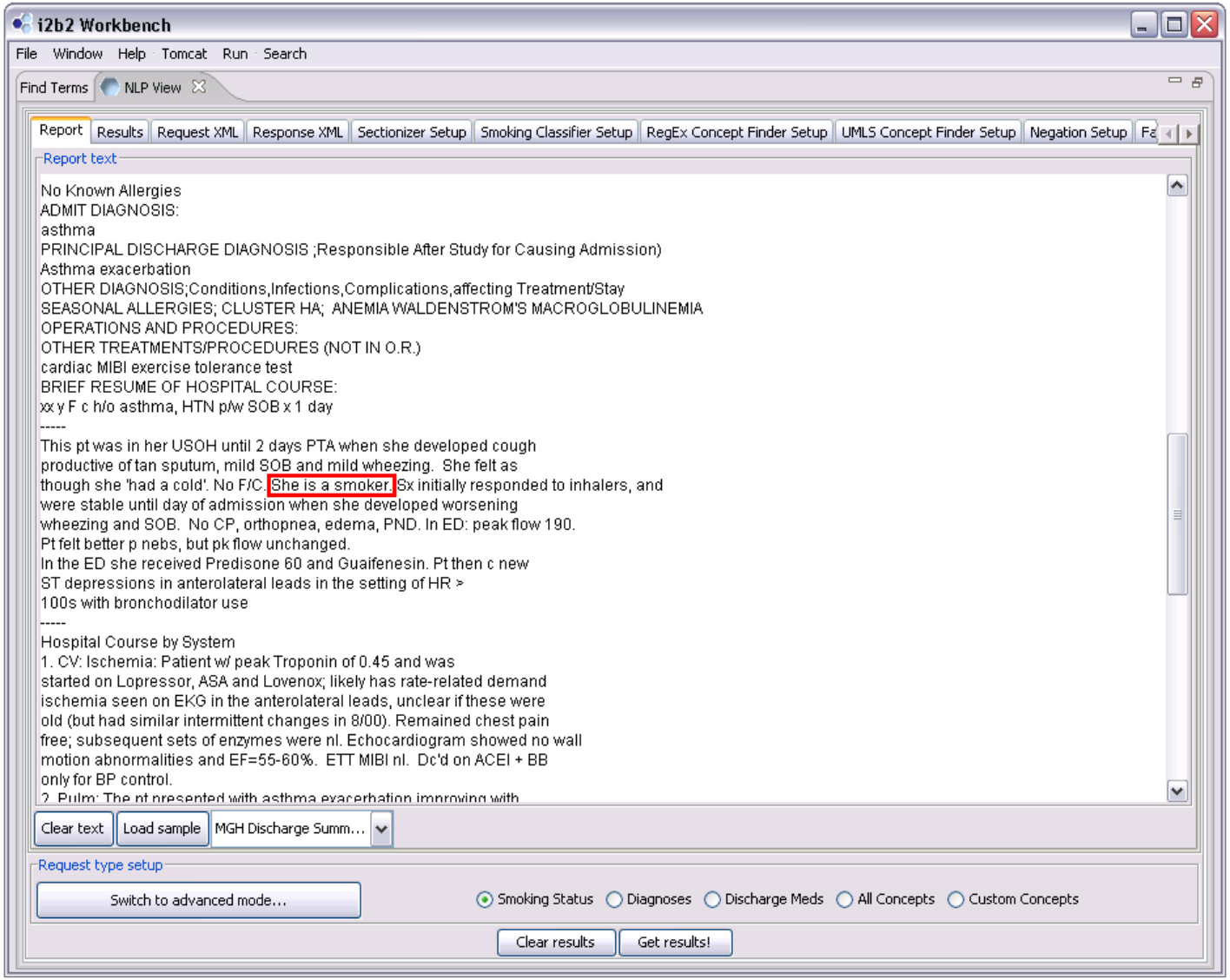


Use of medical record data in clinical studies focused upon genomics and pharmacology

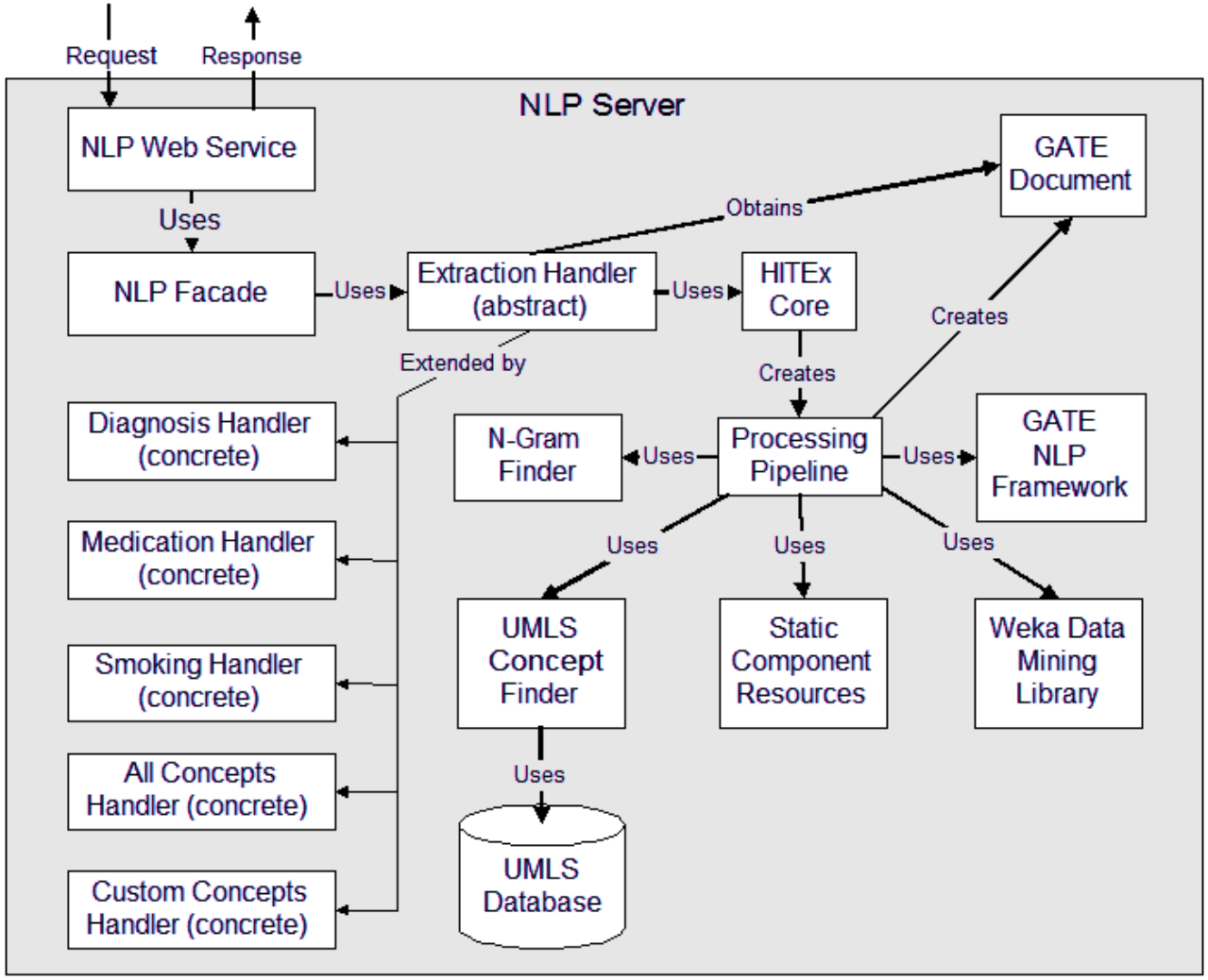
I2b2 Workbench carries hive activity into a detailed patient view for Investigator

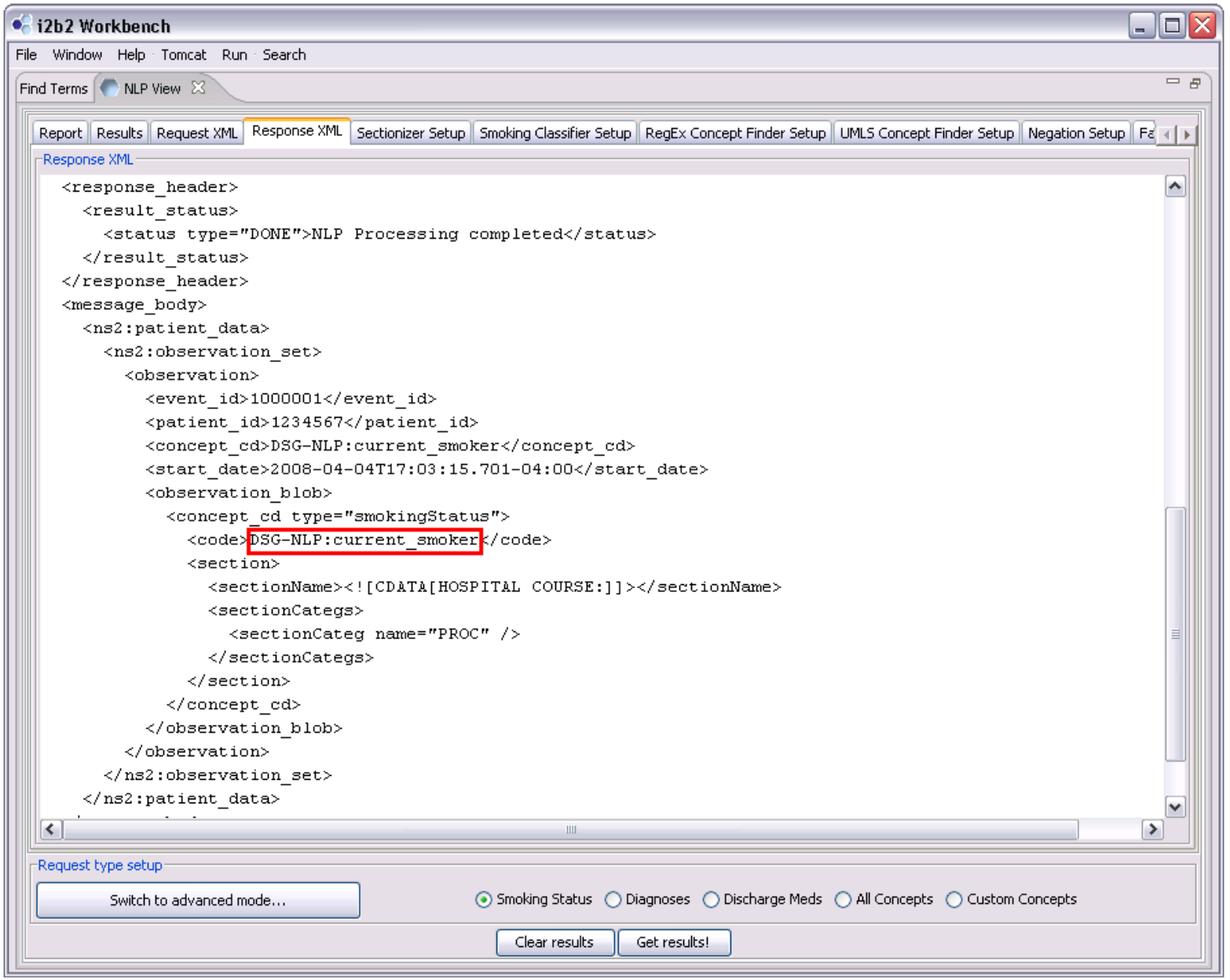


Natural Language Processing Cell

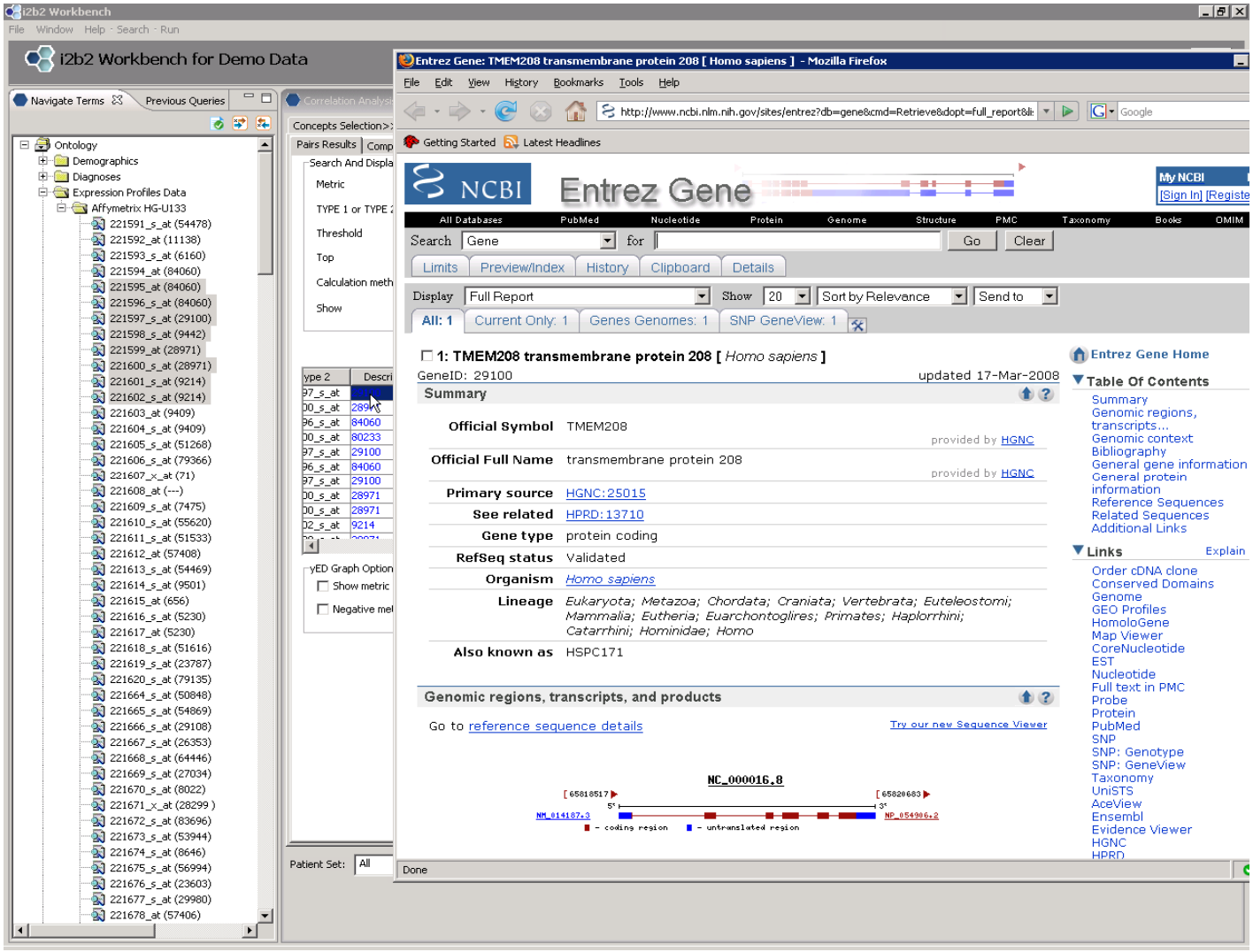


NLP Cell Architecture

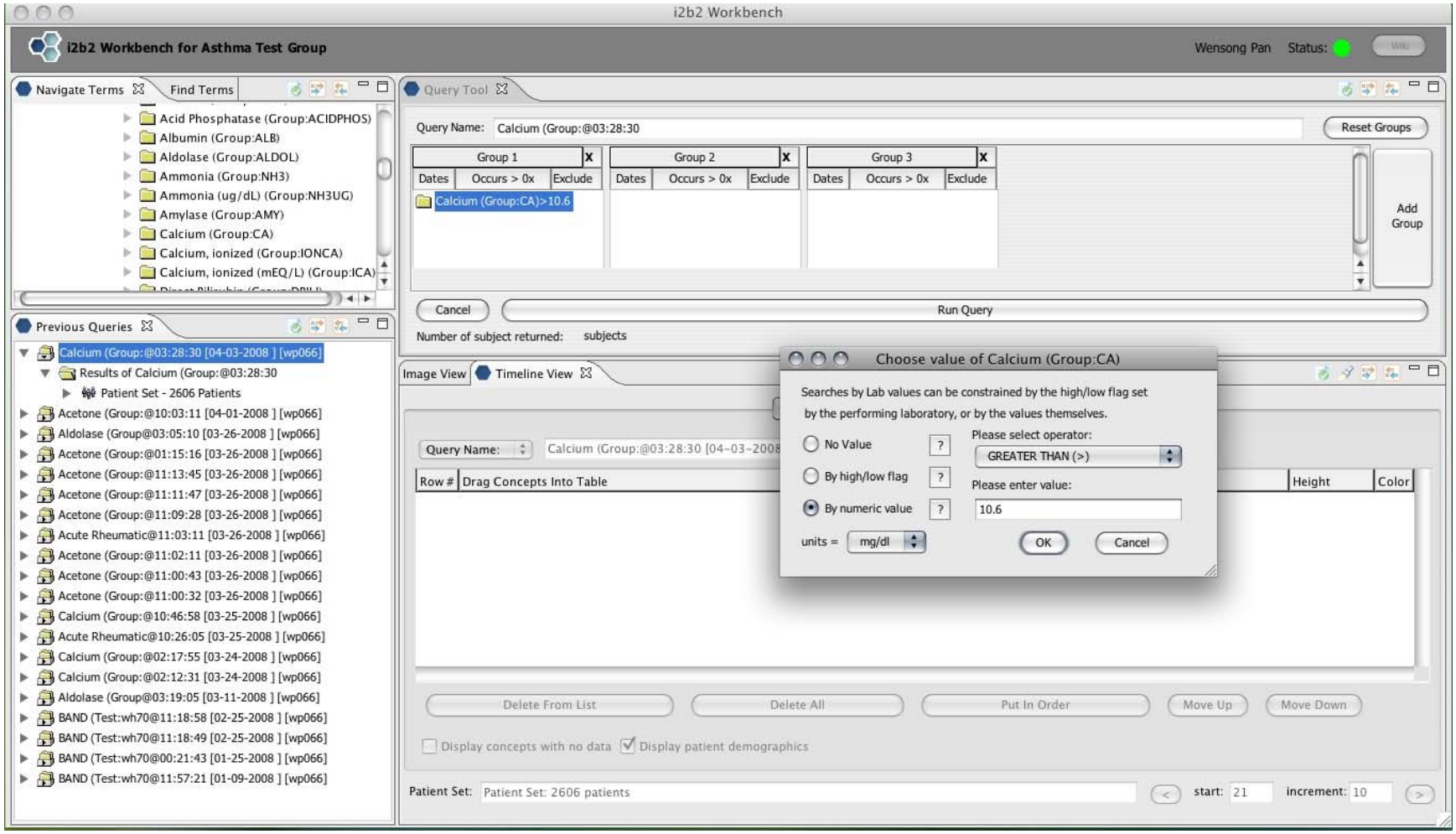



i2b2 Workbench
File Window Help Tomcat Run Search
Find Terms
NLP View
Report
Results
Request XML
Response XML
Sectionizer Setup
Smoking Classifier Setup
RegEx Concept Finder Setup
UMLS Concept Finder Setup
Negation Setup
Response XML
<response_header>
<result_status>
<status type="DONE">NLP Processing completed</status>
</result_status>
</response_header>
<message_body>
<ns2:patient_data>
<ns2:observation_set>
<observation>
<event_id>1000001</event_id>
<patient_id>1234567</patient_id>
<concept_cd>DSG-NLP:current_smoker</concept_cd>
<start_date>2008-04-04T17:03:15.701-04:00</start_date>
<observation_blob>
<concept_cd type="smokingStatus">
<code>DSG-NLP:current_smoker</code>
<section>
<sectionName><![CDATA[HOSPITAL COURSE:]]></sectionName>
<sectionCategs>
<sectionCateg name="PROC" />
</sectionCategs>
</section>
</concept_cd>
</observation_blob>
</observation>
</ns2:observation_set>
</ns2:patient_data>
Request type setup
Switch to advanced mode...
Smoking Status
Diagnoses
Discharge Meds
All Concepts
Custom Concepts
Clear results
Get results!

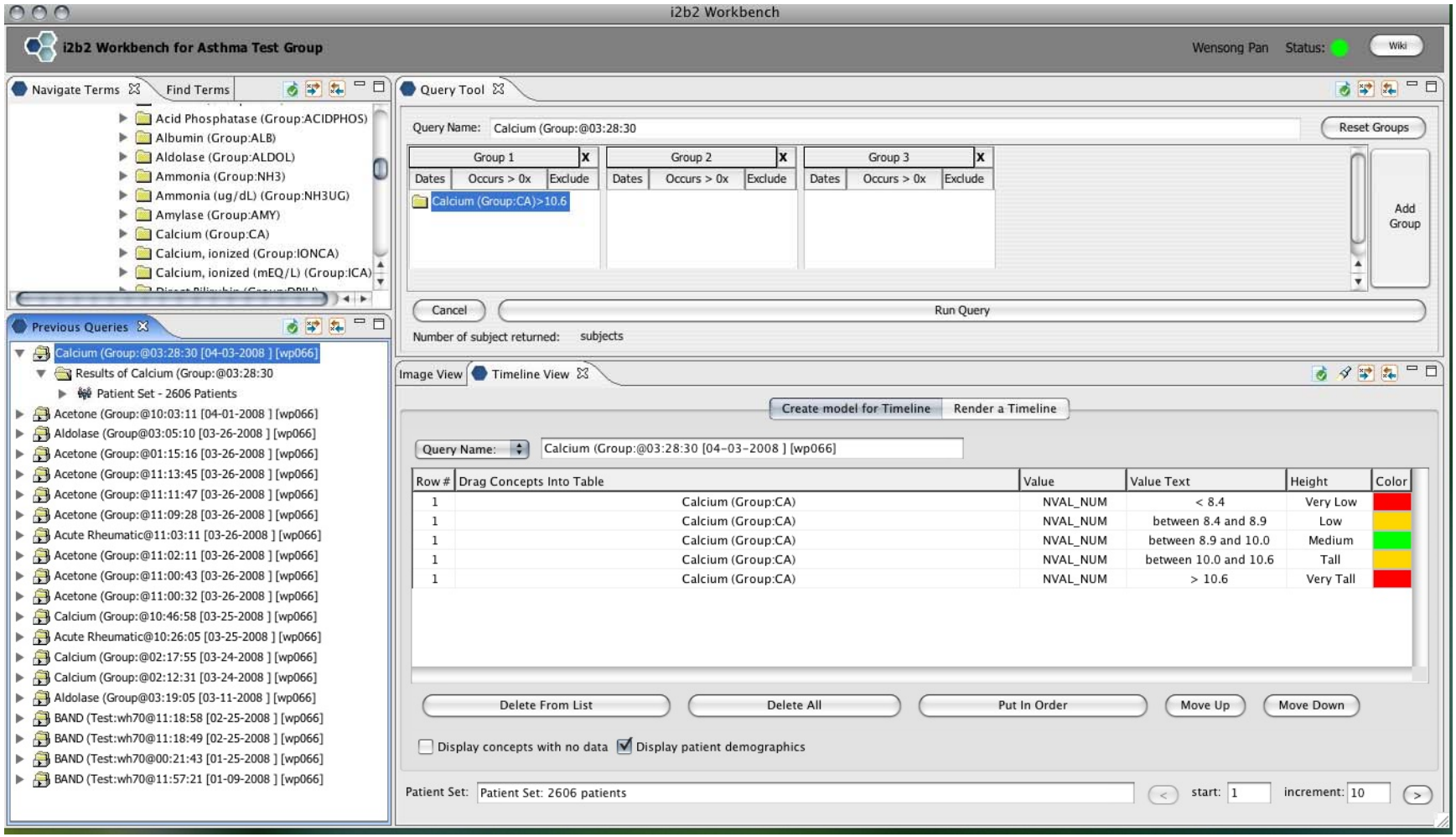
Data integration – Genotype / Phenotype



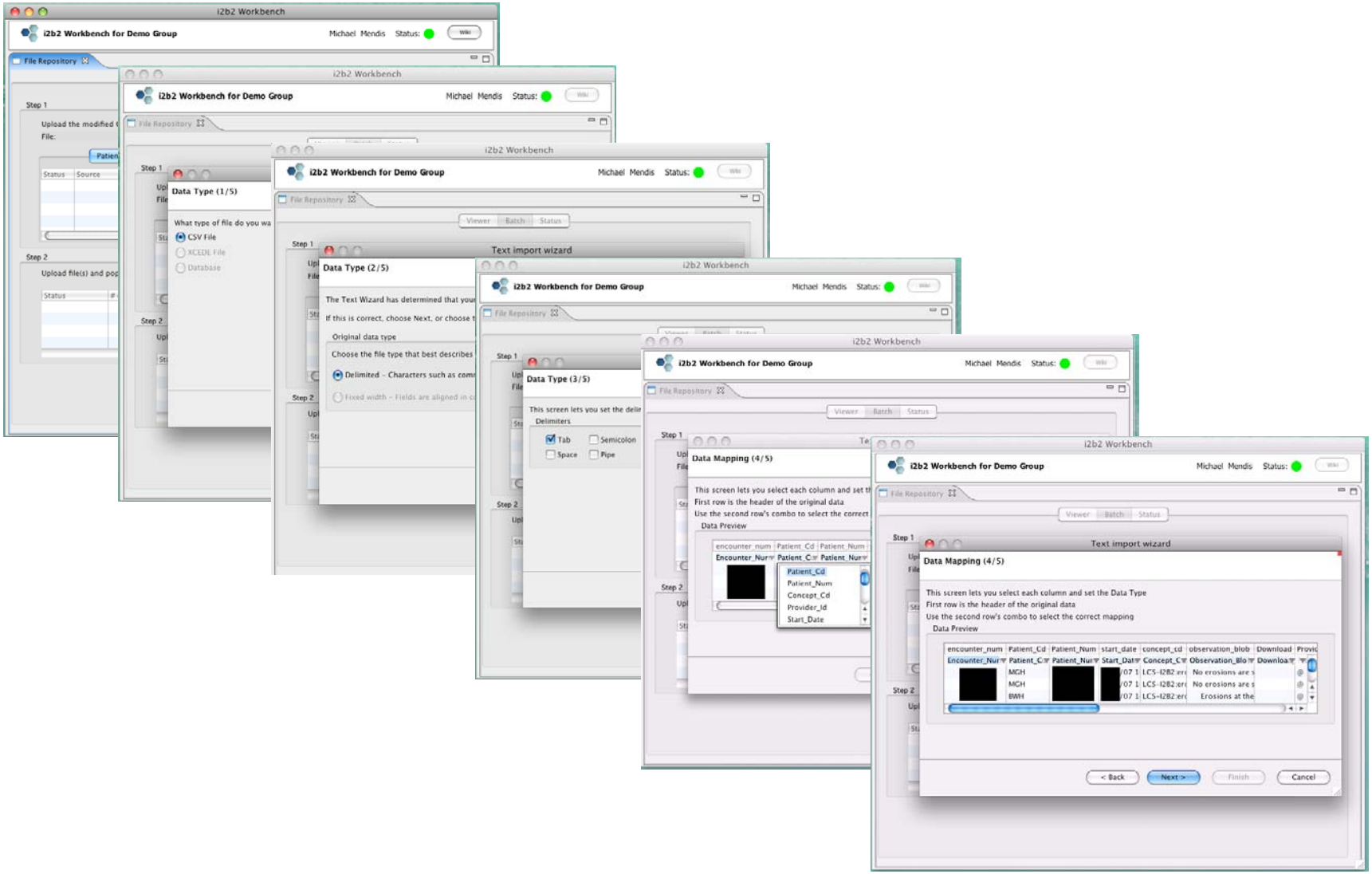
Query by values



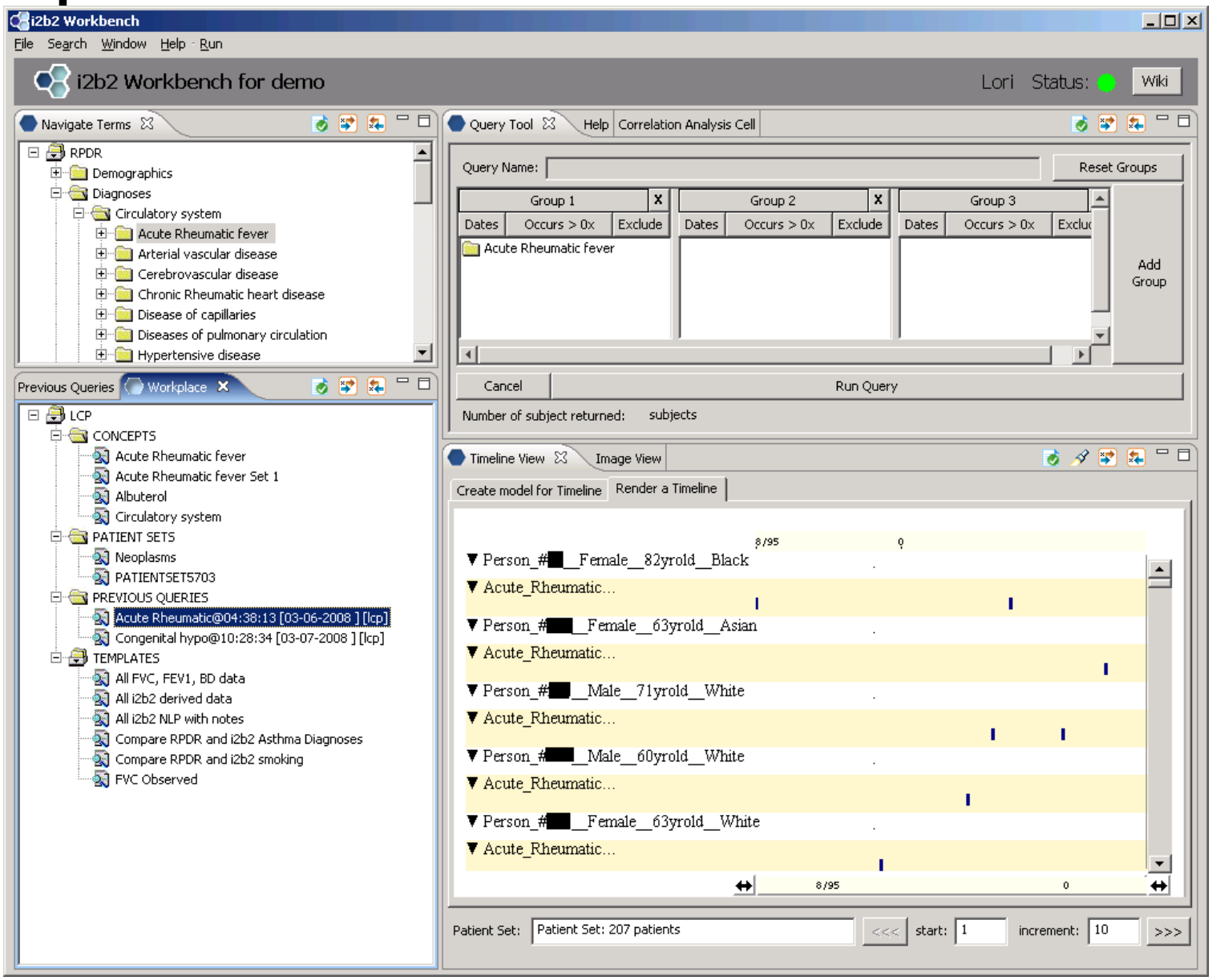
Query by values



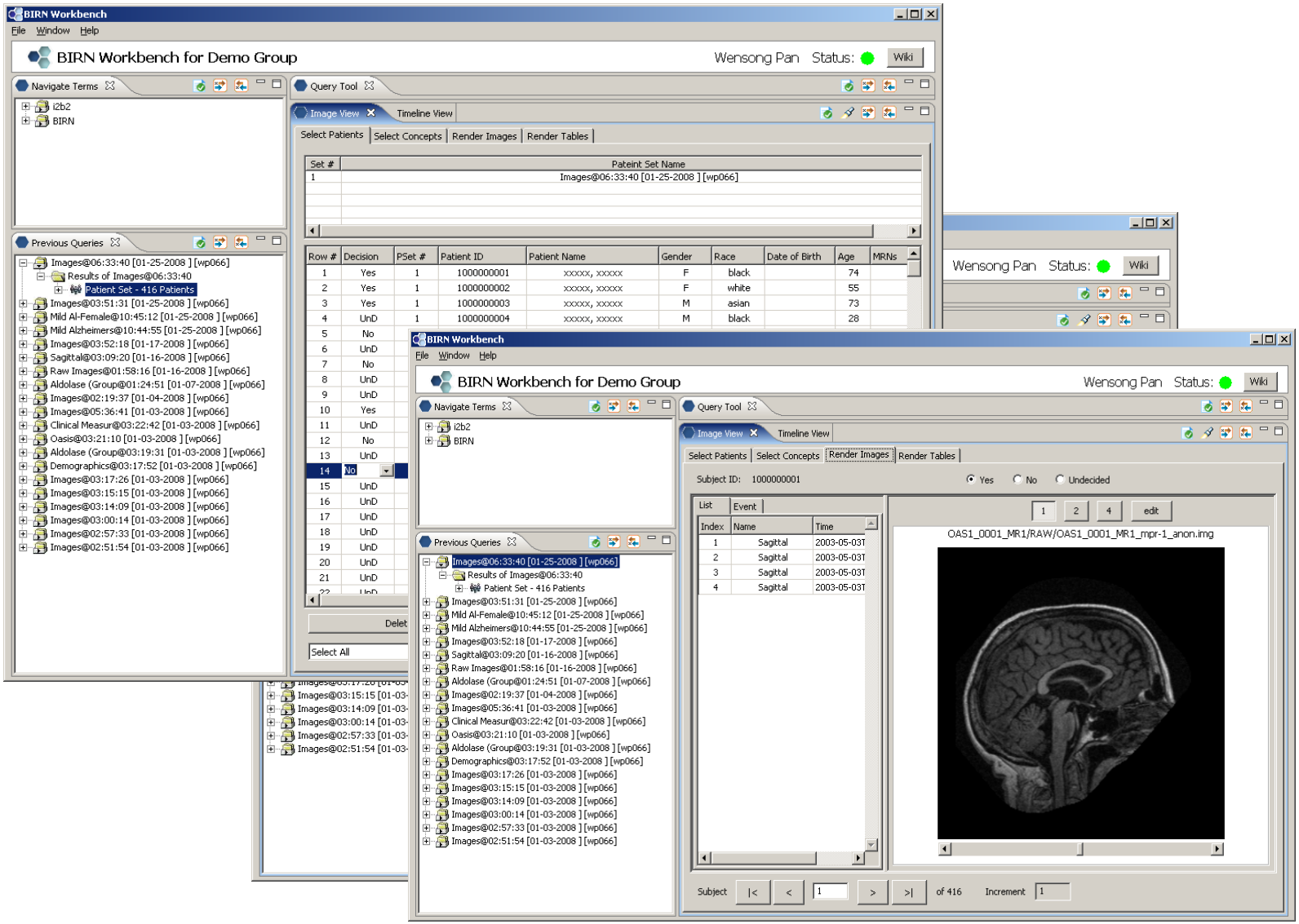
Upload data through i2b2 Workbench



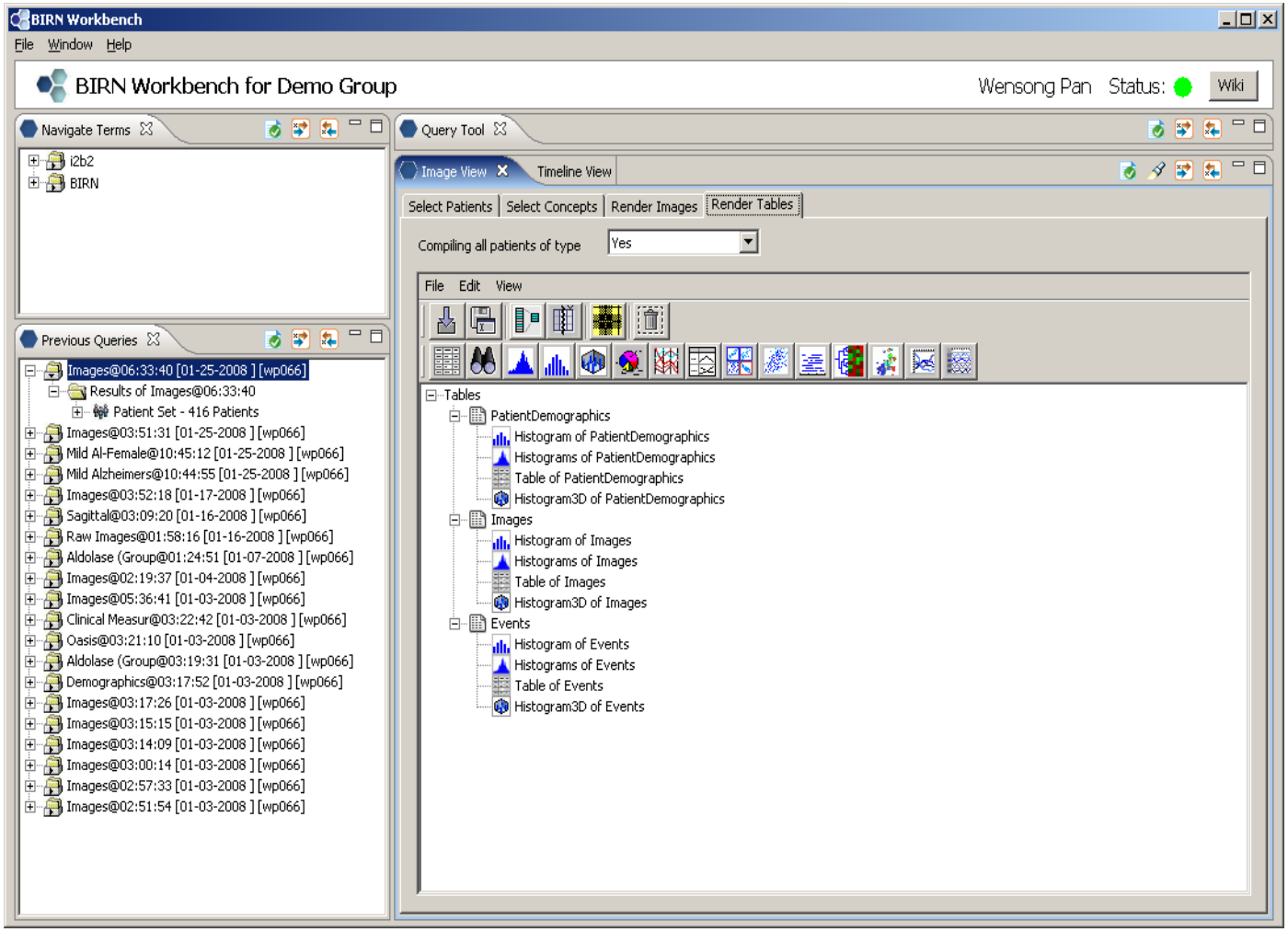
Workplace Interface

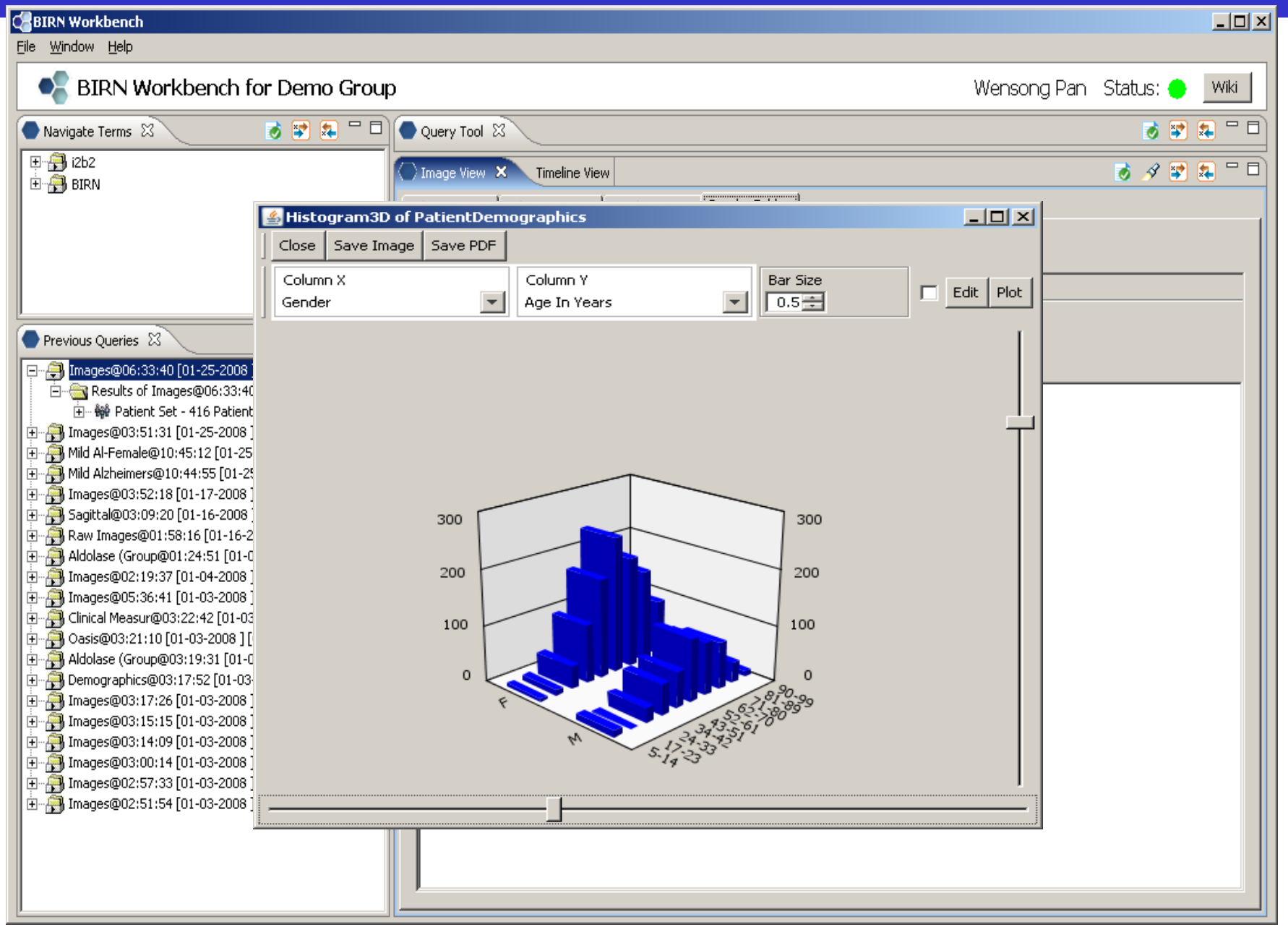


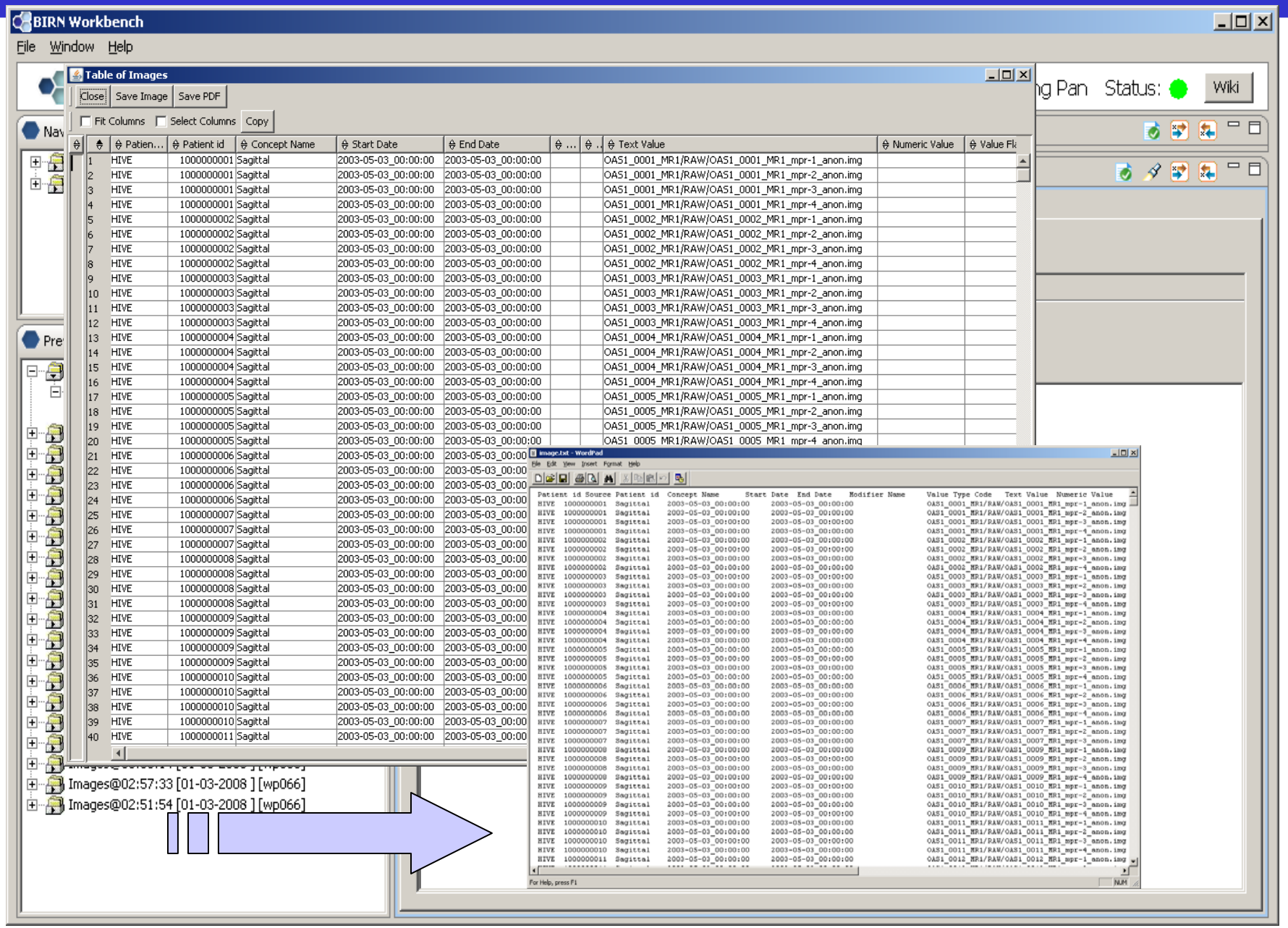
Selecting and Reviewing Patients for Studies

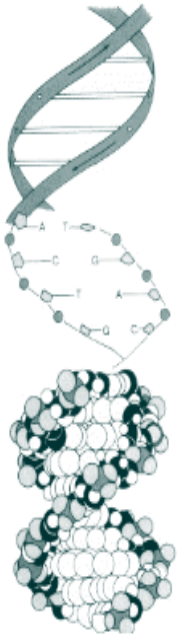


Integration of several data export and analysis tools in i2b2 Workbench




BIRN Workbench
File Window Help
BIRN Workbench for Demo Group
Wensong Pan Status:
Wiki
Navigate Terms
i2b2
BIRN
Query Tool
Image View
Timeline View
Histogram3D of PatientDemographics
Close
Save Image
Save PDF
Column X
Gender
Column Y
Age In Years
Bar Size
0.5
Edit
Plot
300
200
100
0
F
M
Previous Queries
Images@06:33:40 [01-25-2008
Results of Images@06:33:40
Patient Set - 416 Patient
Images@03:51:31 [01-25-2008
Mild Al-Female@10:45:12 [01-25
Mild Alzheimers@10:44:55 [01-25
Images@03:52:18 [01-17-2008
Sagittal@03:09:20 [01-16-2008
Raw Images@01:58:16 [01-16-2
Aldolase (Group@01:24:51 [01-0
Images@02:19:37 [01-04-2008
Images@05:36:41 [01-03-2008
Clinical Measur@03:22:42 [01-03
Oasis@03:21:10 [01-03-2008
Aldolase (Group@03:19:31 [01-0
Demographics@03:17:52 [01-03
Images@03:17:26 [01-03-2008
Images@03:15:15 [01-03-2008
Images@03:14:09 [01-03-2008
Images@03:00:14 [01-03-2008
Images@02:57:33 [01-03-2008
Images@02:51:54 [01-03-2008


BIRN Workbench
File Window Help
Table of Images
Close Save Image Save PDF
Fit Columns Select Columns Copy
Status:
Wiki
Images@02:57:33 [01-03-2008] [wp066]
Images@02:51:54 [01-03-2008] [wp066]



Working Assumptions of i2b2 Data

Shawn Murphy

Vivian Gainer

Data Model: Data Requirements

- **Integration** of data from distributed and differently structured databases in order to perform comprehensive analyses.
- **Separation** of data used for research from daily operational or transactional data.
- **Standardization** of a model across systems.
- **Ease** of use by end-users.

Dimensional Modeling

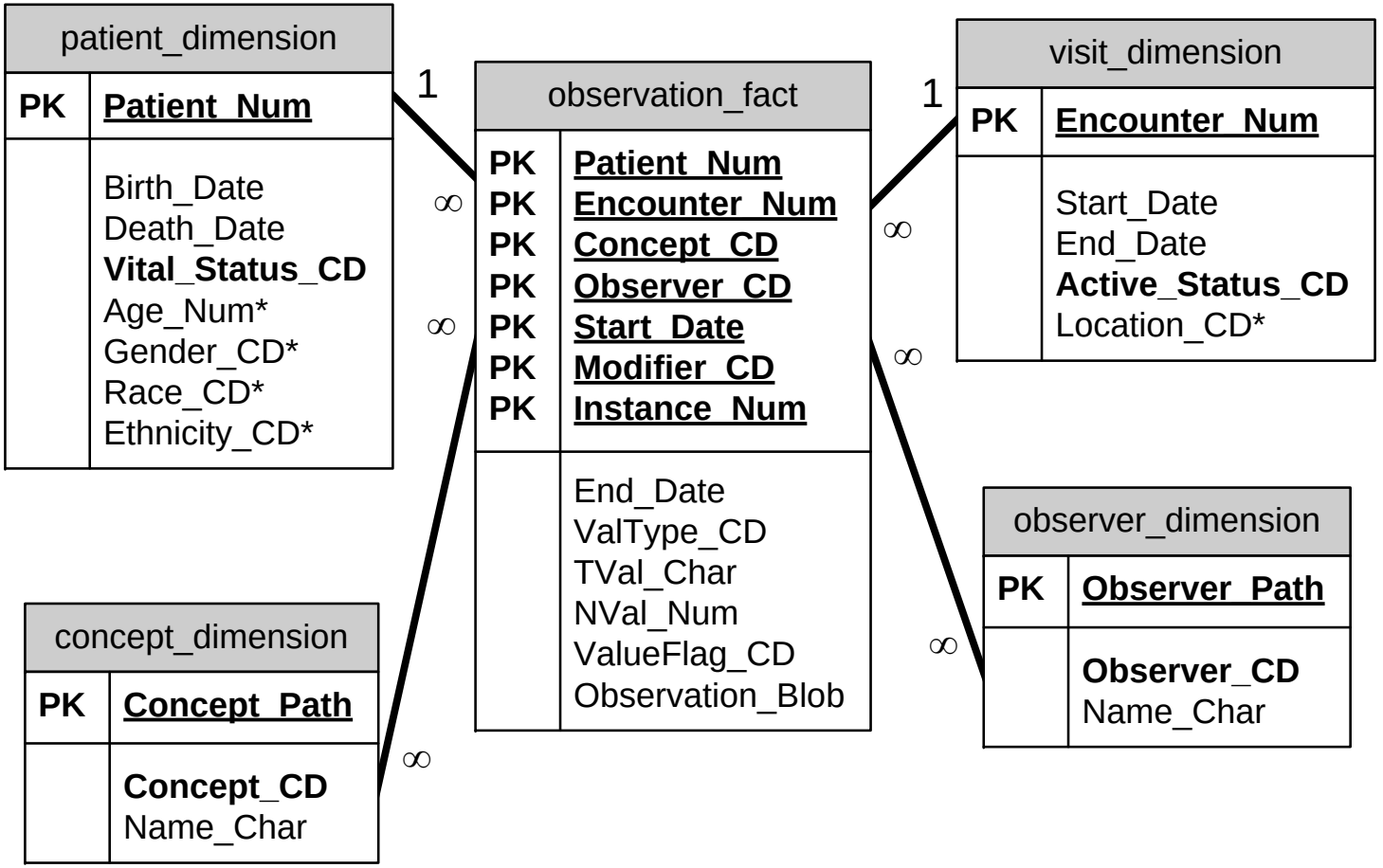
1. **FACTS** - the quantitative or factual data being queried.

2. **DIMENSIONS** – groups of hierarchies and descriptors that define the facts.

Star Schema

One fact table surrounded radially by numerous dimension tables.

i2b2 Star Schema



i2b2 Fact Table

- In i2b2, a fact is an observation on a patient.
- Examples of FACTS:
  - Diagnoses
  - Procedures
  - Health History
  - Genetic Data
  - Lab Data
  - Provider Data
  - Demographics Data
- An observation is not necessarily the same thing as an event

i2b2 Dimension Tables

- Dimension tables contain descriptive information about facts.
- In i2b2 there are four dimension tables

concept_dimension
provider_dimension
visit_dimension
patient_dimension

Indexes

- Very large data warehouses and marts require many indexes for good performance. Use as many indexes as necessary for covering virtually any query
- Consider adding a clustered index (SQL Server) to any table in a data warehouse that needs to produce sorted results.

Values

Valtype_cd	either N for numeric or T for text
Tval_char	if valtype_cd = ‘T’, then the text value goes here.
	if valtype_cd = ‘N’, then tval_char can be ‘E’ for equals, G for greater than, L for less than
Nval_num	if valtype_cd = ‘N’, then the text value goes here
Valueflag_cd	Flag (for high or low values, for example)

Example: Lab Test Values

```
select o.concept_cd, name_char, valtype_cd, tval_char,
nval_num, valueflag_cd, units_cd
from observation_fact o join concept_dimension  c
on o.concept_cd = c.concept_cd
where valtype_cd = 'N‘
```

concept_...	name_char	valtype_cd	tval_char	nval_num	valueflag_cd	units_cd
BC1-20	Alt/gpt (Test:bc1-20)	N	E	6.00000	L	u/l
BC1-21	Ast/got (Test:bc1-21)	N	E	16.00000	@	u/l
BC1-24	Alk phos (Test:bc1-24)	N	E	107.00000	@	u/l
BC1-39	Albumin (Test:bc1-39)	N	E	4.20000	@	g/dl
BC1-7	Creatinine (Test:bc1-7)	N	E	0.70000	@	mg/dl
BC1-10	Chloride (Test:bc1-10)	N	E	106.00000	@	mmol/l
BC1-106	B12 (Test:bc1-106)	N	E	409.00000	@	pg/ml
BC1-11	Total co2 (Test:bc1-11)	N	E	26.00000	@	mmol/l
BC1-110	Ferritin (Test:bc1-110)	N	E	42.00000	@	ug/l
BC1-136	Vldl (Test:bc1-136)	N	E	12.00000	@	mg/dl
BC1-19	Anion gap (Test:bc1-19)	N	E	10.00000	@	mmol/l
BC1-20	Alt/gpt (Test:bc1-20)	N	E	9.00000	@	u/l

Relationship of Metadata to Star Schema

- Star Schema contains one fact and many dimension tables.
- Concepts in these tables are defined in a separate metadata table or tables.
- The structure of the metadata is integral to the visualization of concepts as well as for querying the data.
- All metadata tables have the same basic structure.

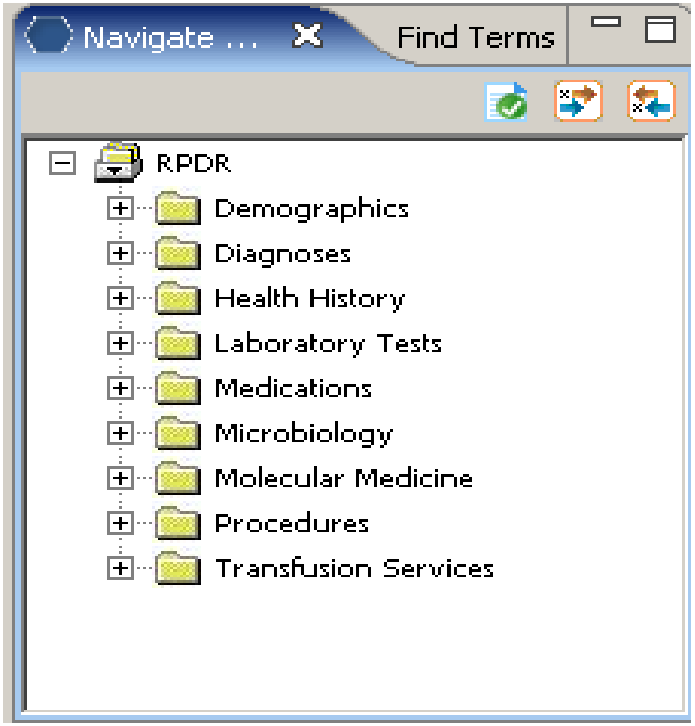
Typical i2b2 Metadata Categories

- Diagnoses
- Procedures
- Demographics
- Lab Tests
- Encounters (visits or observations)
- Providers (observers)
- Health History (physical findings and vital signs)
- Transfusion
- Microbiology

Structure of Metadata Table

METADATA	
C_HLEVEL	INT NULL
C_FULLNAME	VARCHAR(900) NULL
C_NAME	VARCHAR(2000) NULL
C_SYNONYM_CD	CHAR(1) NULL
C_VISUALATTRIBUTES	CHAR(3) NULL
C_TOTALNUM	INT NULL
C_BASECODE	VARCHAR(450) NULL
C_METADATAXML	TEXT NULL
C_FACTTABLECOLUMN	VARCHAR(50) NULL
C_TABLENAME	VARCHAR(50) NULL
C_COLUMNNAME	VARCHAR(50) NULL
C_COLUMNDATATYPE	VARCHAR(50) NULL
C_OPERATOR	VARCHAR(10) NULL
C_DIMCODE	VARCHAR(900) NULL
C_COMMENT	TEXT NULL
C_TOOLTIP	VARCHAR(900) NULL
UPDATE_DATE	DATETIME NULL
DOWNLOAD_DATE	DATETIME NULL
IMPORT_DATE	DATETIME NULL
SOURCESYSTEM_CD	VARCHAR(50) NULL
VALUETYPE_CD	VARCHAR(50) NULL

c_hlevel = 1

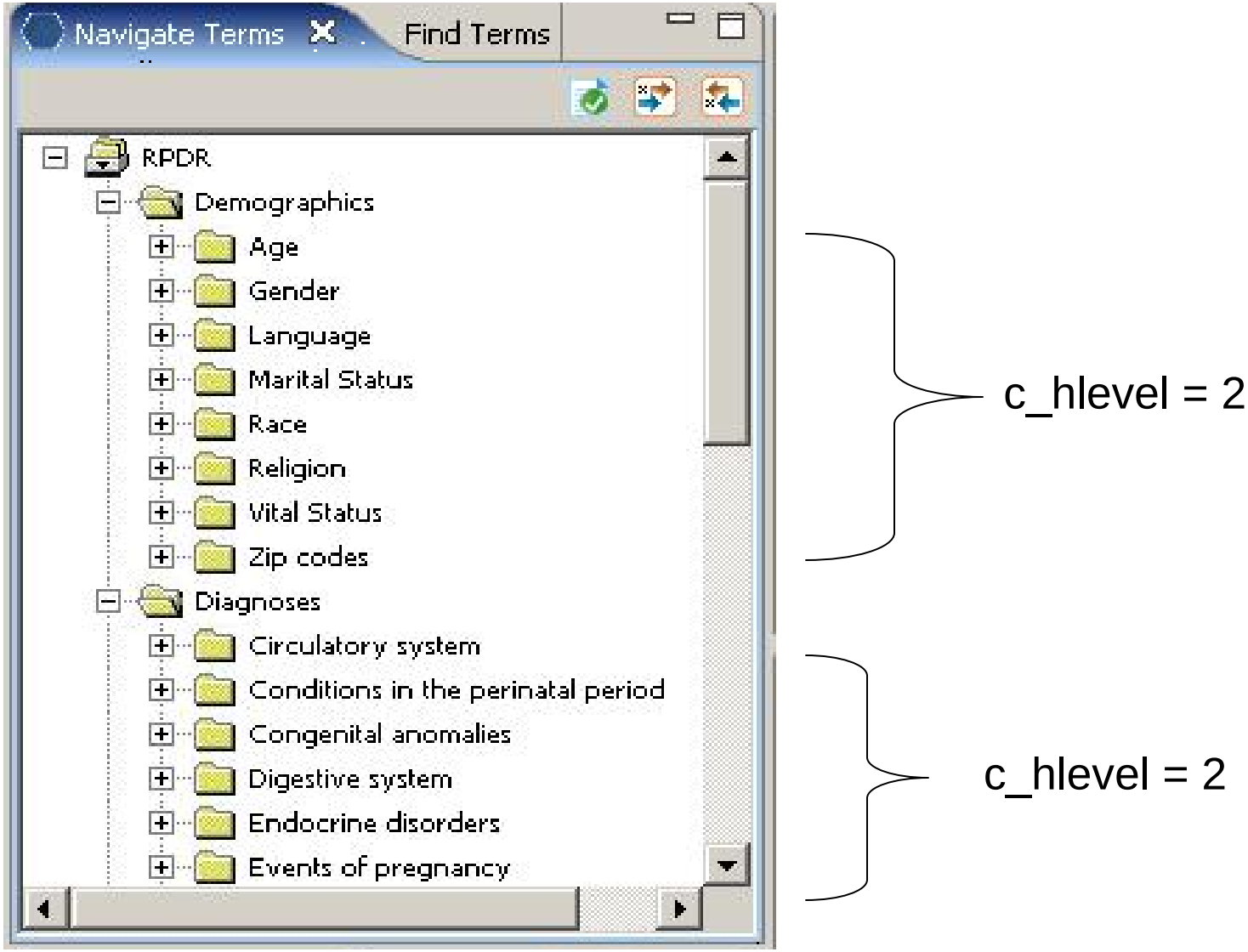



```
select c_hlevel, c_fullname,
c_name,c_visualattributes
from testrpdr
where c_hlevel=1
```

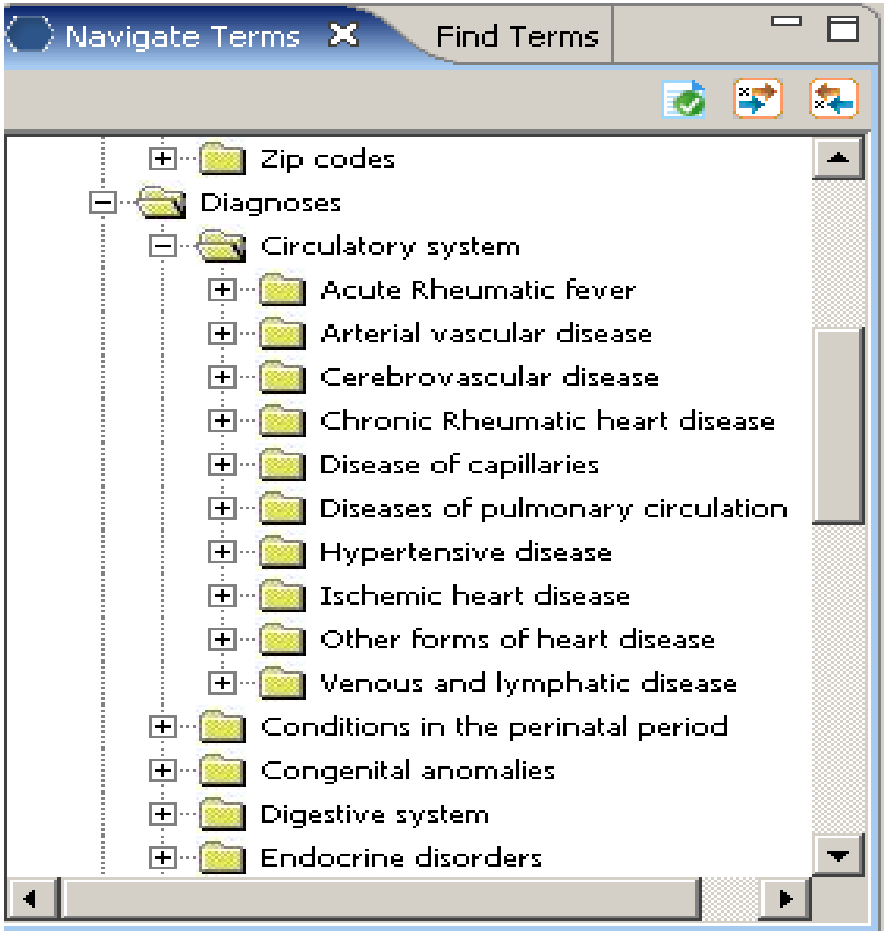
Results | Messages

	c_hlevel	c_fullname	c_name	c_visualattributes
1	1	\RPDR\Microbiology	Microbiology	FA
2	1	\RPDR\Procedures	Procedures	FA
3	1	\RPDR\Labtests	Laboratory Tests	FA
4	1	\RPDR\Medications	Medications	FA
5	1	\RPDR\Transfusions	Transfusion Services	FA
6	1	\RPDR\HealthHistory	Health History	FA
7	1	\RPDR\HPCGG	Molecular Medicine	FA
8	1	\RPDR\Diagnoses	Diagnoses	FA
9	1	\RPDR\Demographics	Demographics	FA

c_hlevel = 2


Navigate Terms
Find Terms
RPDR
Demographics
Age
Gender
Language
Marital Status
Race
Religion
Vital Status
Zip codes
Diagnoses
Circulatory system
Conditions in the perinatal period
Congenital anomalies
Digestive system
Endocrine disorders
Events of pregnancy
c_hlevel = 2
c_hlevel = 2

c_hlevel = 3



c_hlevel=3

c_fullname and c_name

c_fullname is the hierarchical path that leads to the term

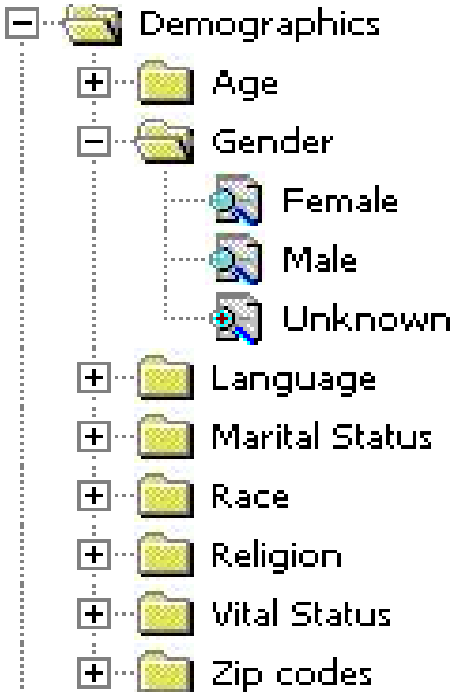
\RPDR
\Diagnoses
\Musculoskeletal and connective tissue (710-739)
\Arthropathies (710-719)
\(714) Rheumatoid arthritis and other arthropathies
\(714-0) Rheumatoid arthritis

c_name is the actual term

Rheumatoid arthritis
Atrophic arthritis
RA [Rheumatoid arthritis]
Chronic rheumatic arthritis

c_fullname	c_name
\RPDR\Diagnoses\Musculoskeletal and connective tissue (710-739)\Arthropathies (710-719)\(714) Rheumatoid arthritis and ot~\(714-0) Rheumatoid arthritis	Rheumatoid arthritis
\RPDR\Diagnoses\Musculoskeletal and connective tissue (710-739)\Arthropathies (710-719)\(714) Rheumatoid arthritis and ot~\(714-0) Rheumatoid arthritis	Atrophic arthritis
\RPDR\Diagnoses\Musculoskeletal and connective tissue (710-739)\Arthropathies (710-719)\(714) Rheumatoid arthritis and ot~\(714-0) Rheumatoid arthritis	RA [Rheumatoid arthritis]
\RPDR\Diagnoses\Musculoskeletal and connective tissue (710-739)\Arthropathies (710-719)\(714) Rheumatoid arthritis and ot~\(714-0) Rheumatoid arthritis	Chronic rheumatic arthritis

c_visualattributes



1st character :

F = Folder

C = Container

M = Multiple

2nd character:

A = Active

I = Inactive

H = Hidden

C_HLEVEL	C_NAME	C_SYNONYM_CD	C_VISUALATTRIBUTES	C_BASECODE	
5	Other specified inflammatory polyarthropathies	N	FA	7148	(non-specific code)
5	Other specified inflammatory polyarthropathies	Y	FH	7148	
6	Other specified inflammatory polyarthropathies	N	LA	71489	(specific code)

c_basecode

The basecode is the coded value for the term.

c_hlevel	c_fullname	c_name	c_visualattributes	c_basecode	c_operator
2	\RPDR\Demographics\Gender	Gender	FA	NULL	LIKE
3	\RPDR\Demographics\Gender\Female	Female	LA	DEM|SEX:f	LIKE
3	\RPDR\Demographics\Gender\Male	Male	LA	DEM|SEX:m	LIKE
3	\RPDR\Demographics\Gender\Unknown	Unknown	MA	NULL	LIKE
4	\RPDR\Demographics\Gender\Unknown\Unknown-@	Unknown-@	LH	DEM|SEX:@	LIKE
4	\RPDR\Demographics\Gender\Unknown\Unknown-U	Unknown-U	LH	DEM|SEX:u	LIKE

It maps to the concept_cd in the star schema tables.

c_facttablecolumn,c_tablename,c_columnname,
c_columndatatype,c_operator,c_dimcode

METADATA	
C_HLEVEL	INT NULL
C_FULLNAME	VARCHAR(900) NULL
C_NAME	VARCHAR(2000) NULL
C_SYNONYM_CD	CHAR(1) NULL
C_VISUALATTRIBUTES	CHAR(3) NULL
C_TOTALNUM	INT NULL
C_BASECODE	VARCHAR(450) NULL
C_METADATAXML	TEXT NULL
C_FACTTABLECOLUMN	VARCHAR(50) NULL
C_TABLENAME	VARCHAR(50) NULL
C_COLUMNNAME	VARCHAR(50) NULL
C_COLUMNDATATYPE	VARCHAR(50) NULL
C_OPERATOR	VARCHAR(10) NULL
C_DIMCODE	VARCHAR(900) NULL
C_COMMENT	TEXT NULL
C_TOOLTIP	VARCHAR(900) NULL
UPDATE_DATE	DATETIME NULL
DOWNLOAD_DATE	DATETIME NULL
IMPORT_DATE	DATETIME NULL
SOURCESYSTEM_CD	VARCHAR(50) NULL
VALUETYPE_CD	VARCHAR(50) NULL

Fields used to construct queries

c_facttablecolumn	c_tablename	c_columnname	c_columndatatype	c_operator	c_dimcode
concept_cd	concept_dimension	concept_path	T	LIKE	\RPDR\Demographics\Gender
concept_cd	concept_dimension	concept_path	T	LIKE	\RPDR\Demographics\Gender\Female
concept_cd	concept_dimension	concept_path	T	LIKE	\RPDR\Demographics\Gender\Male
concept_cd	concept_dimension	concept_path	T	LIKE	\RPDR\Demographics\Gender\Unknown
concept_cd	concept_dimension	concept_path	T	LIKE	\RPDR\Demographics\Gender\Unknown\Unknown-@
concept_cd	concept_dimension	concept_path	T	LIKE	\RPDR\Demographics\Gender\Unknown\Unknown-U

Select * from observation_fact where c_facttablecolumn in
(select concept_cd from c_tablename where c_columname c_operator ‘c_dimcode%’)

Select * from observation_fact where concept_cd in
(select concept_cd from concept_dimension where concept_path like
’\RPDR\Demographics\Gender%’)

c_metadataxml

stores values and information about the concept such as high and low indicators

Navigating concept_dimension

Except for diagnoses, which do not have a standard prefix, the prefix of the location_path identifies the type of concept:

LAB\ is for Laboratory data
TRS\ is for Transfusion data
MIC\ is for Microbiology data
MCP\ is for CPT medications
MUL\ is for hospital medications
PRC\ is for procedures

So, if looking for a list of all possible labs, the query to run is

select * from concept_dimension where concept_path like 'lab\%'

concept_cd can be joined to concept_cd in the Observation_Fact table.

Visit_dimension

- Encounter number
- In/out code
- Hospital of service
- Clinic

Encounter_num can be joined to encounter_num in the Observation_Fact table.

Patient_num can be joined to patient_num in the Observation_Fact and Visit_Dimension tables.

Example 1: Query based on a diagnosis

To find all the patients diagnosed with migraines, use this query:

```
Select p.patient_num
From patient_dimension p join observation_fact f
On p.patient_num = f.patient_num
And f.concept_cd in
(select concept_cd
from concept_dimension
where concept_path like
'Neurologic Disorders (320-389)\(346) Migraine\%')
```

Add Demographics

To find the ages of all patients diagnosed with migraines, use this query:

```
Select p.age_in_years_num
From patient_dimension p, observation_fact f
Where p.patient_num = f.patient_num
And f.concept_cd in
(select concept_cd
from concept_dimension
where concept_path like
'Neurologic Disorders (320-389)\(346) Migraine\%')
```

Add Visit data

To limit the query to only those patients seen at MGH, add the Visit table.

```
Select p.patient_num,p.age_in_years_num
From patient_dimension p join observation_fact f
on p.patient_num = f.patient_num join visit_dimension v
on v.encounter_num = f.encounter_num
And v.location_cd = 'MGH'
And f.concept_cd in
(select concept_cd
from concept_dimension
where concept_path like
'Neurologic Disorders (320-389)\(346) Migraine\%')
```

Multiple paths

```
select distinct(patient_num)  into BoneMarrowTransplants from observation_fact where
concept_cd in
(Select concept_cd
From concept_dimension
Where concept_path LIKE
'PRC\ICD9 (Inpatient)\(40-41) Operations on hemic and lymphatic system\(p41) Operations on bone marrow a~\(p41-0) Bone
    marrow or hematopoie~\%'
or concept_path LIKE 'PRC\CPT\(10021-69990) Surgery\(38100-38999) Hemic and Lymphatic Systems\(38204-38242) Bone
    Marrow or Stem Cell\(38242) Bone marrow or blood-deri~\%'
or concept_path LIKE 'PRC\CPT\(10021-69990) Surgery\(38100-38999) Hemic and Lymphatic Systems\(38204-38242) Bone
    Marrow or Stem Cell\(38240) Bone marrow or blood-deri~\%'
or concept_path LIKE 'PRC\CPT\(10021-69990) Surgery\(38100-38999) Hemic and Lymphatic Systems\(38204-38242) Bone
    Marrow or Stem Cell\(38241) Bone marrow or blood-deri~\%'
or concept_path LIKE '(Pre) Transplants and Tracheostomy\Surgical\(481) Bone Marrow Transplant\%'
or concept_path LIKE 'zz V-codes\Conditions influencing health status (V40-V49)\(V42) Organ or tissue replaced by~\(V42-8) Other
    specified organ or ~\(V42-81) Bone marrow replaced by ~\%'
or concept_path LIKE 'PRC\LMR\(LPA547) bone marrow transplant\%'
or concept_path LIKE 'Injury and poisoning (800-999)\Complications of medical care (996-999)\(996) Complications peculiar to
    c~\(996-8) Complications of transpla~\(996-85) Complications of bone ma~\%')
```

Find CCP lab codes

```
select c_hlevel,c_fullname,c_name,c_basecode into CCPCodes
from labtests where c_fullname like '%ACCP%'
or c_fullname like '%ANTCCP%'
```

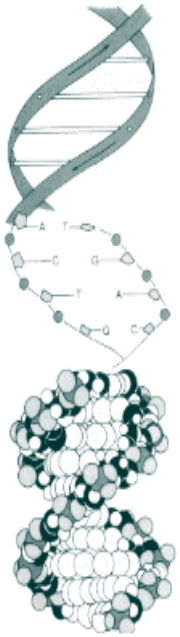
c_hlevel	c_fullname	c_name	c_basecode
3	LAB\(LLB63) DBNull\(LLB64) DBNull\ACCP	CCP Ab (Group:ACCP)	ACCP
4	LAB\(LLB63) DBNull\(LLB64) DBNull\ACCP\BC1-1318	A-CCP (Test:bc1-1318)	BC1-1318
4	LAB\(LLB63) DBNull\(LLB64) DBNull\ACCP\WC890.0334	CYCLIC CITRULLINATED PEPTIDE (Test:wc890.0334)	WC890.0334
3	LAB\(LLB63) DBNull\(LLB64) DBNull\ANTCCP	CCP Ab, IgG (Group:ANTCCP)	ANTCCP
4	LAB\(LLB63) DBNull\(LLB64) DBNull\ANTCCP\FC50.05	ANTI CYCLIC CITRULLINE PEPTIDE (Test:fc50.05)	FC50.05
4	LAB\(LLB63) DBNull\(LLB64) DBNull\ANTCCP\LC6138	CCP IgG Antibodies (Test:lc6138)	LC6138
4	LAB\(LLB63) DBNull\(LLB64) DBNull\ANTCCP\MCSQ-AN...	Anti CCP (Test:mcsq-antccp)	MCSQ-ANTCCP
4	LAB\(LLB63) DBNull\(LLB64) DBNull\ANTCCP\MISQ-CCPT	anti-CCP IgG (Test:misq-ccpt)	MISQ-CCPT
4	LAB\(LLB63) DBNull\(LLB64) DBNull\ANTCCP\NCCCPG	ANTI CCP IGG (Test:ncccpg)	NCCCPG

Determine which patients have a CCP titer >40 and have been seen in either the BWH Arthritis center or MGH Arthritis Associates.

```
select distinct(patient_num)
from observation_fact
where concept_cd in
(select c_basecode from CCPcodes) and nval_num >40
and patient_num in
(select patient_num from vg_ravisitsbypatient)
and patient_num in
(select distinct(patient_num)
from visit_dimension
where location_path like '%arthritis%')
```

Tips

- Look at the tables
- Bear in mind that you won't need to concentrate on every field in every table, but can drill down into the particular fields of interest as needed.
- Figure out how the dimension tables tie into the fact table.
- Check out the mapping tables that are included if you need to further identify the data.
- Try running one of the sample queries.
- Think about what questions you want answered,then try to frame them based on the data in the data mart.
- Write the SQL to perform your queries. Start slowly and gradually build up the complexity of each query.



Principles of Creating an i2b2 Cell

Shawn Murphy
Michael Mendis

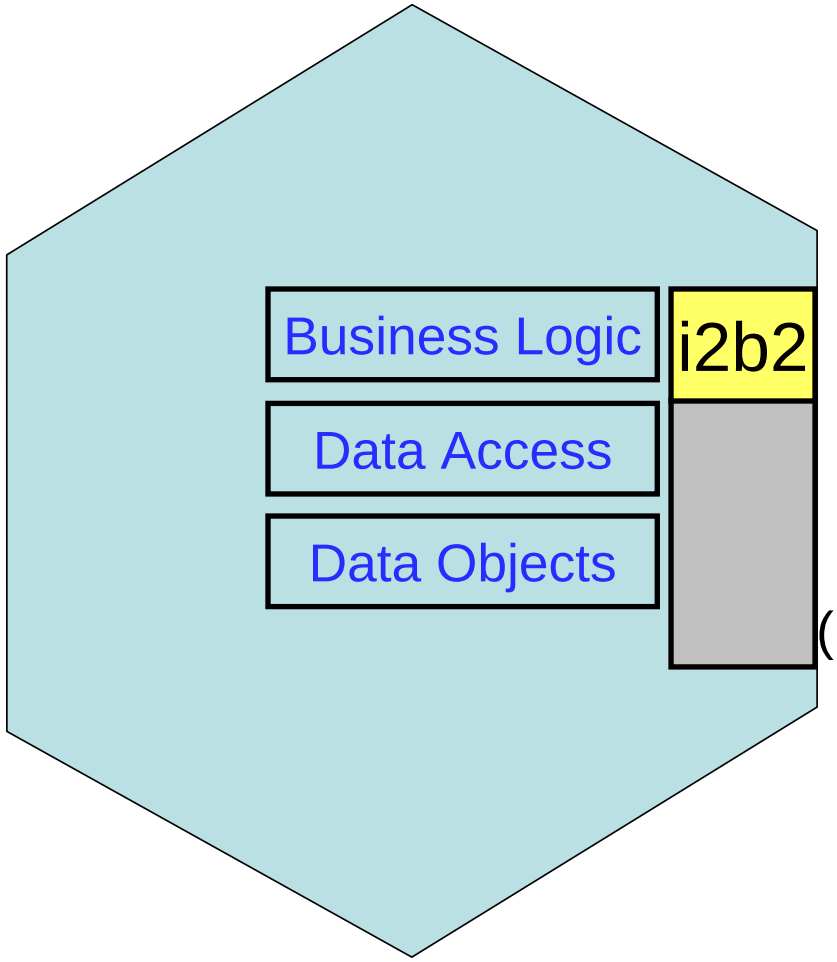
i2b2 Hive - Architecture

- 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 Cell Architecture

- Leverage existing software
- Use Web services as basic form of interaction
- Provide tools to help developers distill complexity into basic automation for clinical investigators
- Emphasize usable open protocols

i2b2 Cell: Canonical Hive Unit



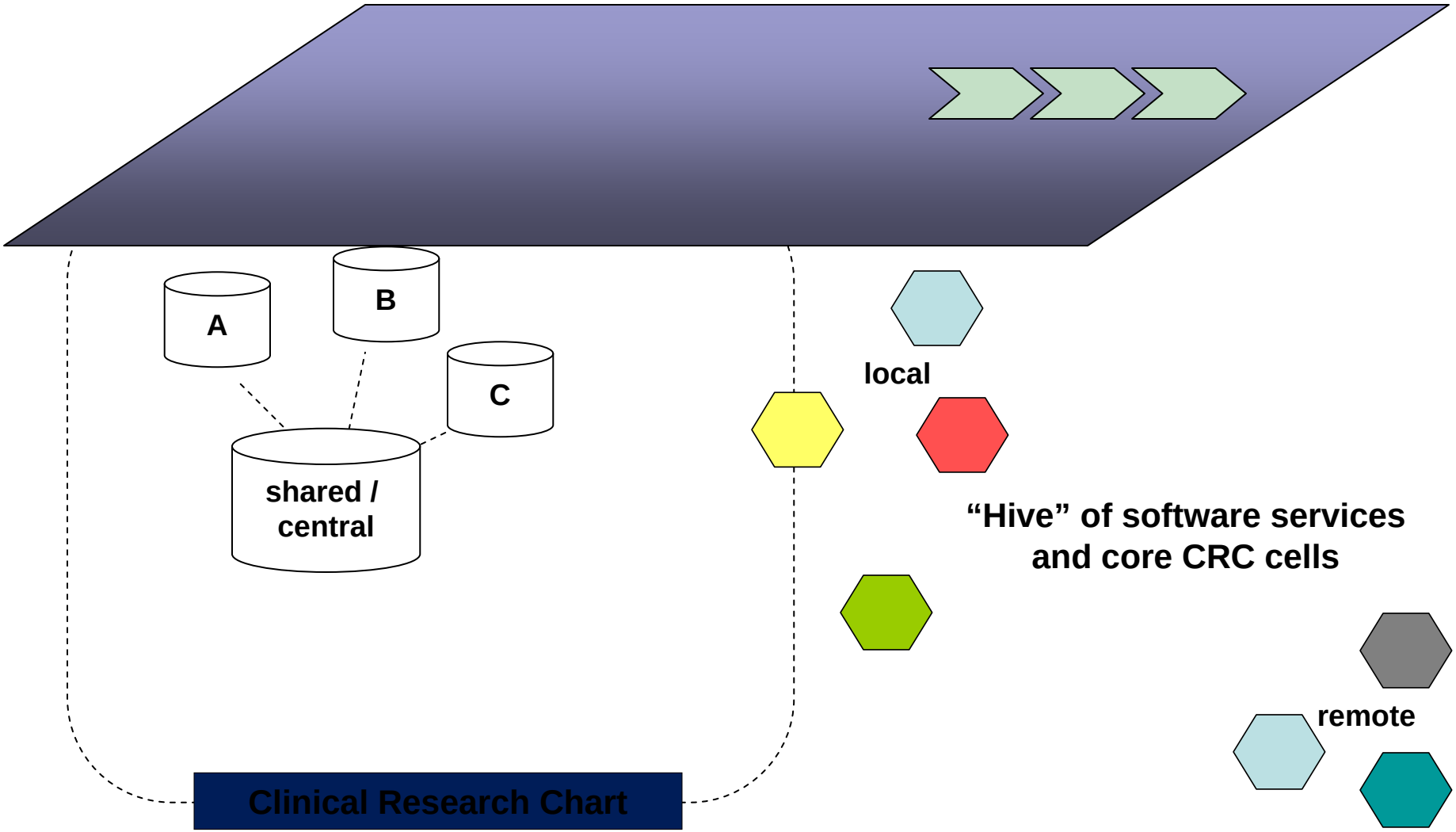
HTTP XML
(minimum: RESTful, others like SOAP optional)

Programmatic Access

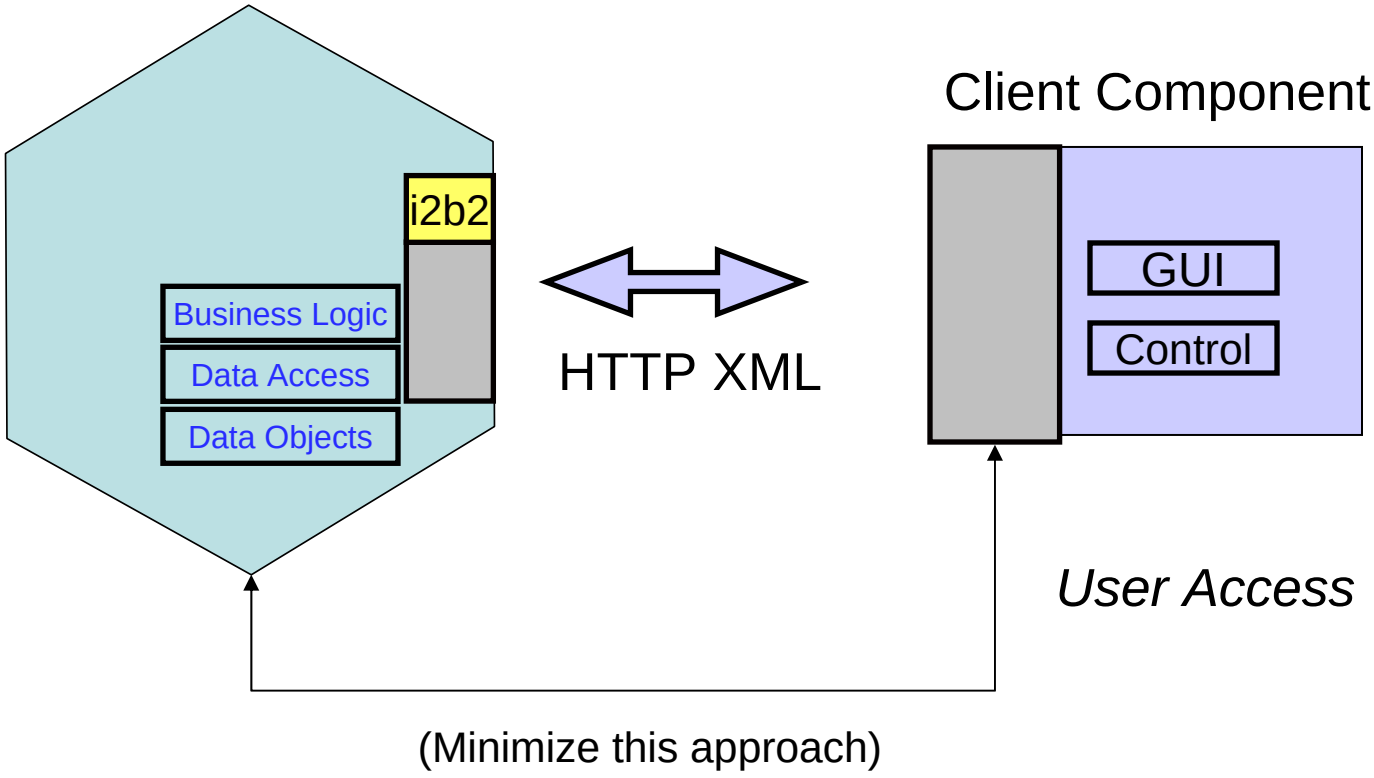
Exposing Cells

- At a low level for integrators; ie, bioinformaticians & software engineers
- At a functional level for investigators
- i2b2 toolkits to allow integrators to expose controlled functionality to investigators so it may be used in workflows.

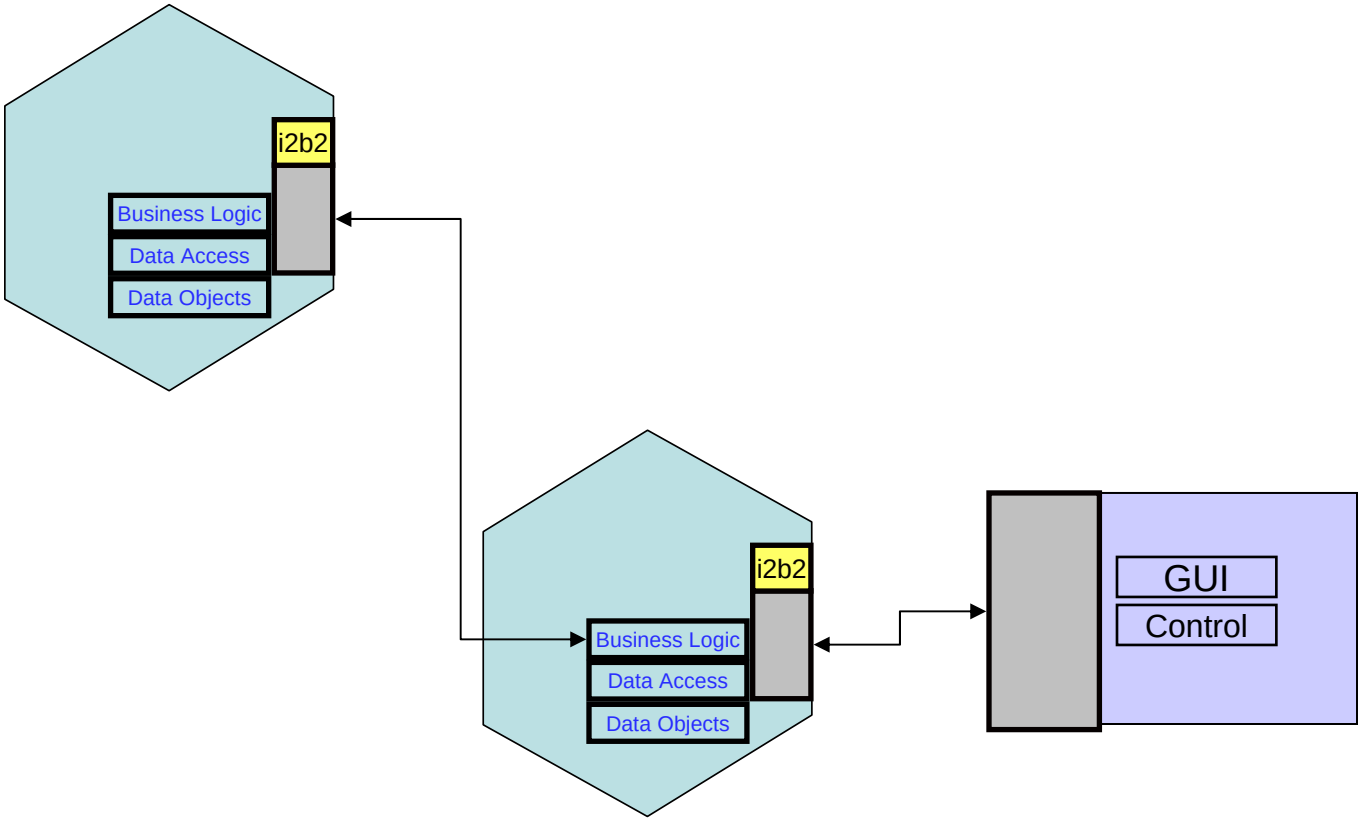
i2b2 Environment

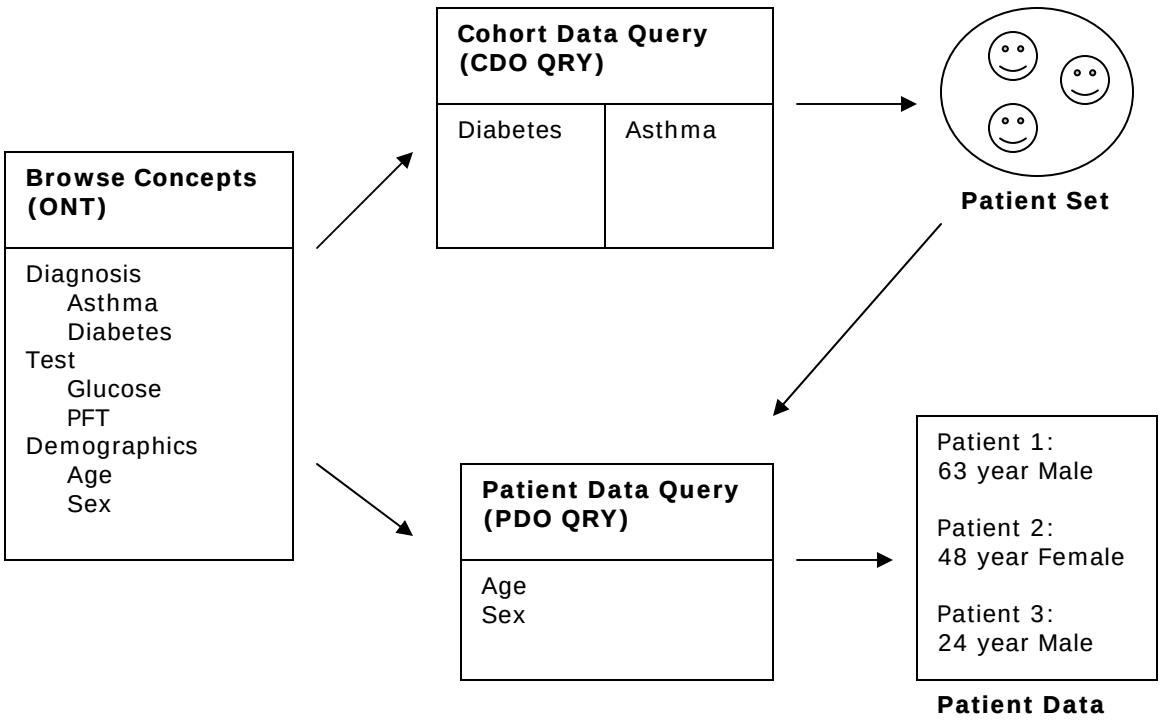


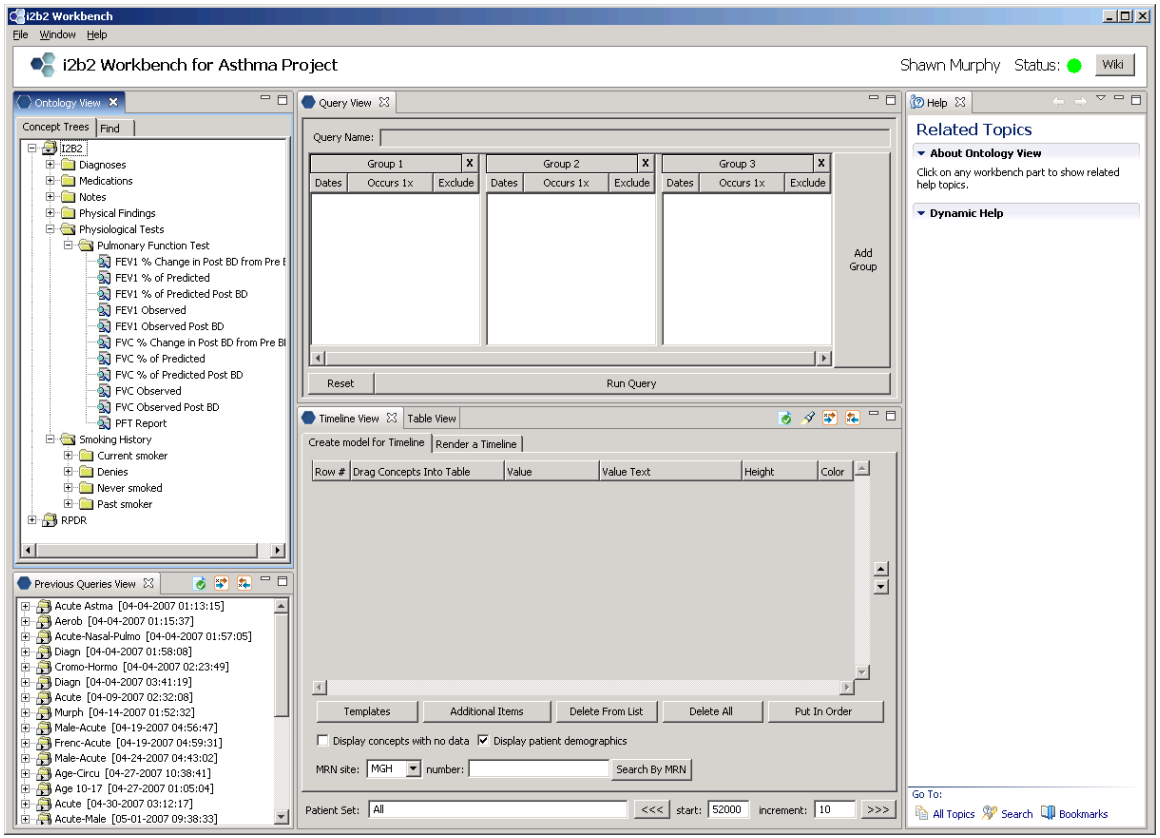
i2b2 Cell:

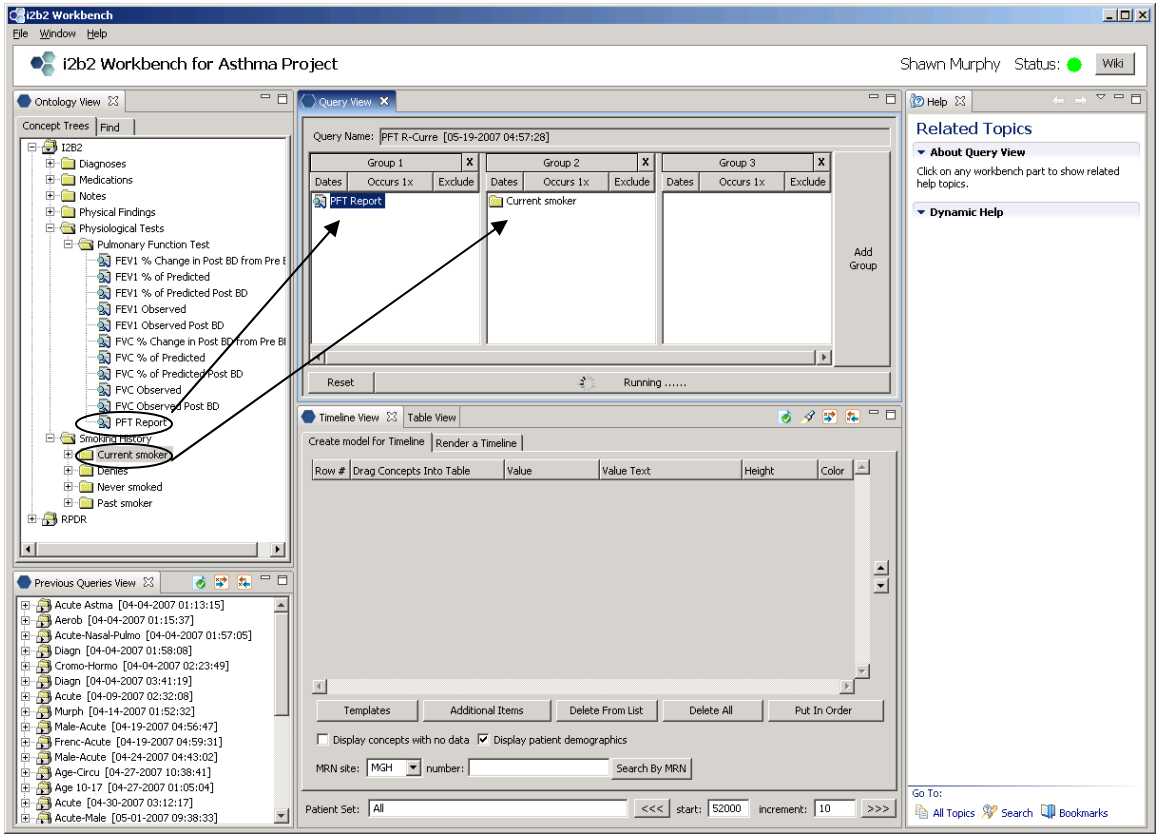

Client Component
i2b2
Business Logic
Data Access
Data Objects
HTTP XML
GUI
Control
User Access
(Minimize this approach)

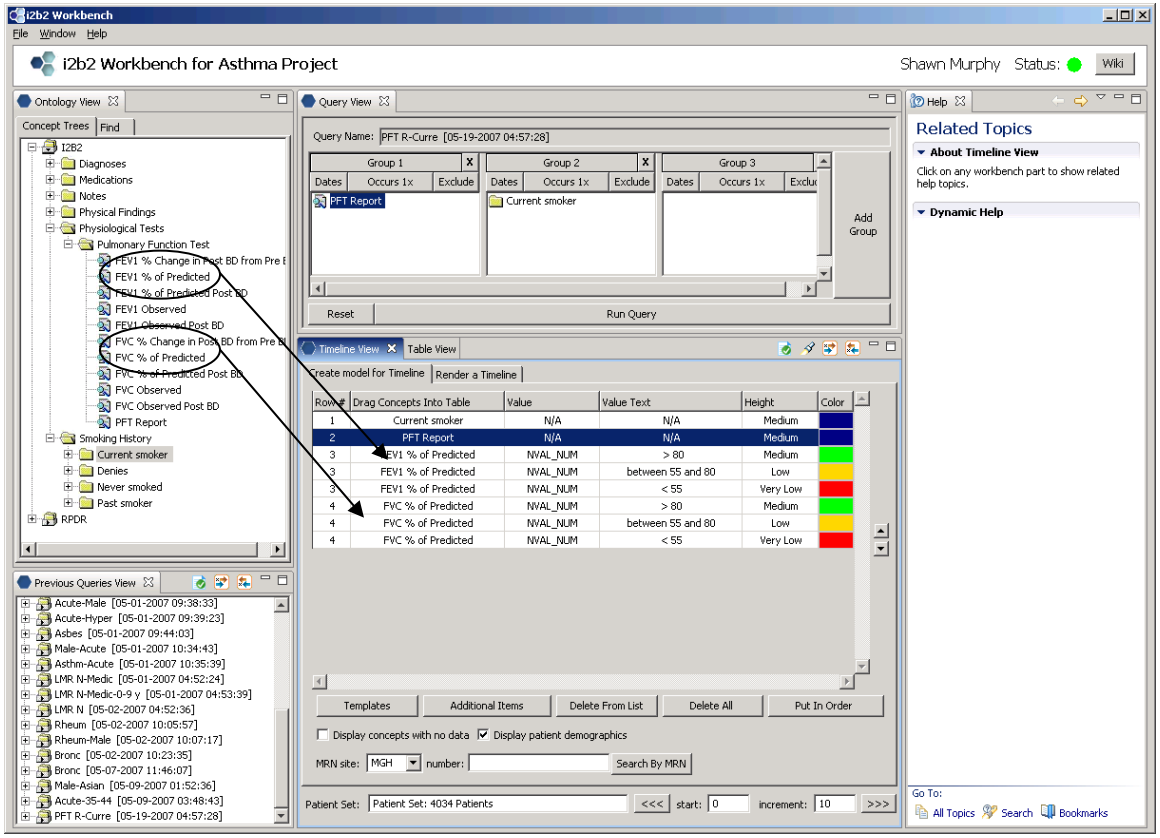
Traversing the i2b2 Hive

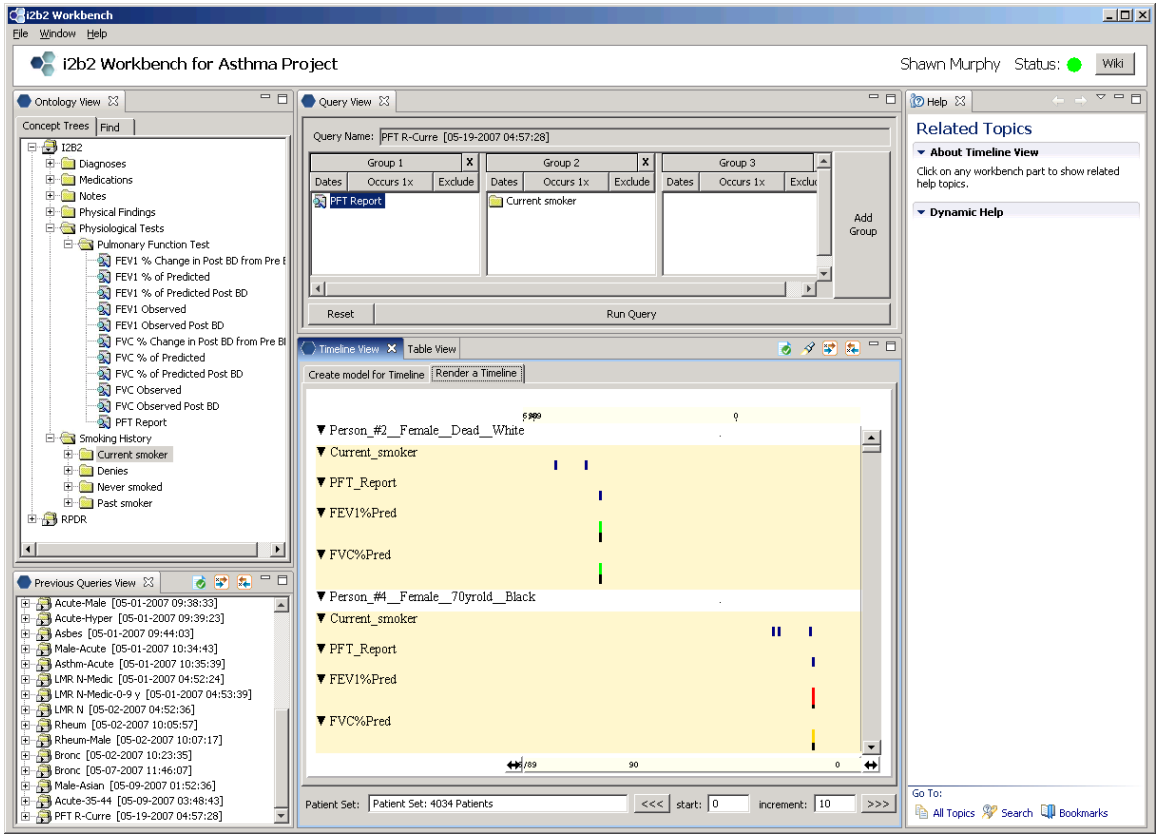



Browse Concepts (ONT)
Diagnosis
Asthma
Diabetes
Test
Glucose
PFT
Demographics
Age
Sex
Cohort Data Query (CDO QRY)
Diabetes
Asthma
Patient Set
Patient Data Query (PDO QRY)
Age
Sex
Patient 1:
63 year Male
Patient 2:
48 year Female
Patient 3:
24 year Male
Patient Data

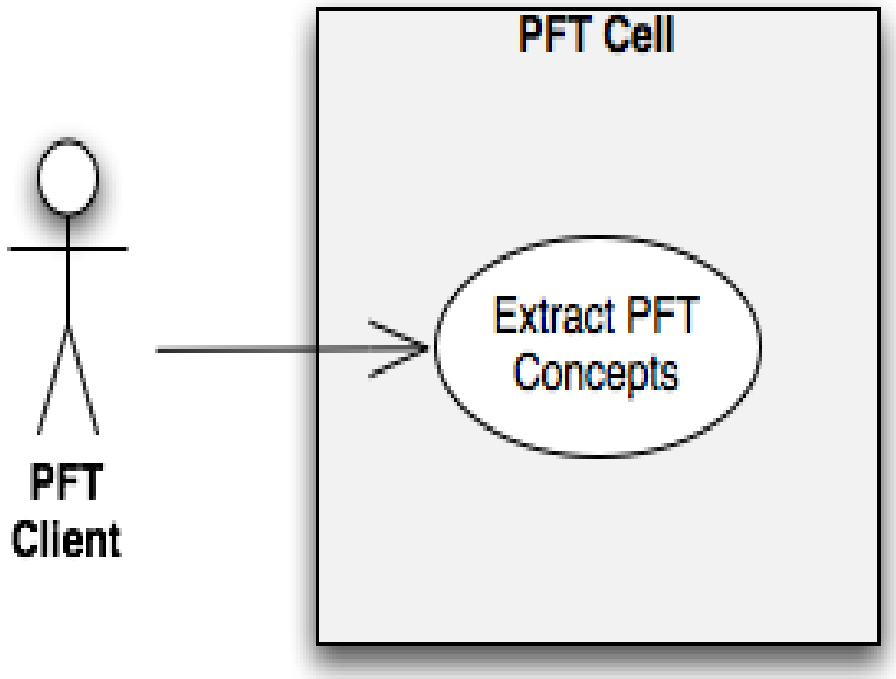

i2b2 Workbench
File Window Help
i2b2 Workbench for Asthma Project
Shawn Murphy Status:
Wiki
Ontology View
Concept Trees Find
I2B2
Diagnoses
Medications
Notes
Physical Findings
Physiological Tests
Pulmonary Function Test
FEV1 % Change in Post BD from Pre
FEV1 % of Predicted
FEV1 % of Predicted Post BD
FEV1 Observed
FEV1 Observed Post BD
FVC % Change in Post BD from Pre
FVC % of Predicted
FVC % of Predicted Post BD
FVC Observed
FVC Observed Post BD
PFT Report
Smoking History
Current smoker
Denies
Never smoked
Past smoker
RPDR
Previous Queries View
Acute Astma [04-04-2007 01:13:15]
Aerob [04-04-2007 01:15:37]
Acute-Nasal-Pulmo [04-04-2007 01:57:05]
Diagn [04-04-2007 01:58:08]
Cromo-Hormo [04-04-2007 02:23:49]
Diagn [04-04-2007 03:41:19]
Acute [04-09-2007 02:32:08]
Murph [04-14-2007 01:52:32]
Male-Acute [04-19-2007 04:56:47]
Frenc-Acute [04-19-2007 04:59:31]
Male-Acute [04-24-2007 04:43:02]
Age-Circu [04-27-2007 10:38:41]
Age 10-17 [04-27-2007 01:05:04]
Acute [04-30-2007 03:12:17]
Acute-Male [05-01-2007 09:38:33]
Query View
Query Name:
Group 1
Group 2
Group 3
Dates
Occurs 1x
Exclude
Add Group
Reset
Run Query
Timeline View
Table View
Create model for Timeline
Render a Timeline
Row #
Drag Concepts Into Table
Value
Value Text
Height
Color
Templates
Additional Items
Delete From List
Delete All
Put In Order
Display concepts with no data
Display patient demographics
MRN site:
MGH
number:
Search By MRN
Patient Set:
All
<<<
start:
52000
increment:
10
>>>
Help
Related Topics
About Ontology View
Click on any workbench part to show related help topics.
Dynamic Help
Go To:
All Topics
Search
Bookmarks

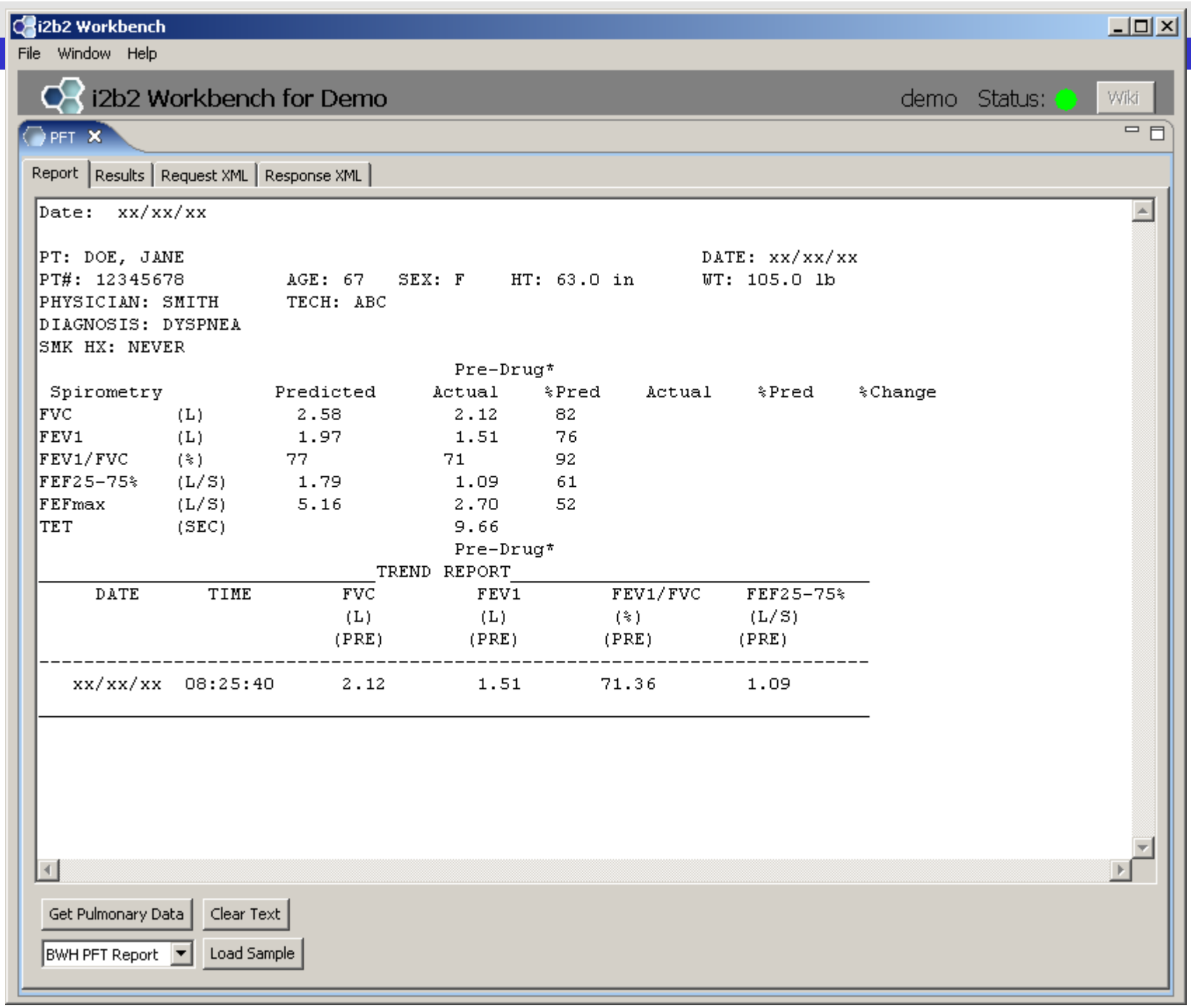

i2b2 Workbench
File Window Help
i2b2 Workbench for Asthma Project
Shawn Murphy Status:
Wiki
Ontology View
Concept Trees Find
I2B2
Diagnoses
Medications
Notes
Physical Findings
Physiological Tests
Pulmonary Function Test
FEV1 % Change in Post BD from Pre
FEV1 % of Predicted
FEV1 % of Predicted Post BD
FEV1 Observed
FEV1 Observed Post BD
FVC % Change in Post BD from Pre Bl
FVC % of Predicted
FVC % of Predicted Post BD
FVC Observed
FVC Observed Post BD
PFT Report
Smoking History
Current smoker
Denies
Never smoked
Past smoker
RPDR
Previous Queries View
Acute Astma [04-04-2007 01:13:15]
Aerob [04-04-2007 01:15:37]
Acute-Nasal-Pulmo [04-04-2007 01:57:05]
Diagn [04-04-2007 01:58:08]
Cromo-Hormo [04-04-2007 02:23:49]
Diagn [04-04-2007 03:41:19]
Acute [04-09-2007 02:32:08]
Murph [04-14-2007 01:52:32]
Male-Acute [04-19-2007 04:56:47]
Frenc-Acute [04-19-2007 04:59:31]
Male-Acute [04-24-2007 04:43:02]
Age-Circu [04-27-2007 10:38:41]
Age 10-17 [04-27-2007 01:05:04]
Acute [04-30-2007 03:12:17]
Acute-Male [05-01-2007 09:38:33]
Query View
Query Name: PFT R-Curre [05-19-2007 04:57:28]
Group 1
Group 2
Group 3
Dates Occurs 1x Exclude
PFT Report
Current smoker
Add Group
Reset
Running .....
Timeline View
Table View
Create model for Timeline
Render a Timeline
Row # Drag Concepts Into Table Value Value Text Height Color
Templates
Additional Items
Delete From List
Delete All
Put In Order
Display concepts with no data
Display patient demographics
MRN site: MGH number:
Search By MRN
Patient Set: All
<<<
start: 52000
increment: 10
>>>
Help
Related Topics
About Query View
Click on any workbench part to show related help topics.
Dynamic Help
Go To:
All Topics
Search
Bookmarks


i2b2 Workbench
File Window Help
i2b2 Workbench for Asthma Project
Shawn Murphy Status:
Wiki
Ontology View
Concept Trees Find
I2B2
Diagnoses
Medications
Notes
Physical Findings
Physiological Tests
Pulmonary Function Test
FEV1 % of Predicted
FEV1 % of Predicted Post BD
FEV1 Observed
FEV1 Observed Post BD
FVC % of Predicted
FVC % of Predicted Post BD
FVC Observed
FVC Observed Post BD
PFT Report
Smoking History
Current smoker
Denies
Never smoked
Past smoker
RPDR
Previous Queries View
Acute-Male [05-01-2007 09:38:33]
Acute-Hyper [05-01-2007 09:39:23]
Asbes [05-01-2007 09:44:03]
Male-Acute [05-01-2007 10:34:43]
Asthm-Acute [05-01-2007 10:35:39]
LMR N-Medic [05-01-2007 04:52:24]
LMR N-Medic-0-9 y [05-01-2007 04:53:39]
LMR N [05-02-2007 04:52:36]
Rheum [05-02-2007 10:05:57]
Rheum-Male [05-02-2007 10:07:17]
Bronc [05-02-2007 10:23:35]
Bronc [05-07-2007 11:46:07]
Male-Asian [05-09-2007 01:52:36]
Acute-35-44 [05-09-2007 03:48:43]
PFT R-Curre [05-19-2007 04:57:28]
Query View
Query Name: PFT R-Curre [05-19-2007 04:57:28]
Group 1
Group 2
Group 3
Dates Occurs 1x Exclude
PFT Report
Current smoker
Add Group
Reset
Run Query
Timeline View
Table View
Create model for Timeline
Render a Timeline
Row # Drag Concepts Into Table Value Value Text Height Color
1 Current smoker N/A N/A Medium
2 PFT Report N/A N/A Medium
3 FEV1 % of Predicted NVAL_NUM > 80 Medium
3 FEV1 % of Predicted NVAL_NUM between 55 and 80 Low
3 FEV1 % of Predicted NVAL_NUM < 55 Very Low
4 FVC % of Predicted NVAL_NUM > 80 Medium
4 FVC % of Predicted NVAL_NUM between 55 and 80 Low
4 FVC % of Predicted NVAL_NUM < 55 Very Low
Templates
Additional Items
Delete From List
Delete All
Put In Order
Display concepts with no data
Display patient demographics
MRN site: MGH number:
Search By MRN
Patient Set: Patient Set: 4034 Patients
<<<
start: 0
increment: 10
>>>
Help
Related Topics
About Timeline View
Click on any workbench part to show related help topics.
Dynamic Help
Go To:
All Topics
Search
Bookmarks

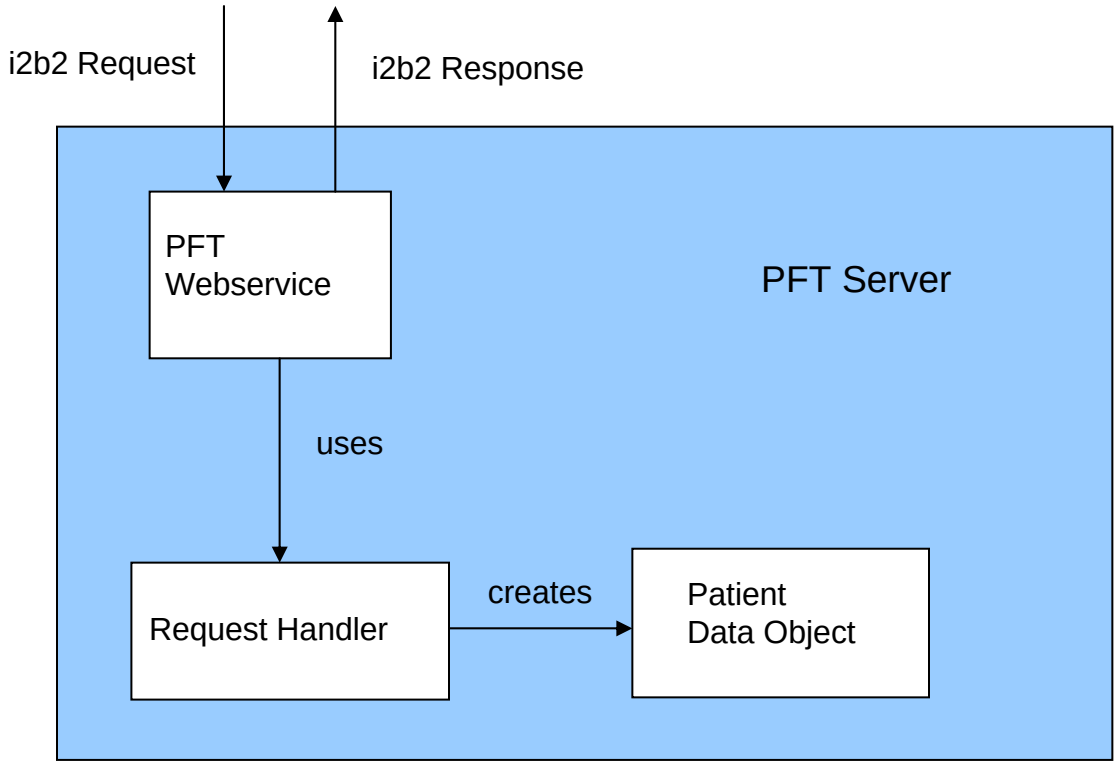

i2b2 Workbench
File Window Help
i2b2 Workbench for Asthma Project
Shawn Murphy Status:
Wiki
Ontology View
Concept Trees Find
I2B2
Diagnoses
Medications
Notes
Physical Findings
Physiological Tests
Pulmonary Function Test
FEV1 % Change in Post BD from Pre
FEV1 % of Predicted
FEV1 % of Predicted Post BD
FEV1 Observed
FEV1 Observed Post BD
FVC % Change in Post BD from Pre
FVC % of Predicted
FVC % of Predicted Post BD
FVC Observed
FVC Observed Post BD
PFT Report
Smoking History
Current smoker
Denies
Never smoked
Past smoker
RPDR
Previous Queries View
Acute-Male [05-01-2007 09:38:33]
Acute-Hyper [05-01-2007 09:39:23]
Asbes [05-01-2007 09:44:03]
Male-Acute [05-01-2007 10:34:43]
Asthm-Acute [05-01-2007 10:35:39]
LMR N-Medic [05-01-2007 04:52:24]
LMR N-Medic-0-9 y [05-01-2007 04:53:39]
LMR N [05-02-2007 04:52:36]
Rheum [05-02-2007 10:05:57]
Rheum-Male [05-02-2007 10:07:17]
Bronc [05-02-2007 10:23:35]
Bronc [05-07-2007 11:46:07]
Male-Asian [05-09-2007 01:52:36]
Acute-35-44 [05-09-2007 03:48:43]
PFT R-Curre [05-19-2007 04:57:28]
Query View
Query Name: PFT R-Curre [05-19-2007 04:57:28]
Group 1
Group 2
Group 3
Dates Occurs 1x Exclude
PFT Report
Current smoker
Add Group
Reset
Run Query
Timeline View
Table View
Create model for Timeline
Render a Timeline
Person_#2__Female__Dead__White
Current_smoker
PFT_Report
FEV1%Pred
FVC%Pred
Person_#4__Female__70yrold__Black
Current_smoker
PFT_Report
FEV1%Pred
FVC%Pred
Patient Set:
Patient Set: 4034 Patients
<<<
start: 0
increment: 10
>>>
Help
Related Topics
About Timeline View
Click on any workbench part to show related help topics.
Dynamic Help
Go To:
All Topics
Search
Bookmarks

Cells perform services

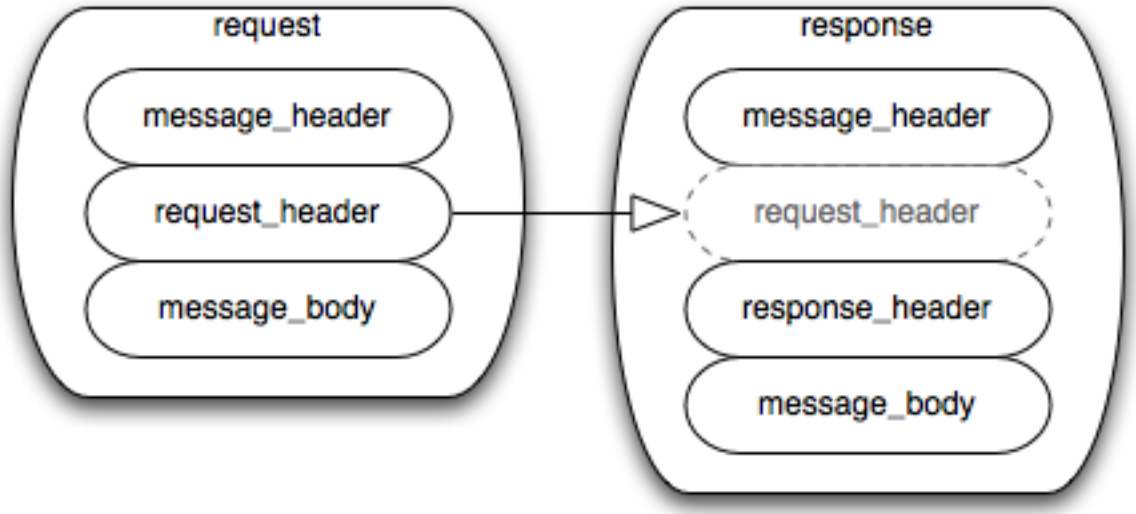


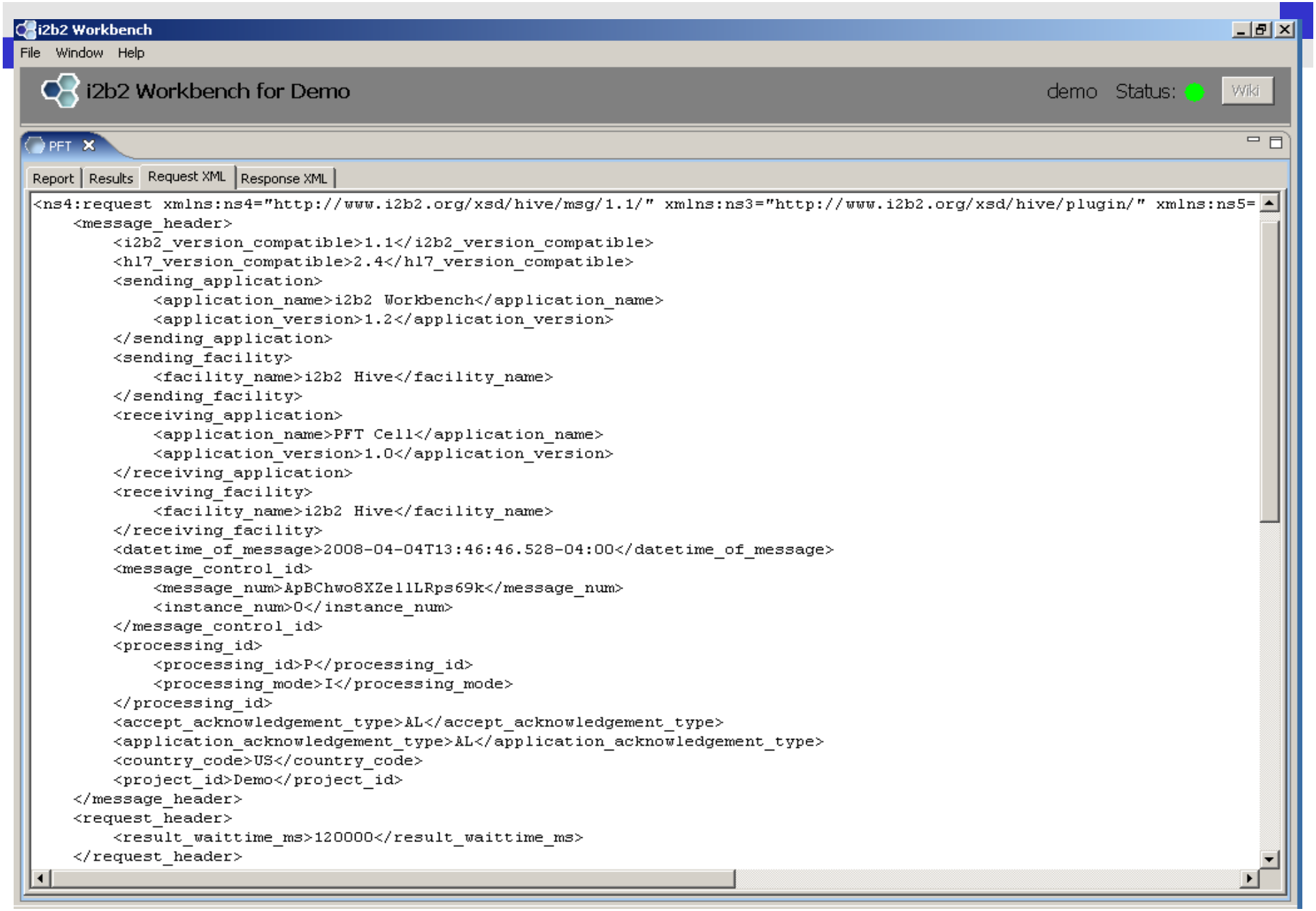

i2b2 Workbench
File Window Help
i2b2 Workbench for Demo
demo Status:
Wiki
PFT
Report
Results
Request XML
Response XML
Date: xx/xx/xx
PT: DOE, JANE
DATE: xx/xx/xx
PT#: 12345678 AGE: 67 SEX: F HT: 63.0 in WT: 105.0 lb
PHYSICIAN: SMITH TECH: ABC
DIAGNOSIS: DYSPNEA
SMK HX: NEVER
Pre-Drug*
Spirometry Predicted Actual %Pred Actual %Pred %Change
FVC (L) 2.58 2.12 82
FEV1 (L) 1.97 1.51 76
FEV1/FVC (%) 77 71 92
FEF25-75% (L/S) 1.79 1.09 61
FEFmax (L/S) 5.16 2.70 52
TET (SEC) 9.66
Pre-Drug*
TREND REPORT
DATE TIME FVC (L) (PRE) FEV1 (L) (PRE) FEV1/FVC (%) (PRE) FEF25-75% (L/S) (PRE)
xx/xx/xx 08:25:40 2.12 1.51 71.36 1.09
Get Pulmonary Data
Clear Text
BWH PFT Report
Load Sample

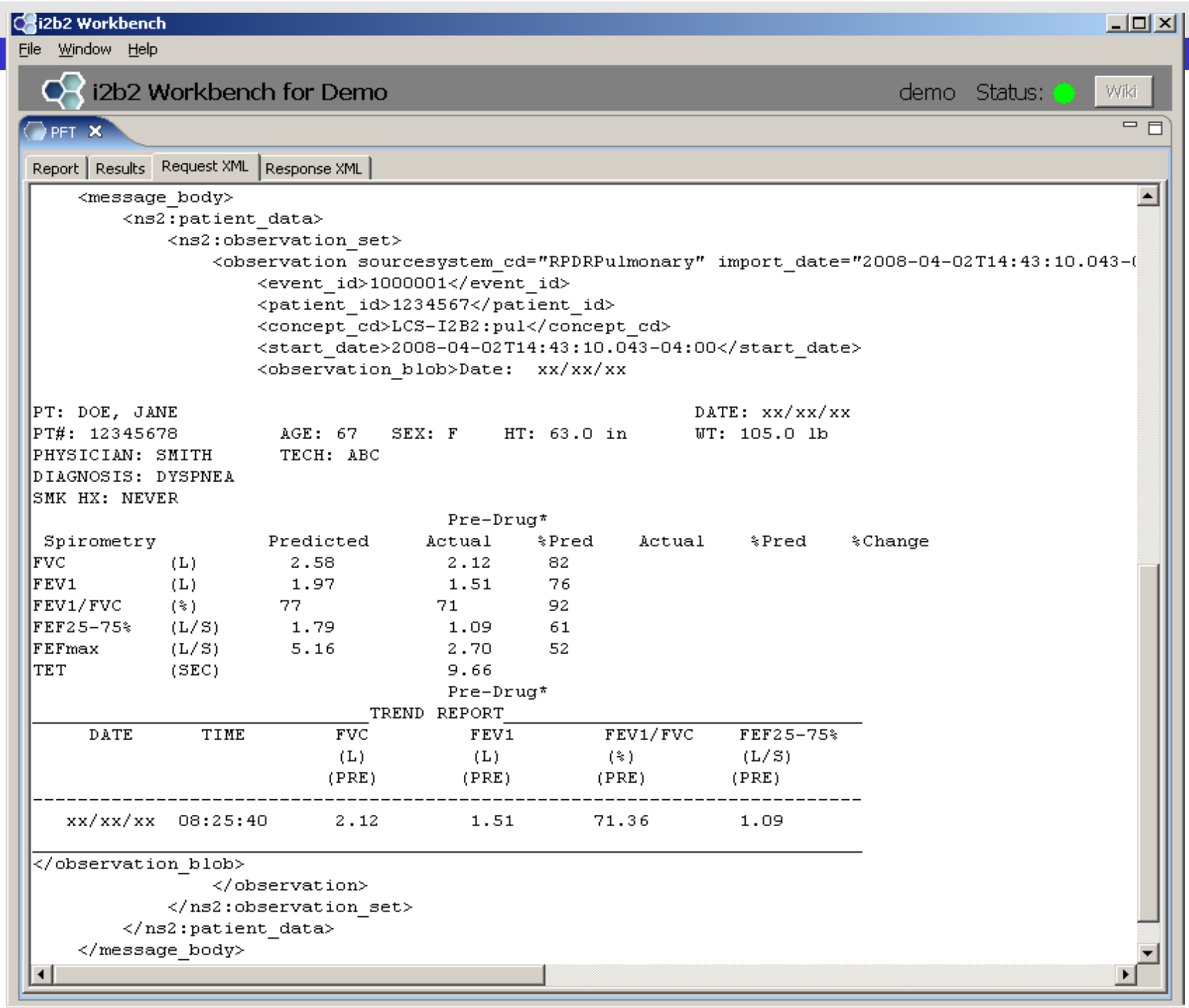
PFT Cell

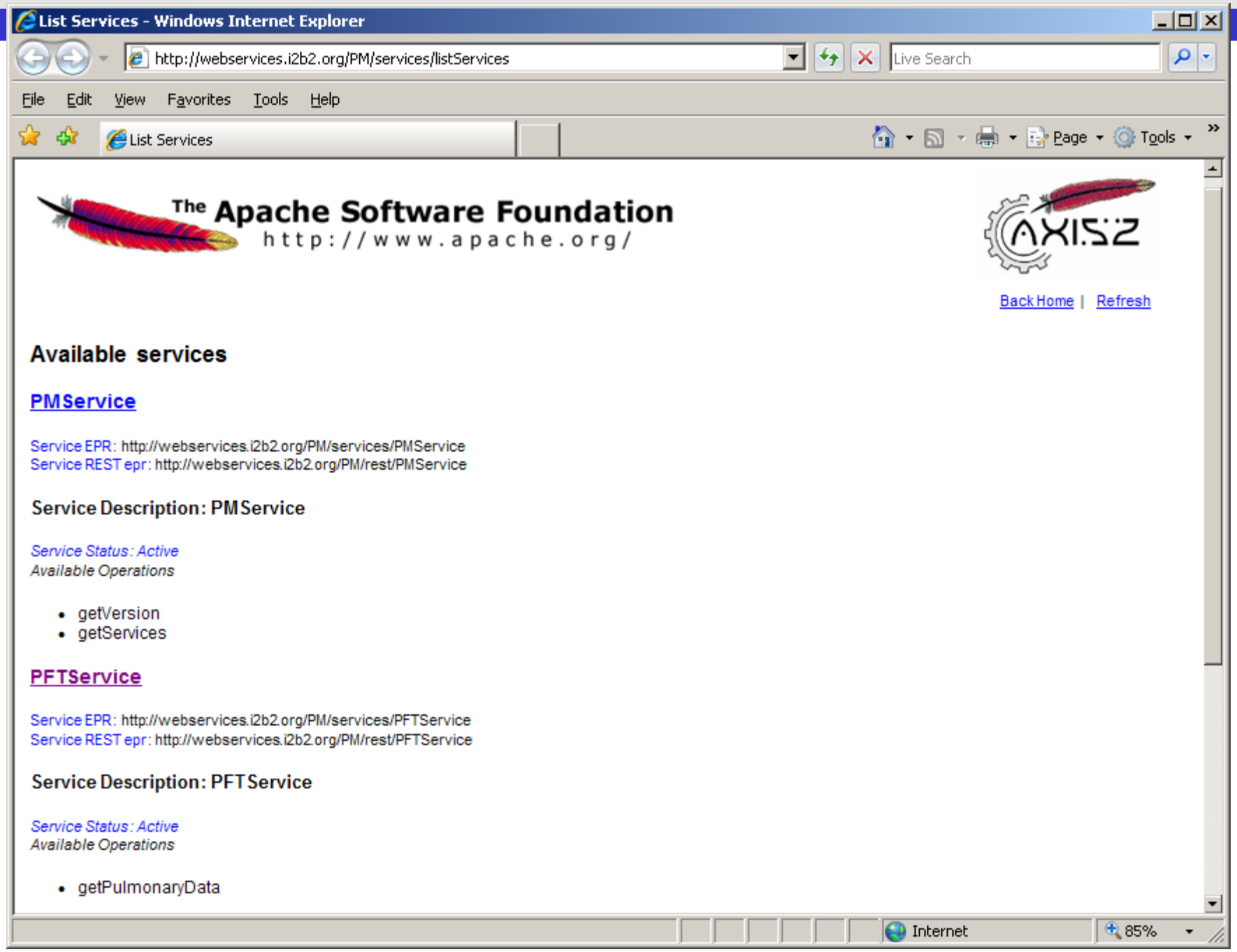

i2b2 Request
i2b2 Response
PFT
Webservice
PFT Server
uses
Request Handler
creates
Patient
Data Object

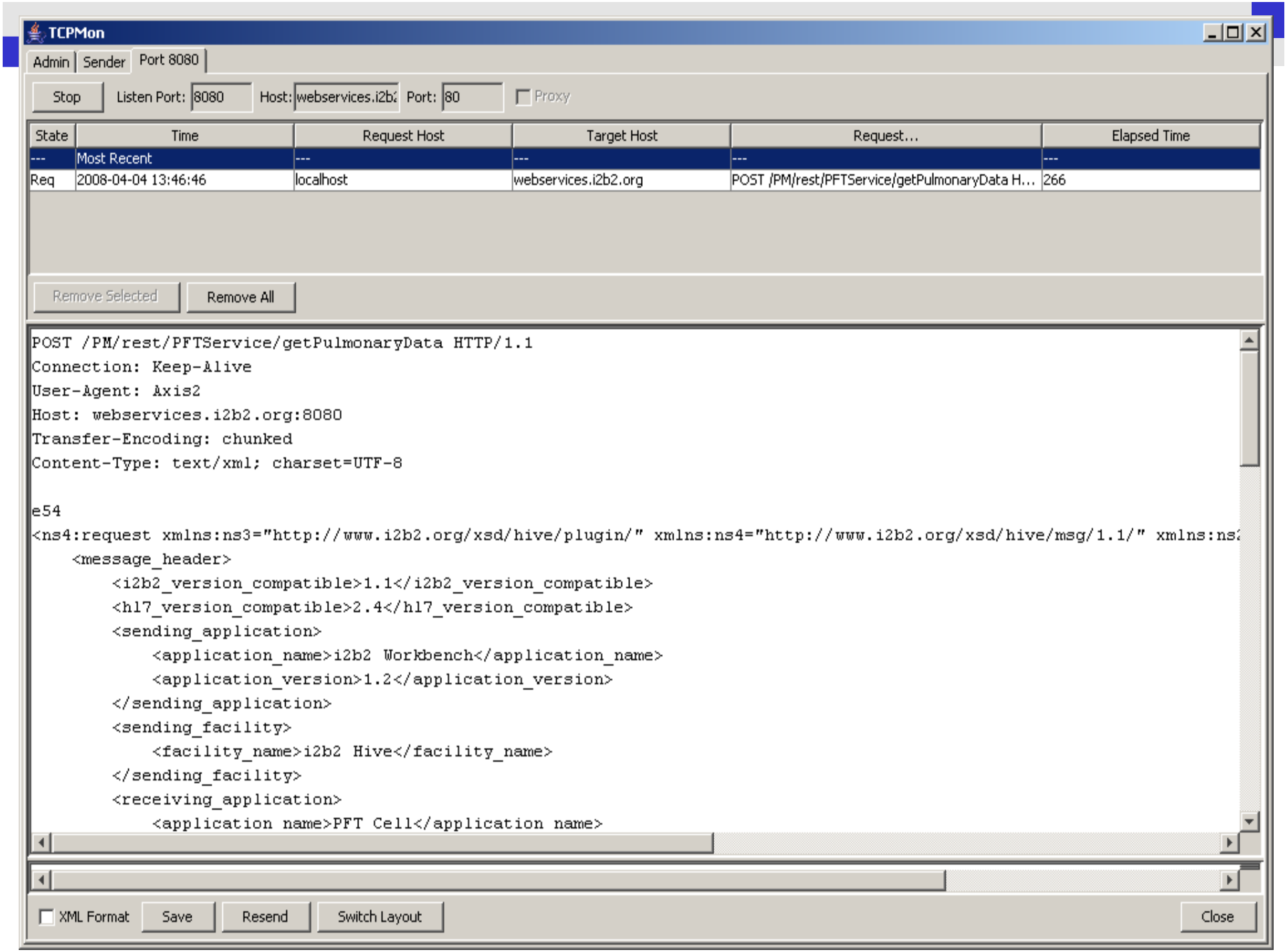
i2b2 Message Format



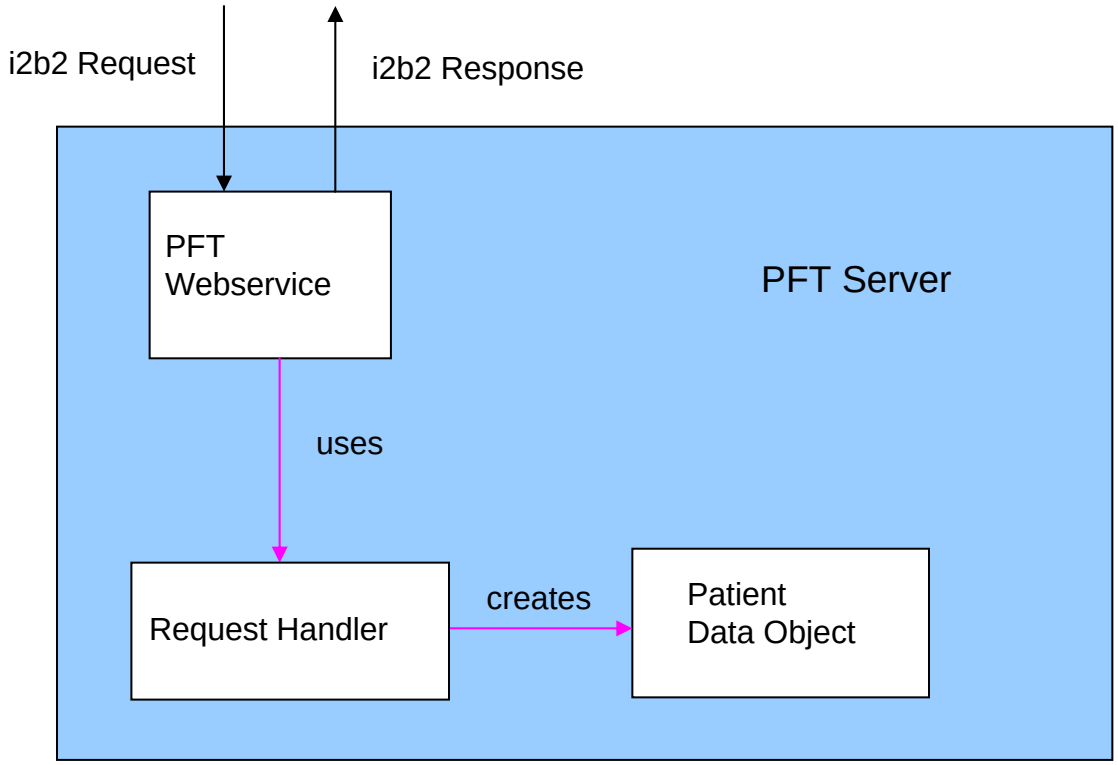

i2b2 Workbench
File Window Help
i2b2 Workbench for Demo
demo Status:
Wiki
PFT
Report Results Request XML Response XML
<ns4:request xmlns:ns4="http://www.i2b2.org/xsd/hive/msg/1.1/" xmlns:ns3="http://www.i2b2.org/xsd/hive/plugin/" xmlns:ns5=
<message_header>
<i2b2_version_compatible>1.1</i2b2_version_compatible>
<hl7_version_compatible>2.4</hl7_version_compatible>
<sending_application>
<application_name>i2b2 Workbench</application_name>
<application_version>1.2</application_version>
</sending_application>
<sending_facility>
<facility_name>i2b2 Hive</facility_name>
</sending_facility>
<receiving_application>
<application_name>PFT Cell</application_name>
<application_version>1.0</application_version>
</receiving_application>
<receiving_facility>
<facility_name>i2b2 Hive</facility_name>
</receiving_facility>
<datetime_of_message>2008-04-04T13:46:46.528-04:00</datetime_of_message>
<message_control_id>
<message_num>ApBChwo8XZellLRps69k</message_num>
<instance_num>0</instance_num>
</message_control_id>
<processing_id>
<processing_id>P</processing_id>
<processing_mode>I</processing_mode>
</processing_id>
<accept_acknowledgement_type>AL</accept_acknowledgement_type>
<application_acknowledgement_type>AL</application_acknowledgement_type>
<country_code>US</country_code>
<project_id>Demo</project_id>
</message_header>
<request_header>
<result_waittime_ms>120000</result_waittime_ms>
</request_header>

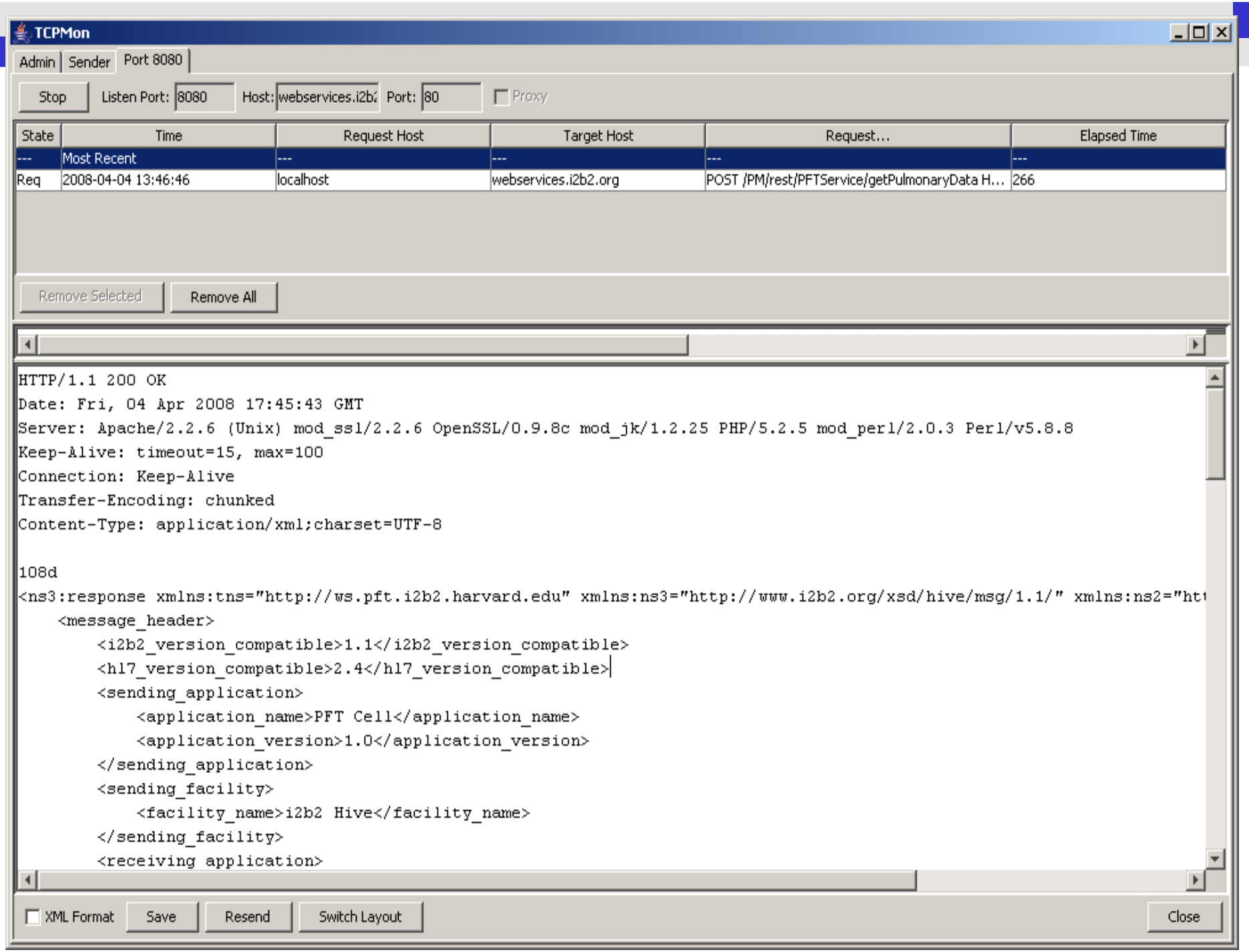

i2b2 Workbench
File Window Help
i2b2 Workbench for Demo
demo Status:
Wiki
PFT
Report Results Request XML Response XML
<message_body>
<ns2:patient_data>
<ns2:observation_set>
<observation sourcesystem_cd="RPDRPulmonary" import_date="2008-04-02T14:43:10.043-(
<event_id>1000001</event_id>
<patient_id>1234567</patient_id>
<concept_cd>LCS-I2B2:pul</concept_cd>
<start_date>2008-04-02T14:43:10.043-04:00</start_date>
<observation_blob>Date: xx/xx/xx
PT: DOE, JANE DATE: xx/xx/xx
PT#: 12345678 AGE: 67 SEX: F HT: 63.0 in WT: 105.0 lb
PHYSICIAN: SMITH TECH: ABC
DIAGNOSIS: DYSPNEA
SMK HX: NEVER
Pre-Drug*
Spirometry Predicted Actual %Pred Actual %Pred %Change
FVC (L) 2.58 2.12 82
FEV1 (L) 1.97 1.51 76
FEV1/FVC (%) 77 71 92
FEF25-75% (L/S) 1.79 1.09 61
FEFmax (L/S) 5.16 2.70 52
TET (SEC) 9.66
Pre-Drug*
TREND REPORT
DATE TIME FVC (L) (PRE) FEV1 (L) (PRE) FEV1/FVC (%) (PRE) FEF25-75% (L/S) (PRE)
xx/xx/xx 08:25:40 2.12 1.51 71.36 1.09
</observation_blob>
</observation>
</ns2:observation_set>
</ns2:patient_data>
</message_body>


List Services - Windows Internet Explorer
http://webservices.i2b2.org/PM/services/listServices
Live Search
File Edit View Favorites Tools Help
List Services
Page Tools
The Apache Software Foundation
http://www.apache.org/
AXIS2
Back Home | Refresh
Available services
PMService
Service EPR: http://webservices.i2b2.org/PM/services/PMService
Service REST epr: http://webservices.i2b2.org/PM/rest/PMService
Service Description: PMService
Service Status: Active
Available Operations
getVersion
getServices
PFTService
Service EPR: http://webservices.i2b2.org/PM/services/PFTService
Service REST epr: http://webservices.i2b2.org/PM/rest/PFTService
Service Description: PFTService
Service Status: Active
Available Operations
getPulmonaryData
Internet
85%


TCPMon
Admin
Sender
Port 8080
Stop
Listen Port: 8080
Host: webservices.i2b2
Port: 80
Proxy
State
Time
Request Host
Target Host
Request...
Elapsed Time
--- Most Recent --- --- --- ---
Req 2008-04-04 13:46:46 localhost webservices.i2b2.org POST /PM/rest/PFTService/getPulmonaryData H... 266
Remove Selected
Remove All
POST /PM/rest/PFTService/getPulmonaryData HTTP/1.1
Connection: Keep-Alive
User-Agent: Axis2
Host: webservices.i2b2.org:8080
Transfer-Encoding: chunked
Content-Type: text/xml; charset=UTF-8
e54
<ns4:request xmlns:ns3="http://www.i2b2.org/xsd/hive/plugin/" xmlns:ns4="http://www.i2b2.org/xsd/hive/msg/1.1/" xmlns:ns
<message_header>
<i2b2_version_compatible>1.1</i2b2_version_compatible>
<hl7_version_compatible>2.4</hl7_version_compatible>
<sending_application>
<application_name>i2b2 Workbench</application_name>
<application_version>1.2</application_version>
</sending_application>
<sending_facility>
<facility_name>i2b2 Hive</facility_name>
</sending_facility>
<receiving_application>
<application name>PFT Cell</application name>
XML Format
Save
Resend
Switch Layout
Close

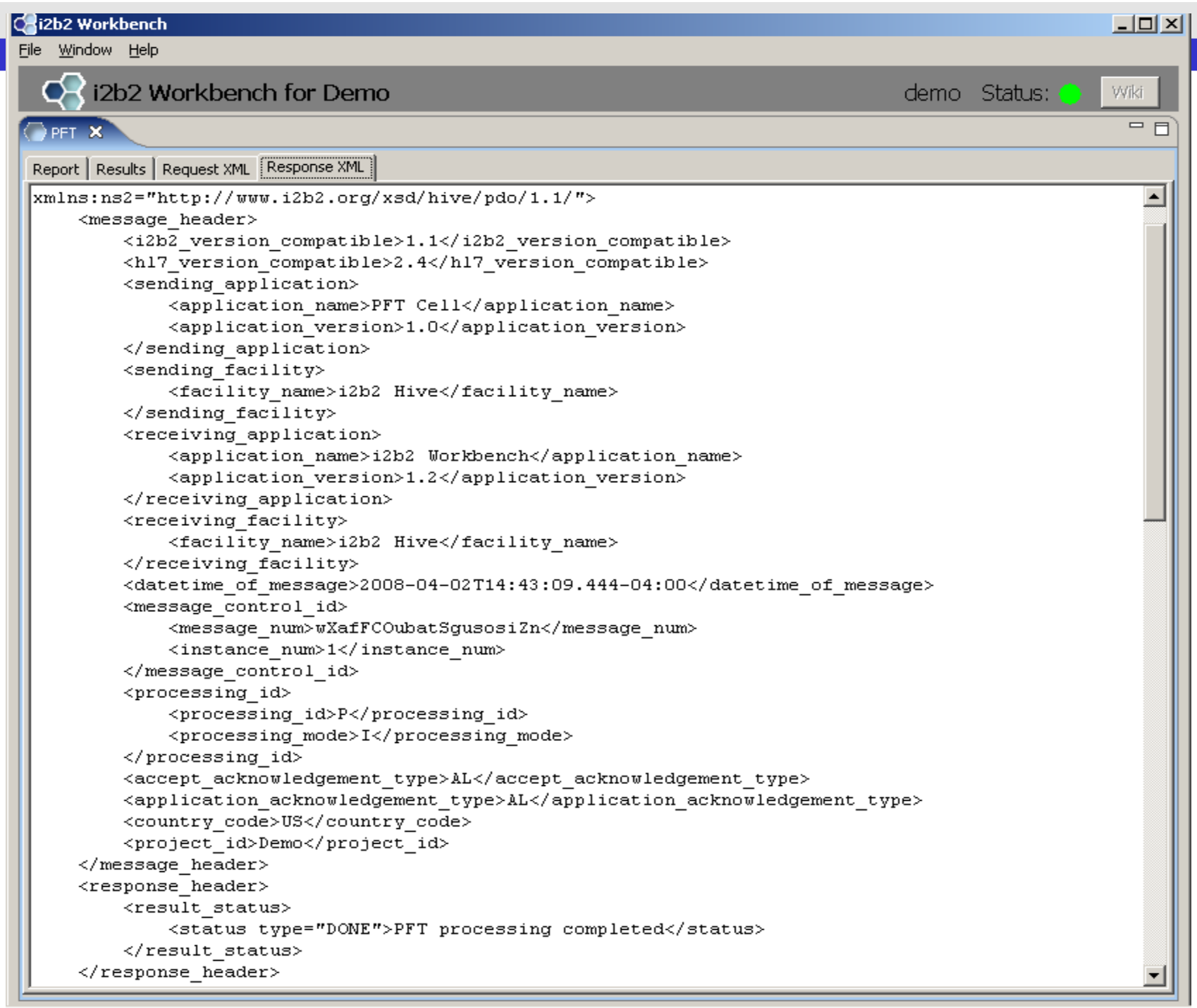
PFT Cell

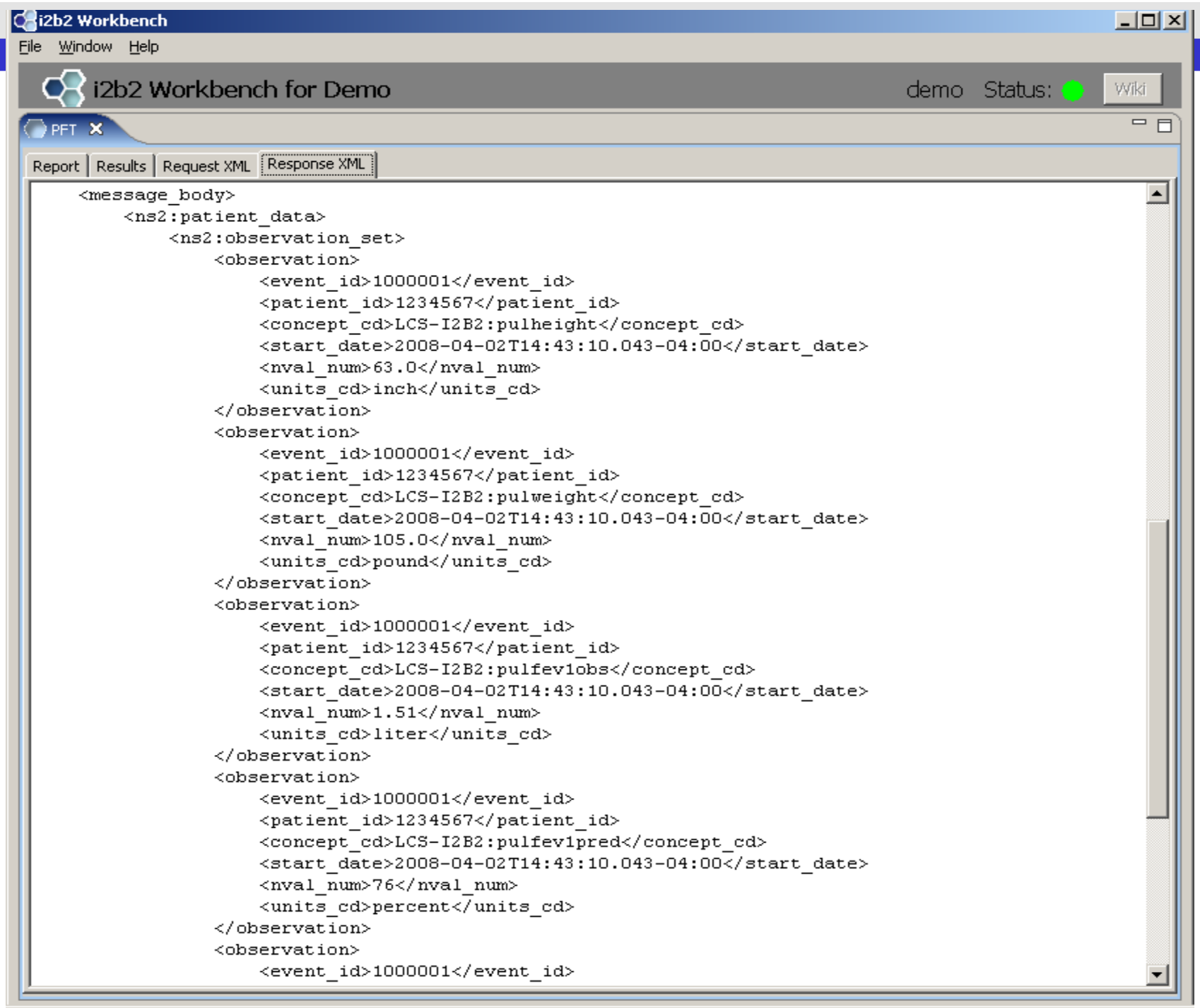

i2b2 Request
i2b2 Response
PFT
Webservice
PFT Server
uses
Request Handler
creates
Patient
Data Object

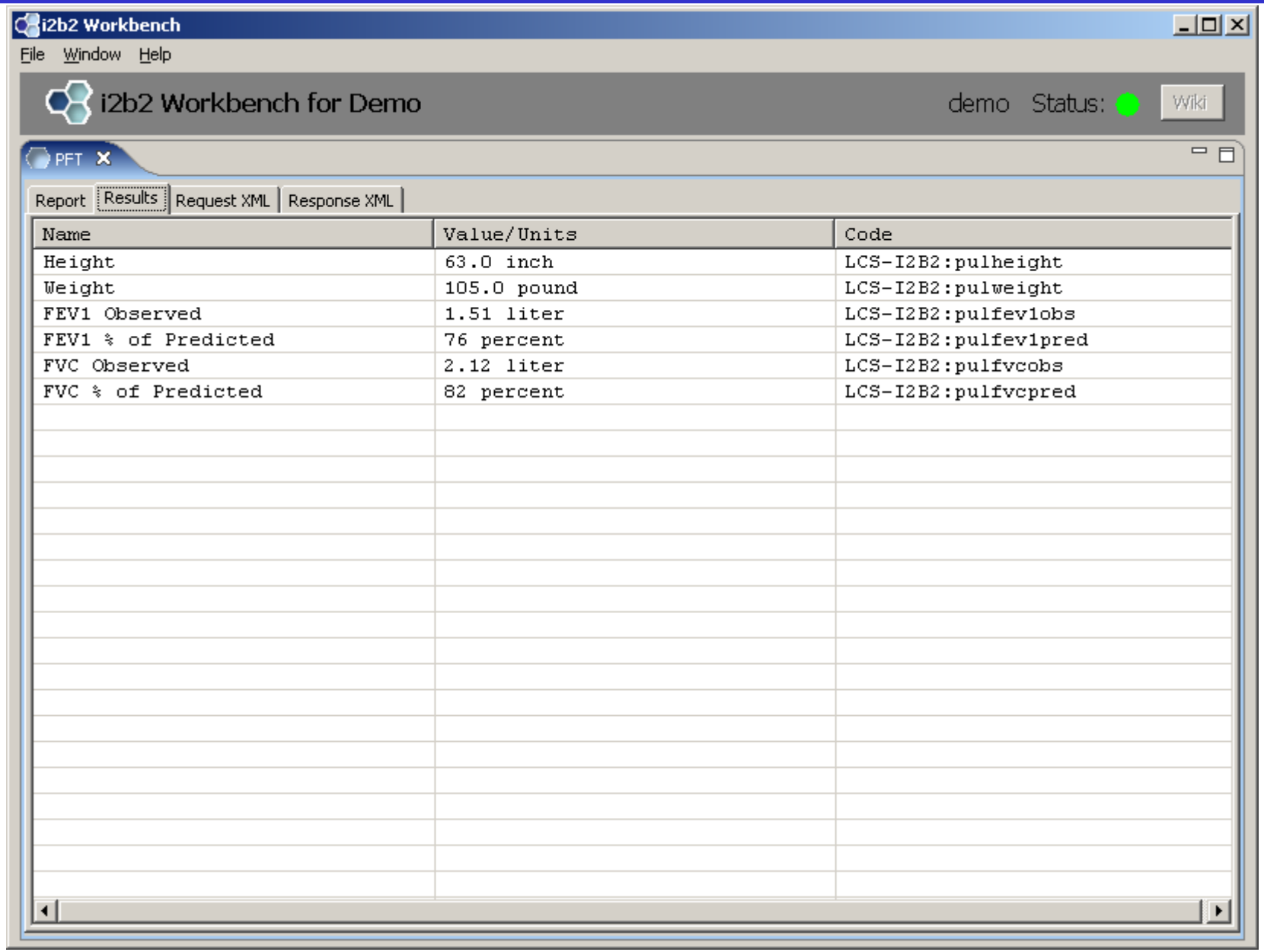

TCPMon
Admin
Sender
Port 8080
Stop
Listen Port: 8080
Host: webservices.i2b2
Port: 80
Proxy
State
Time
Request Host
Target Host
Request...
Elapsed Time

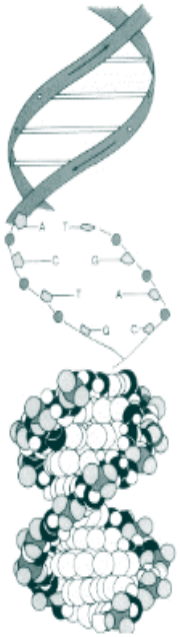
Most Recent

Req
2008-04-04 13:46:46
localhost
webservices.i2b2.org
POST /PM/rest/PFTService/getPulmonaryData H...
266
Remove Selected
Remove All
HTTP/1.1 200 OK
Date: Fri, 04 Apr 2008 17:45:43 GMT
Server: Apache/2.2.6 (Unix) mod_ssl/2.2.6 OpenSSL/0.9.8c mod_jk/1.2.25 PHP/5.2.5 mod_perl/2.0.3 Perl/v5.8.8
Keep-Alive: timeout=15, max=100
Connection: Keep-Alive
Transfer-Encoding: chunked
Content-Type: application/xml;charset=UTF-8
108d
<ns3:response xmlns:tns="http://ws.pft.i2b2.harvard.edu" xmlns:ns3="http://www.i2b2.org/xsd/hive/msg/1.1/" xmlns:ns2="ht
<message_header>
<i2b2_version_compatible>1.1</i2b2_version_compatible>
<hl7_version_compatible>2.4</hl7_version_compatible>
<sending_application>
<application_name>PFT Cell</application_name>
<application_version>1.0</application_version>
</sending_application>
<sending_facility>
<facility_name>i2b2 Hive</facility_name>
</sending_facility>
<receiving application>
XML Format
Save
Resend
Switch Layout
Close


i2b2 Workbench
File Window Help
i2b2 Workbench for Demo
demo Status:
Wiki
PFT
Report Results Request XML Response XML
xmlns:ns2="http://www.i2b2.org/xsd/hive/pdo/1.1/">
<message_header>
<i2b2_version_compatible>1.1</i2b2_version_compatible>
<hl7_version_compatible>2.4</hl7_version_compatible>
<sending_application>
<application_name>PFT Cell</application_name>
<application_version>1.0</application_version>
</sending_application>
<sending_facility>
<facility_name>i2b2 Hive</facility_name>
</sending_facility>
<receiving_application>
<application_name>i2b2 Workbench</application_name>
<application_version>1.2</application_version>
</receiving_application>
<receiving_facility>
<facility_name>i2b2 Hive</facility_name>
</receiving_facility>
<datetime_of_message>2008-04-02T14:43:09.444-04:00</datetime_of_message>
<message_control_id>
<message_num>wXafFCOubatSgusosiZn</message_num>
<instance_num>1</instance_num>
</message_control_id>
<processing_id>
<processing_id>P</processing_id>
<processing_mode>I</processing_mode>
</processing_id>
<accept_acknowledgement_type>AL</accept_acknowledgement_type>
<application_acknowledgement_type>AL</application_acknowledgement_type>
<country_code>US</country_code>
<project_id>Demo</project_id>
</message_header>
<response_header>
<result_status>
<status type="DONE">PFT processing completed</status>
</result_status>
</response_header>


i2b2 Workbench
File Window Help
i2b2 Workbench for Demo
demo Status:
Wiki
PFT
Report Results Request XML Response XML
<message_body>
<ns2:patient_data>
<ns2:observation_set>
<observation>
<event_id>1000001</event_id>
<patient_id>1234567</patient_id>
<concept_cd>LCS-I2B2:pulheight</concept_cd>
<start_date>2008-04-02T14:43:10.043-04:00</start_date>
<nval_num>63.0</nval_num>
<units_cd>inch</units_cd>
</observation>
<observation>
<event_id>1000001</event_id>
<patient_id>1234567</patient_id>
<concept_cd>LCS-I2B2:pulweight</concept_cd>
<start_date>2008-04-02T14:43:10.043-04:00</start_date>
<nval_num>105.0</nval_num>
<units_cd>pound</units_cd>
</observation>
<observation>
<event_id>1000001</event_id>
<patient_id>1234567</patient_id>
<concept_cd>LCS-I2B2:pulfev1obs</concept_cd>
<start_date>2008-04-02T14:43:10.043-04:00</start_date>
<nval_num>1.51</nval_num>
<units_cd>liter</units_cd>
</observation>
<observation>
<event_id>1000001</event_id>
<patient_id>1234567</patient_id>
<concept_cd>LCS-I2B2:pulfev1pred</concept_cd>
<start_date>2008-04-02T14:43:10.043-04:00</start_date>
<nval_num>76</nval_num>
<units_cd>percent</units_cd>
</observation>
<observation>
<event_id>1000001</event_id>


i2b2 Workbench
File Window Help
i2b2 Workbench for Demo
demo Status:
Wiki
PFT
Report Results Request XML Response XML
Name Value/Units Code
Height 63.0 inch LCS-I2B2:pulheight
Weight 105.0 pound LCS-I2B2:pulweight
FEV1 Observed 1.51 liter LCS-I2B2:pulfev1obs
FEV1 % of Predicted 76 percent LCS-I2B2:pulfev1pred
FVC Observed 2.12 liter LCS-I2B2:pulfvcobs
FVC % of Predicted 82 percent LCS-I2B2:pulfvcpred

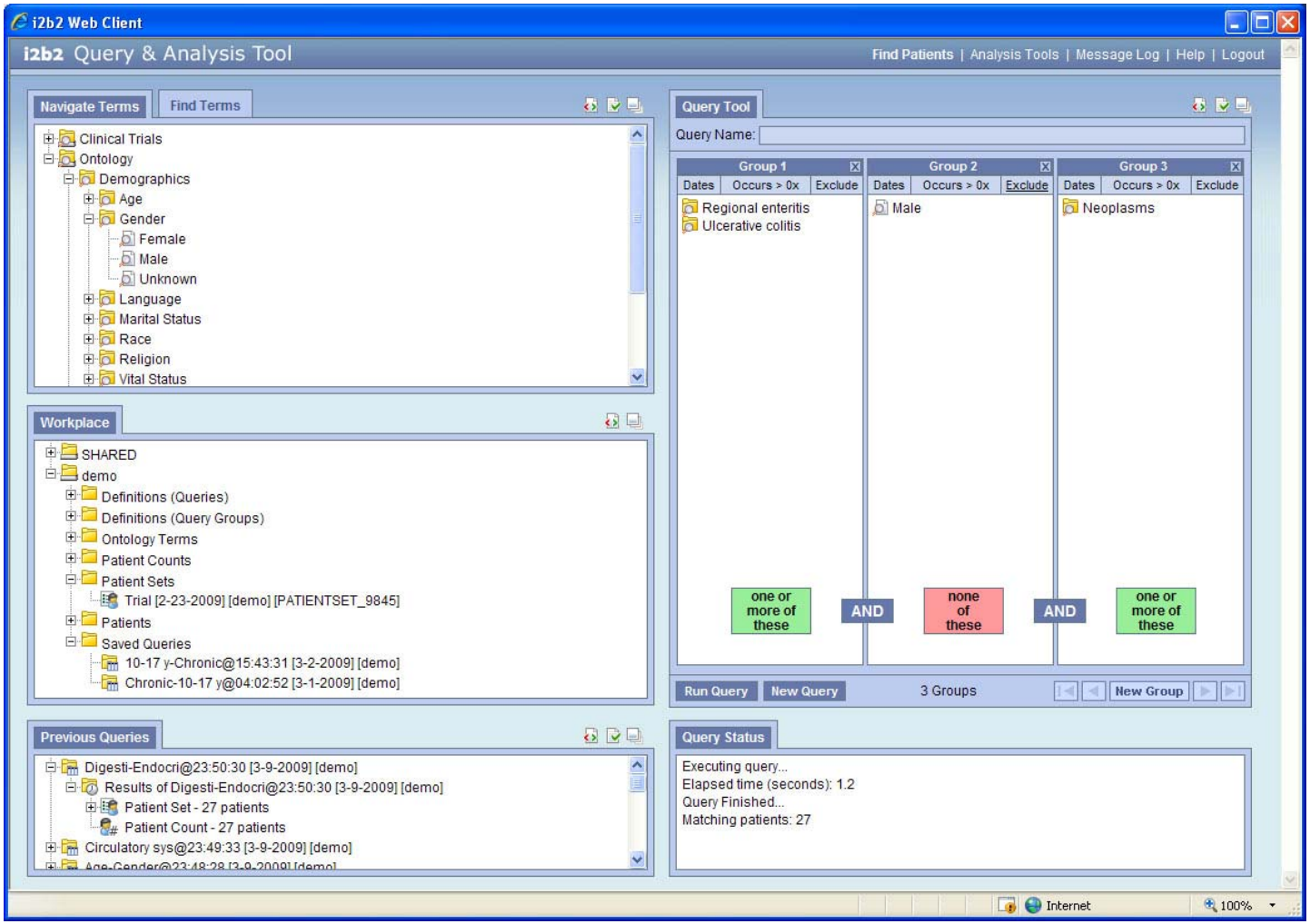


i2b2 Web Client

Griffin Weber, MD, PhD
weber@hms.harvard.edu
Chief Technology Officer
Harvard Medical School
Assistant Professor of Medicine
Beth Israel Deaconess Medical Center

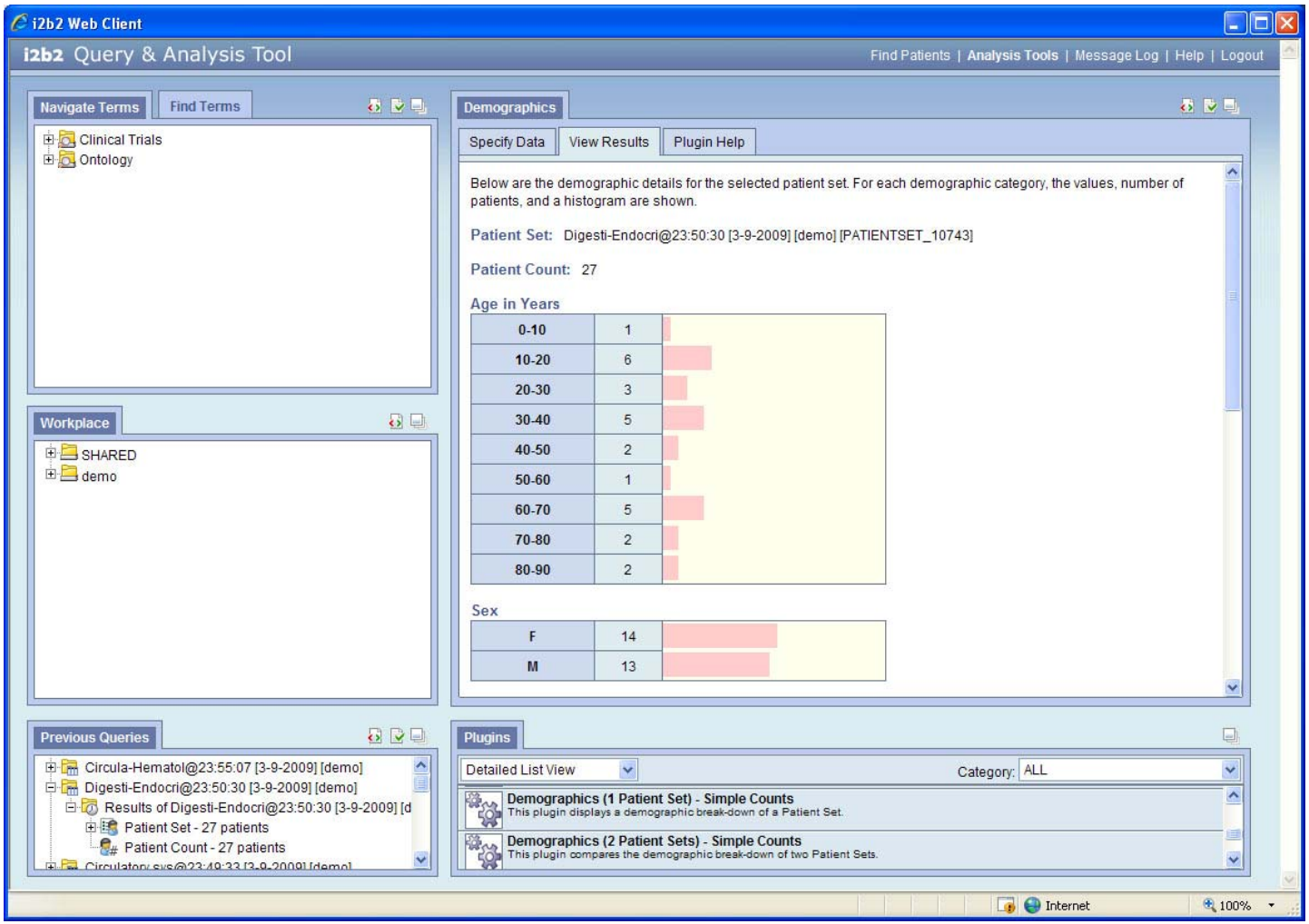
Nicholas Benik
nick_benik@hms.harvard.edu
Senior Software Architect

Find Patients



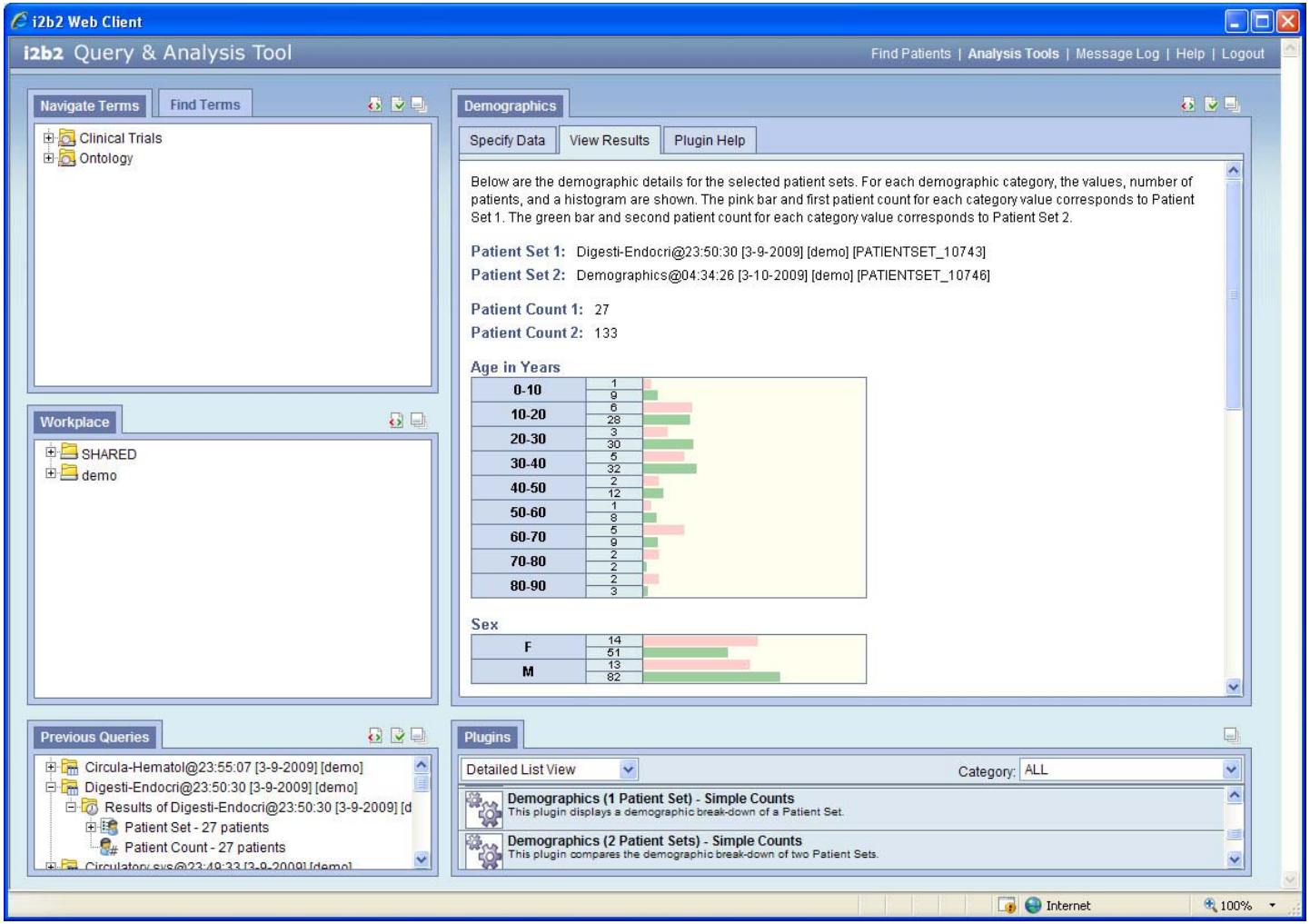
Drag-and-drop query design interface

Demographics Plugin



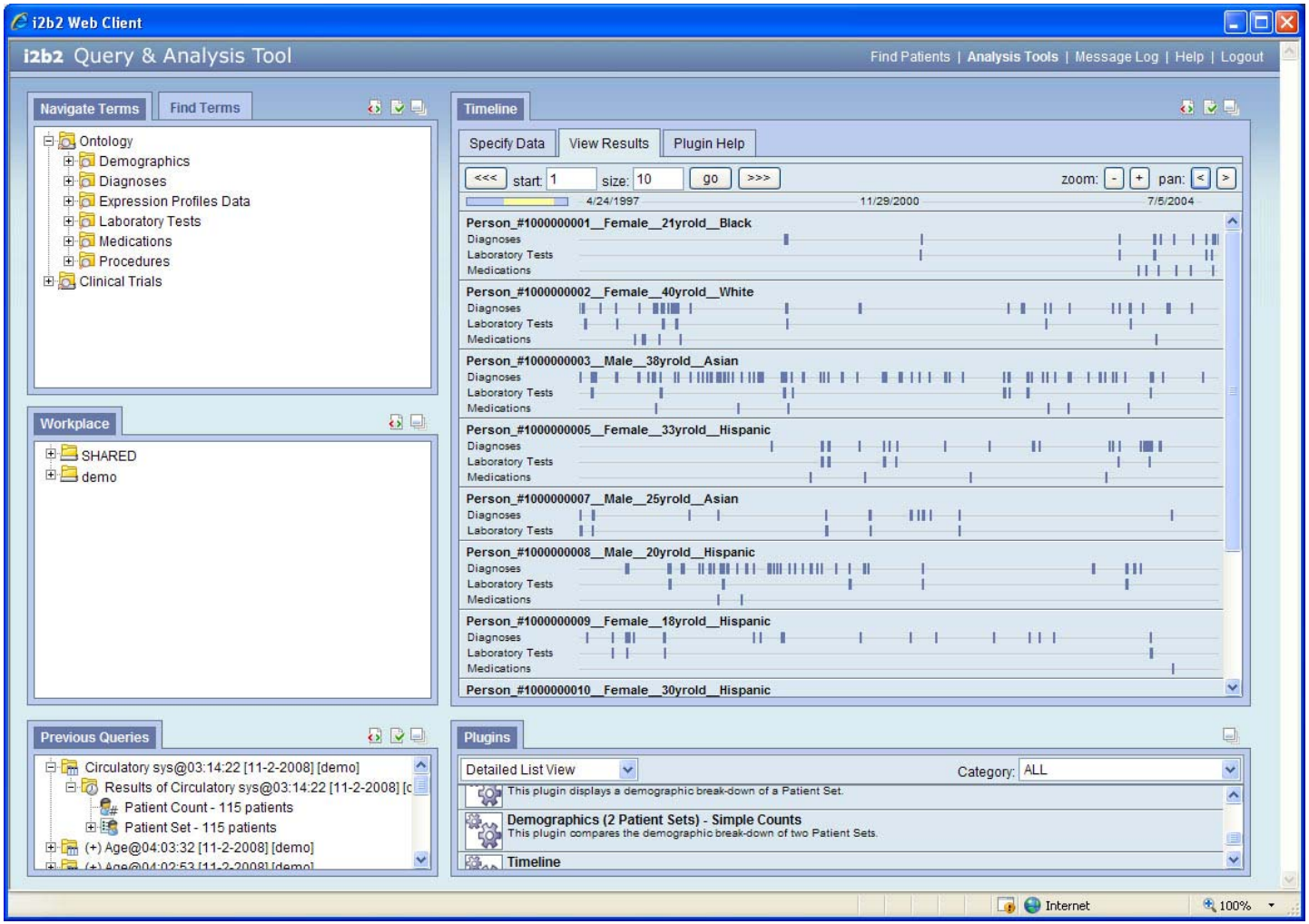
Analysis of a saved patient set using a “plugin”

Two Patient List Plugin



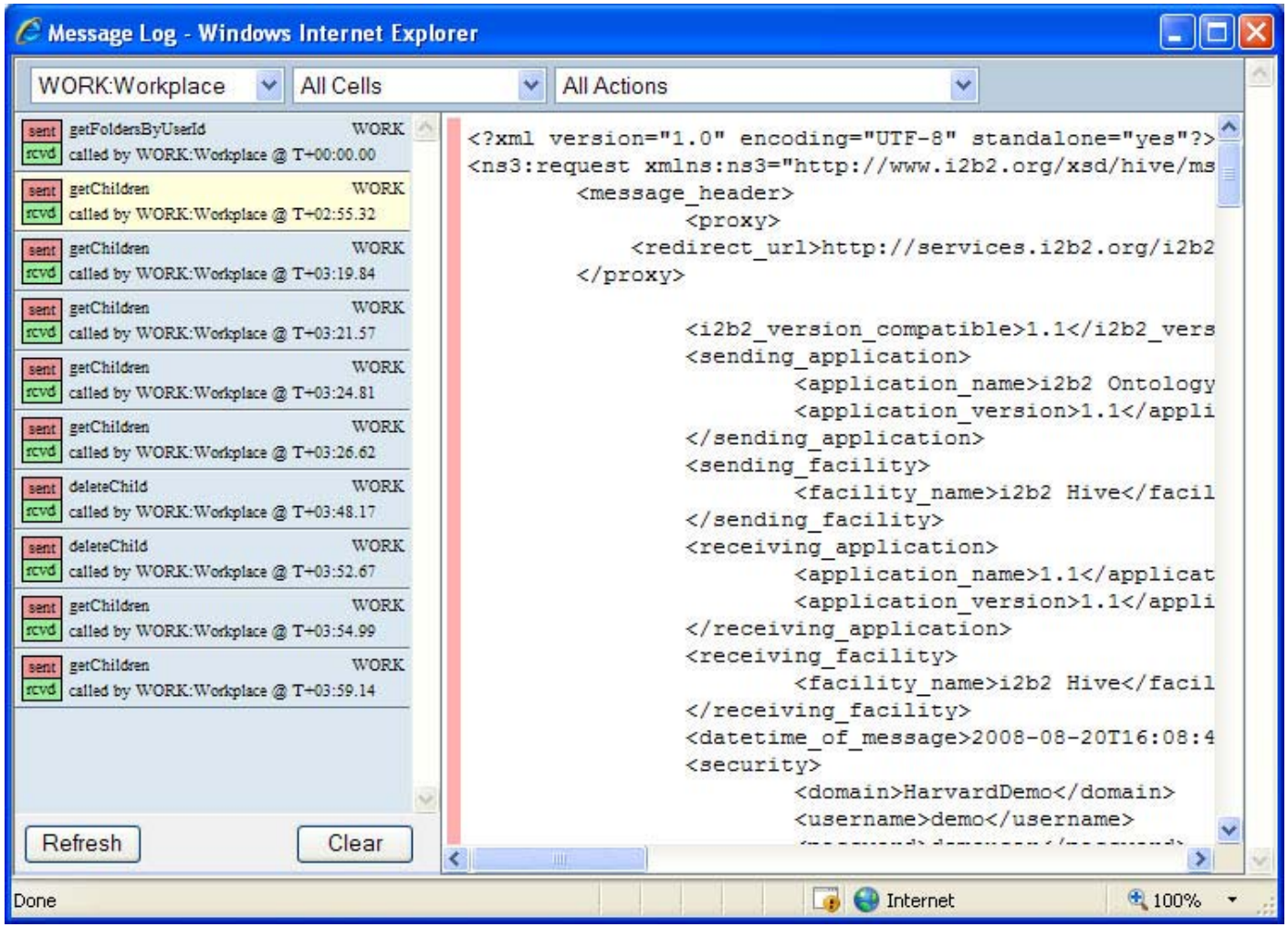
Compare multiple patient sets

Timeline Plugin



An example of visualization of patient data

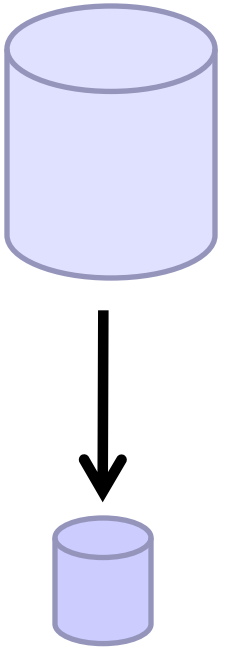
Debugging Tools



View all client-server i2b2 communication

Enterprise vs. Project

- Enterprise
  - All investigators, no IRB needed
  - All patients, aggregate only
  - Exploratory analysis, preliminary data
  - Web client
- Project
  - Small group with IRB protocol number
  - Selected patients, limited or identified data
  - Detailed data analysis
  - Web client or desktop client

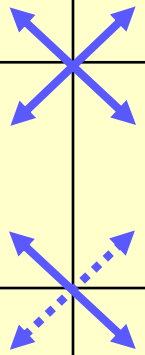


Web Client vs. Desktop Client

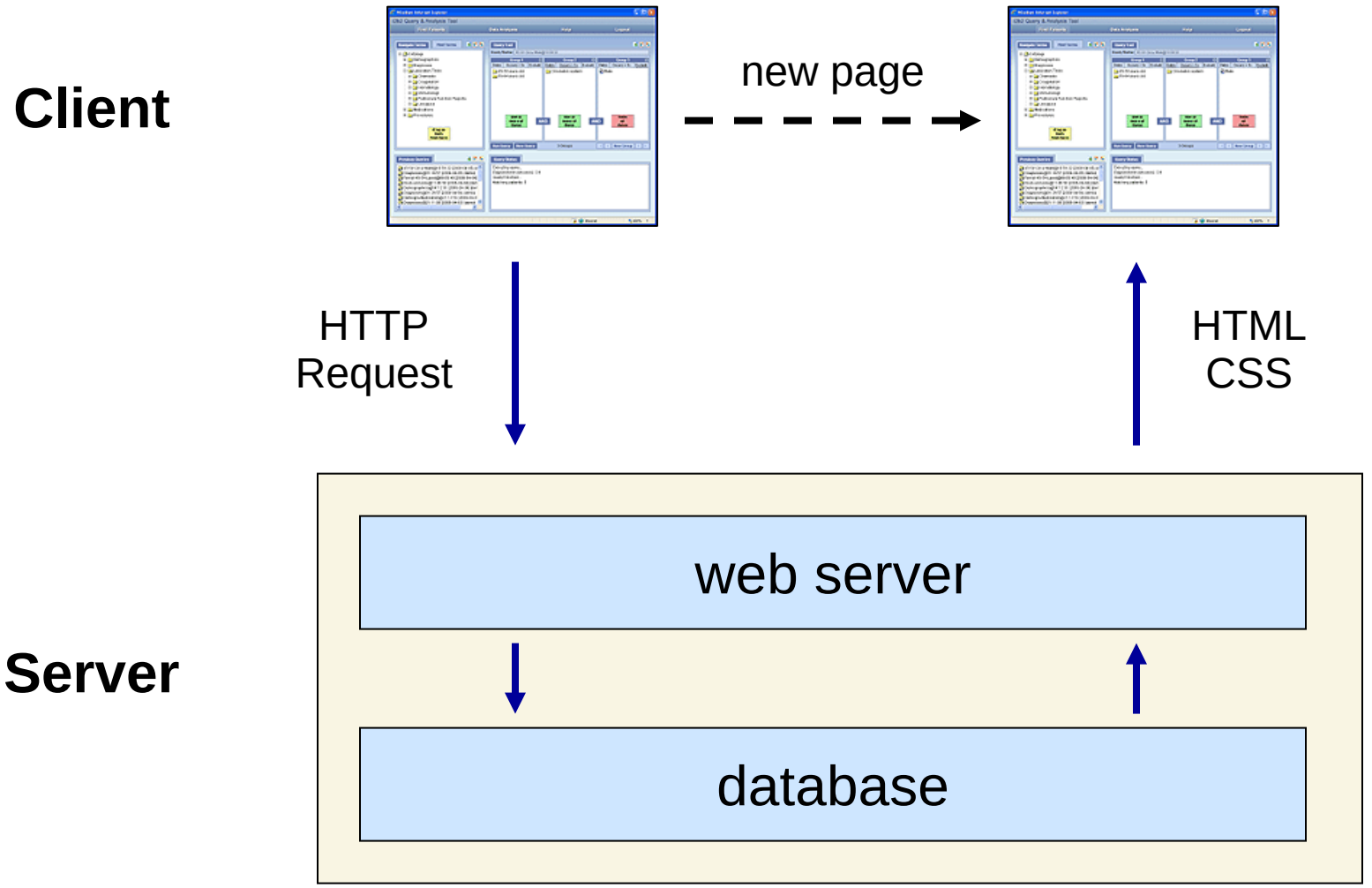
- Web Client
  - Written entirely in JavaScript, HTML, CSS
  - Same code runs on Windows IIS, Linux, etc
  - Easy to deploy and update
  - Designed for enterprise-based use
- Desktop Client
  - Written in Java, Eclipse plugin framework
  - Good for heavy client-side processing
  - Can use other Java/Eclipse plugins
  - Designed for project-based use

i2b2 Platforms

Client	Desktop App	Web Client
Web Services	Java	C# .NET
Database	Oracle	SQL Server



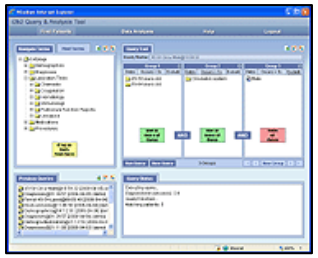
Traditional Web Application



AJAX

(Asynchronous JavaScript and XML)

Client



AJAX Request

XML Data

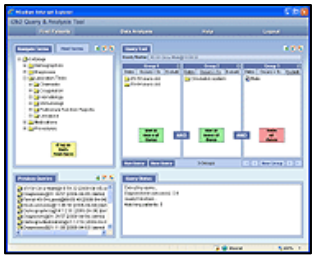
Server

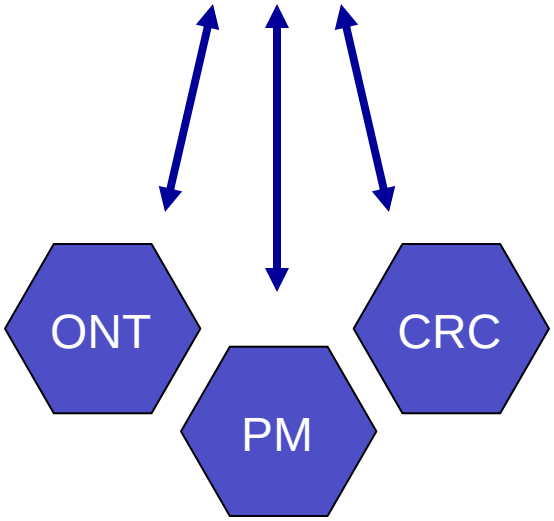
web services

database

Traditional Single Domain Website

Client



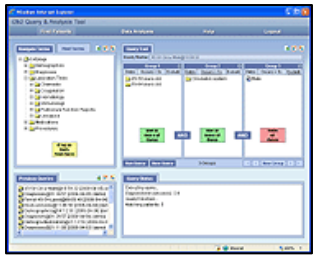


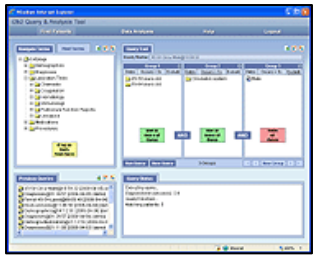
Server

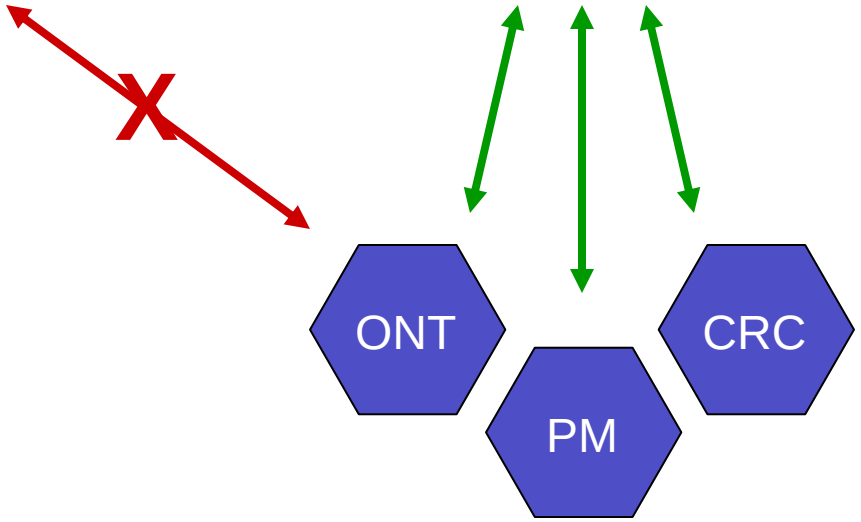
https://harvard.edu

Cross Domain AJAX (Not Possible)

Client





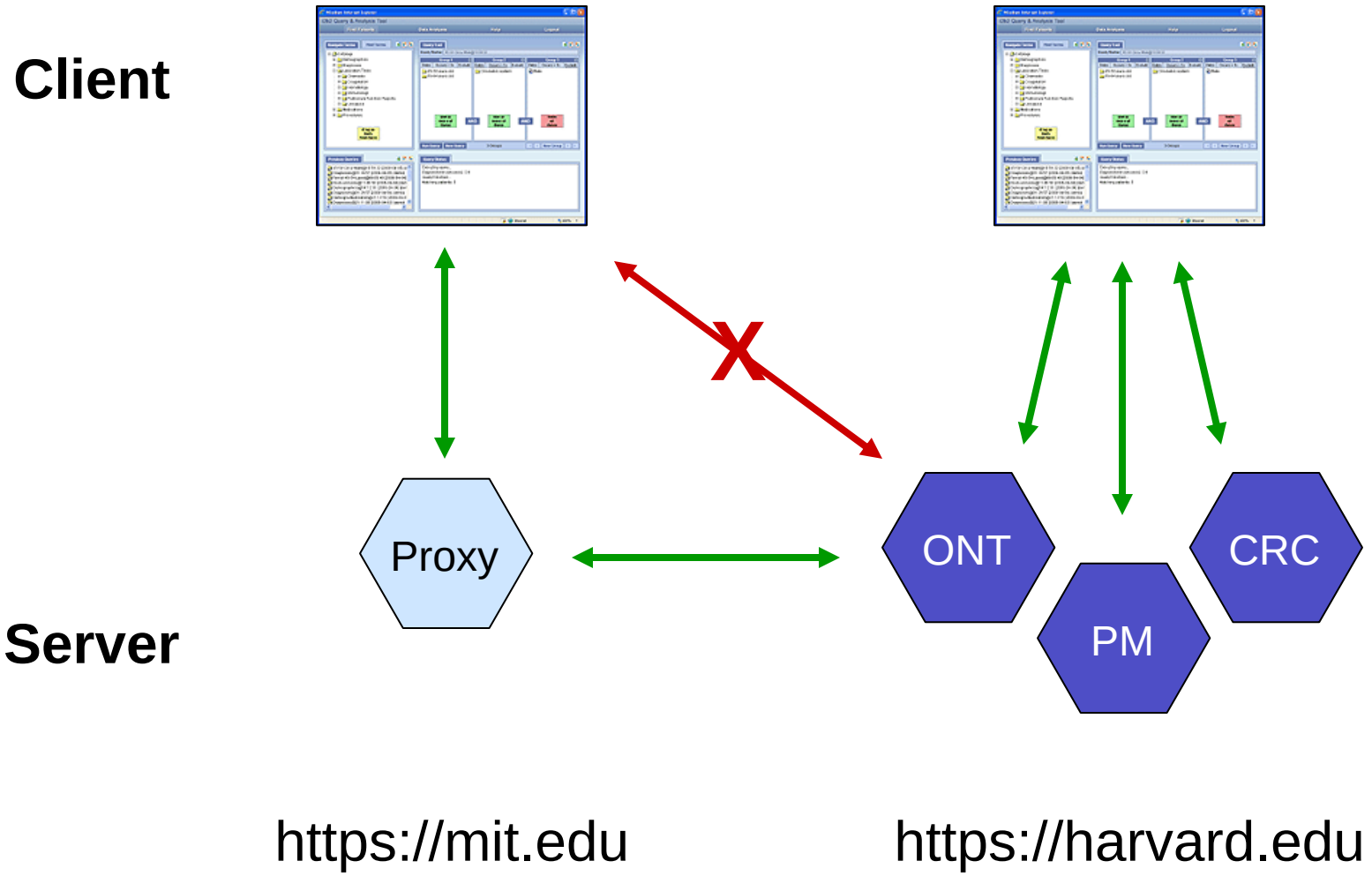


Server

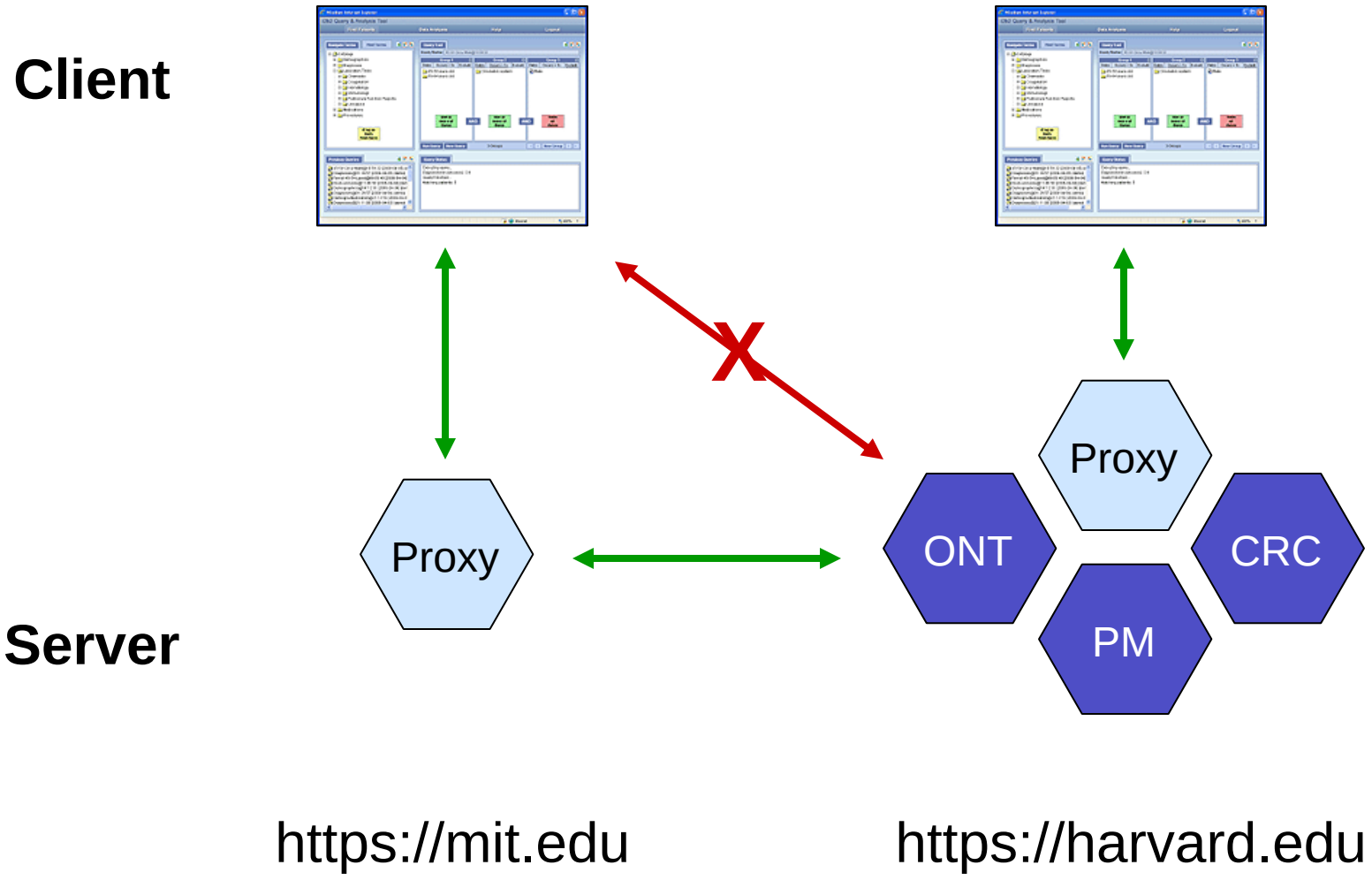
https://mit.edu

https://harvard.edu

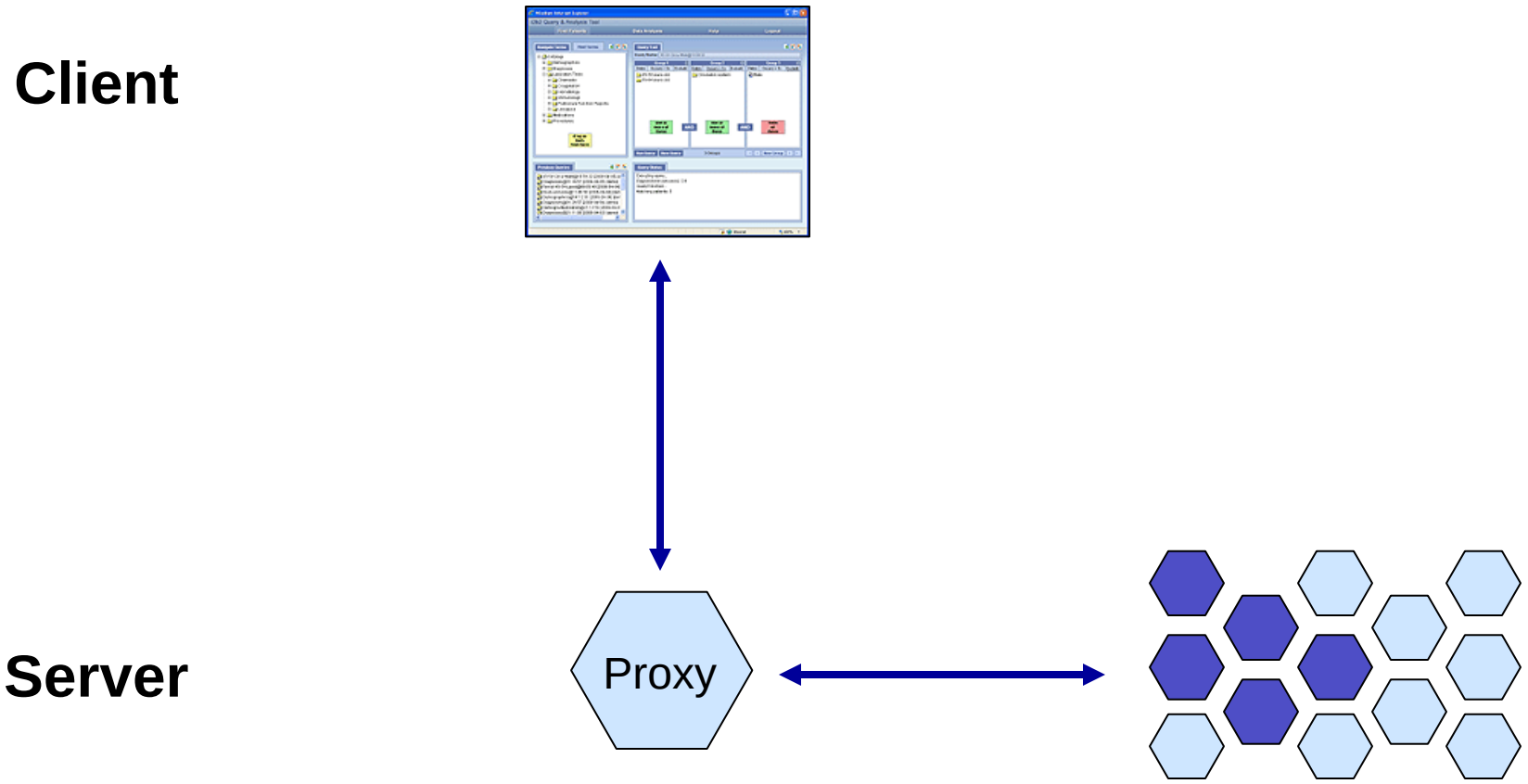
Cross Domain AJAX (With a Proxy)



Cross Domain AJAX (With a Proxy)



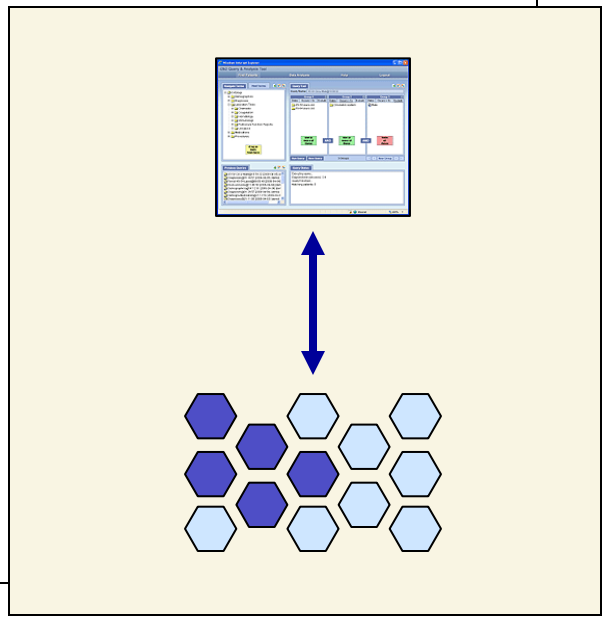
Cross Domain AJAX (With a Proxy)



i2b2 Messaging with the Proxy Cell

POST to **https://i2b2/ontology**

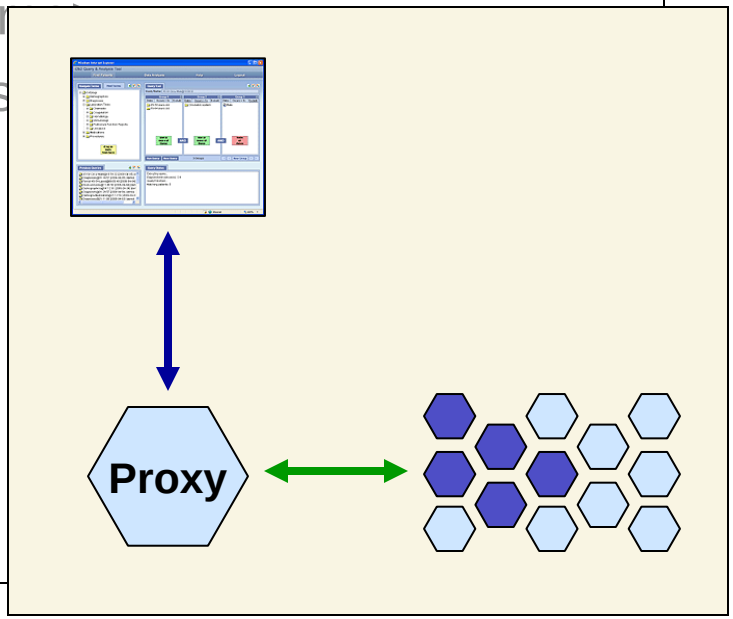
```
<request>
    <message_header>
        <security>
            <username>demo</username>
            <password>demouser</password>
        </security>
        <project_id>demo</project_id>
    </message_header>
    <request_header />
    <message_body />
</request>
```



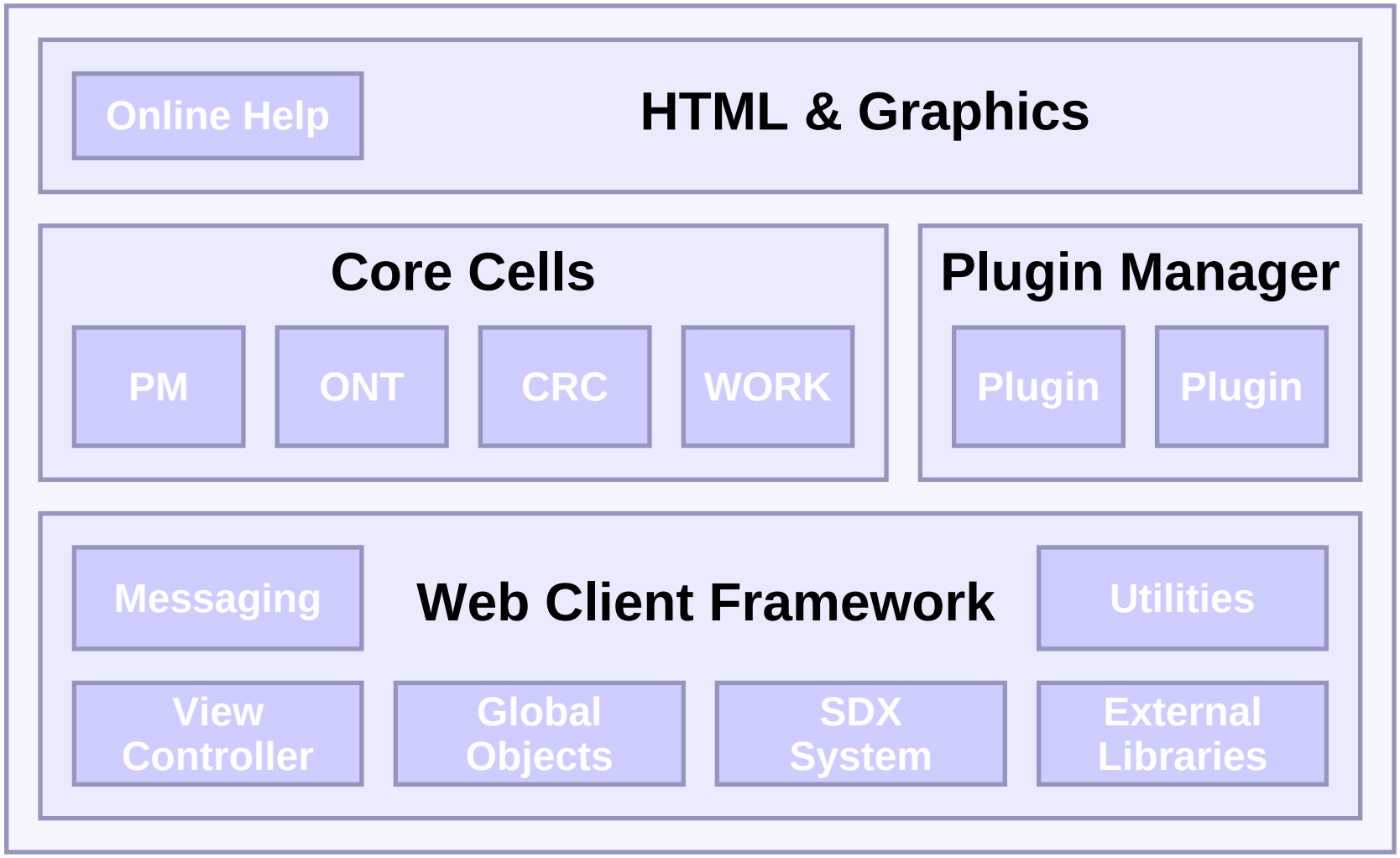
i2b2 Messaging with the Proxy Cell

POST to **https://localserver/Proxy**

```
<request>
    <message_header>
        <proxy>
            <redirect_url>https://i2b2/ontology</redirect_url>
        </proxy>
        <security>
            <username>demo</usernam
            <password>demouser</pas
        </security>
        <project_id>demo</project_id>
    </message_header>
    <request_header />
    <message_body />
</request>
```



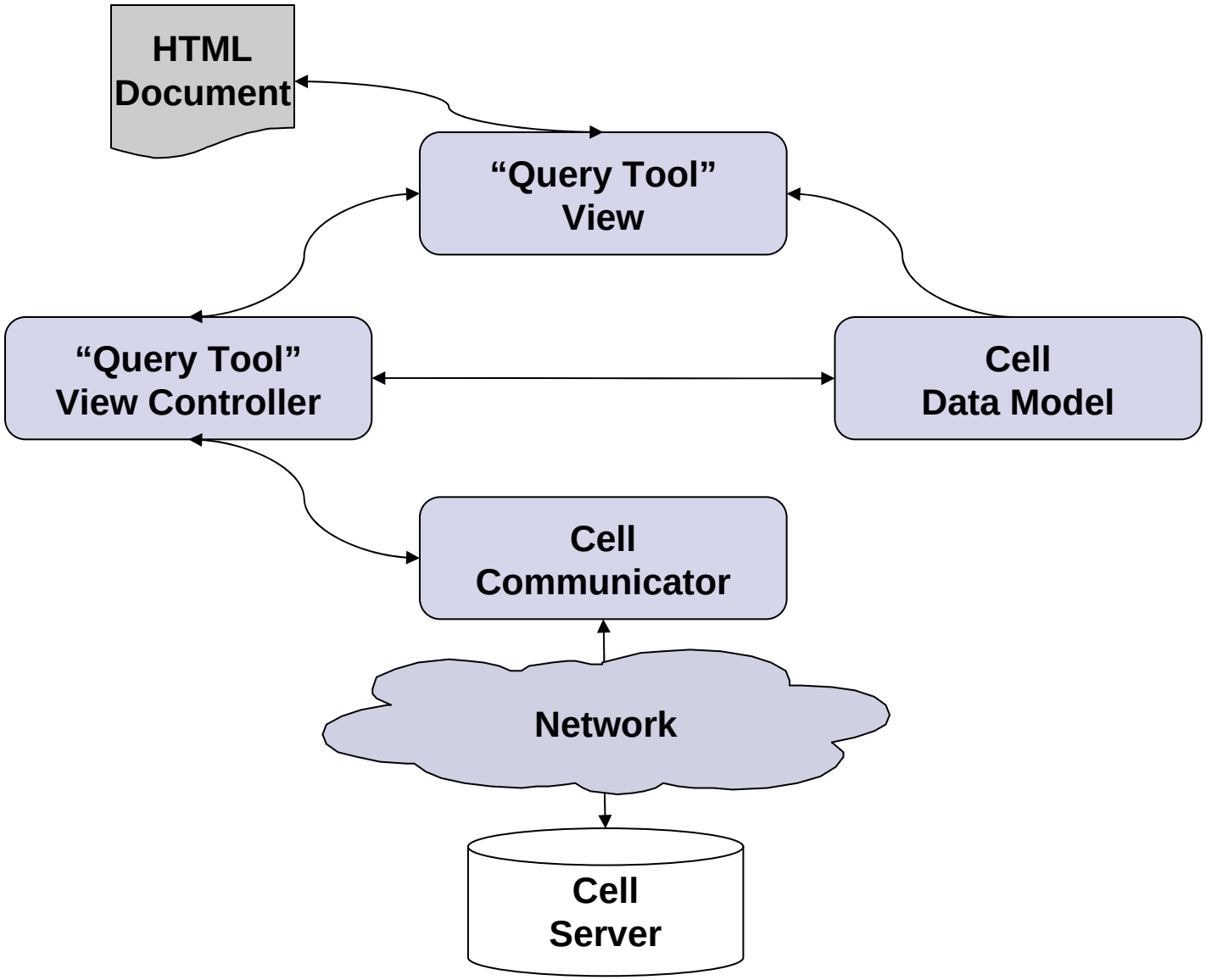
Client Architecture



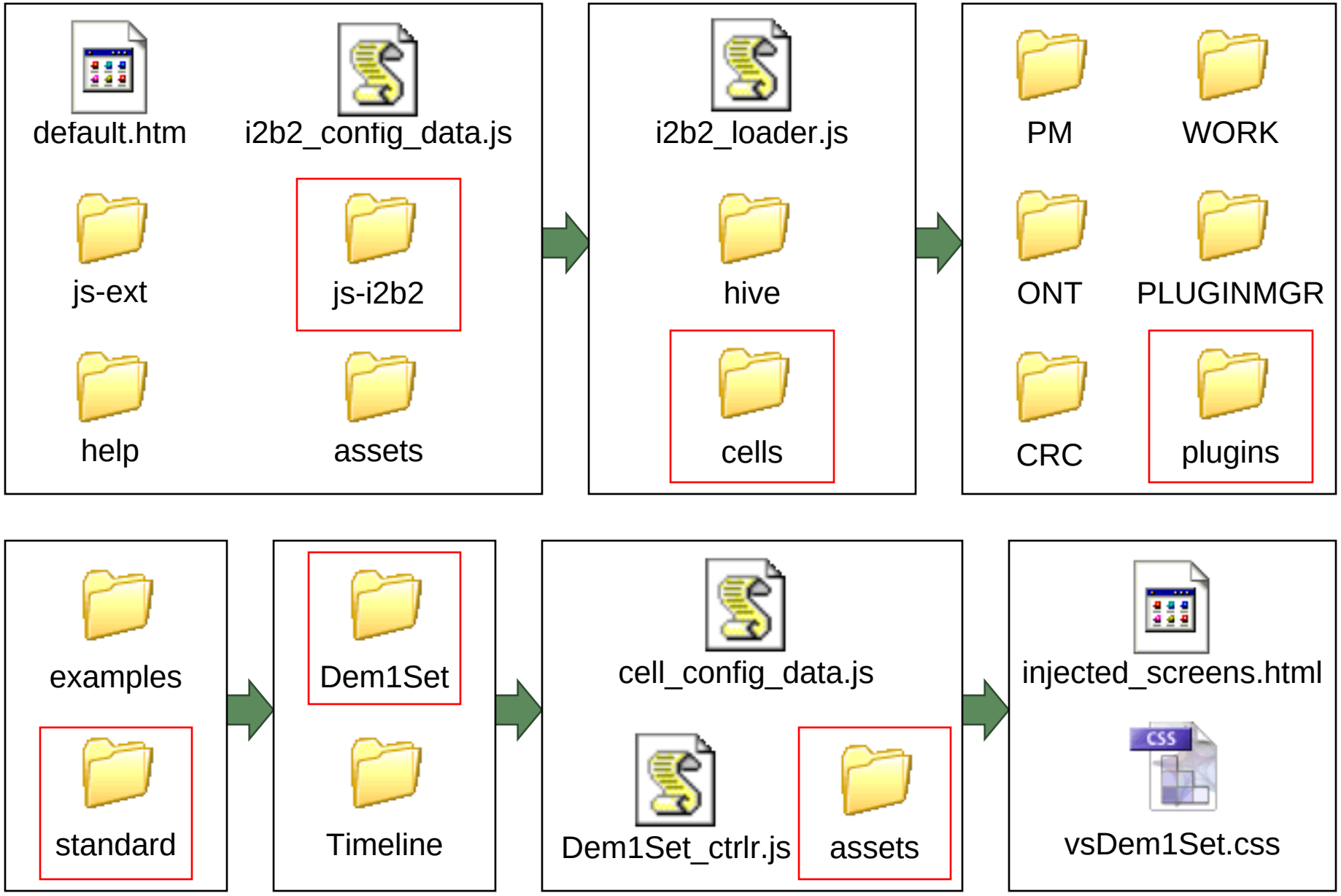
Framework Components

- Messaging – AJAX calls to server, debugging
- View Controller – Common GUI elements
- Global Objects – Shared variables
- SDX System – Standard data exchange (drag-drop)
- Utilities – XML parsing, browser compatibility
- External Libraries – YUI, Prototype, Firebug

Model-View-Controller



Directory Structure



i2b2_config_data.js

```
{
   urlProxy: "index.php",
   urlFramework: "js-i2b2/",
   //--------------------------------------------------------------
   // THESE ARE ALL THE DOMAINS A USER CAN LOGIN TO
   lstDomains: [
           { name: "localhost",
             domain: "demo",
             debug: true,
             urlCellPM: "http://127.0.0.1:9090/axis2/rest/PMService/"
           },
           { name: "i2b2.org (1.3)",
             domain: "HarvardDemo",
             debug: true,
             urlCellPM: "http://services.i2b2.org/PM/rest/PMService/"
           }
   ]
   //--------------------------------------------------------------
}
```

i2b2_loader.js

```
i2b2.hive.tempCellsList = [
          { code: "PM",
            forceLoading: true   <--- Must be true for PM cell
          },
          { code: "ONT"         },
          { code: "CRC"         },
          { code: "WORK"        },
          { code: "PLUGINMGR",
             forceLoading: true,
          },
          { code:    "Dem1Set",
             forceLoading: true,
             forceDir: "cells/plugins/standard"
          },
          { code:    "Timeline",
             forceLoading: true,
             forceConfigMsg: { params: [] },
             forceDir: "cells/plugins/standard"
          }
  ];
```

cell_config_data.js

```
{
   files:[
           "Dem1Set_ctrlr.js"
   ],
   css:[
           "vwDem1Set.css"
   ],
   config: {   // additional configuration variables
           short_name: "Demographics",
           name: "Demographics (1 Patient Set) - Simple Counts",
           description: "Demographics of a Patient Set.",
           category: ["celless","plugin","standard","demographics"],
           plugin: {
                       isolateHtml: false,  // do not use an IFRAME
                       isolateComm: false,  // use built-in communicator
                       standardTabs: true,  // use default tabs
                       html: {
                                   source: 'injected_screens.html',
                                   mainDivId: 'Dem1Set-mainDiv'
                       }
           }
   }
}
```

Regulatory Issues for Sharing EHR Data across Institutional Boundaries

A Survival Guide Based on Our Experience at Harvard's CTSA

Susanne Churchill, PhD

OUR MODEL

- Four + Major Teaching Hospitals
- Fiscally/Administratively Independent
- Historical Competitors
- Married at Gunpoint
- Concerned about Inappropriate Use of Data

MAJOR CONCERNS

- Protecting the Patients
- Protecting the Institutions
- Enabling the Investigators

****Inherent Conflicts****

Protecting the Patients

- IRB Review – Depends on Data Type
- HIPAA Authorization – Implied Consent….
- Trust

Protecting the Institution

- Patient Data = Proprietary Property?
  - View from Management
  - View from Clinicians
- Competitive Mischief
- Patient Backlash/Public Perception

Enabling the Investigator

- Easy to Use System
- Minimize the Bureaucracy
  - Registration
  - Oversight
  - Number of steps
- Support?

BIG QUESTIONS

- Who are the key institutional stakeholders and what is the process for gaining their support and approval?
- What kind(s) of data will you provide?
- How will you manage
  - Access to Data?
  - Limitations on Use of Data?
  - Insuring Integrity of Data Use?

Institutional Stakeholders

- Go wide and deep
- Go wide and deep
- Go wide and deep
- …….

- E.g., IRBs, Privacy Officers/Committees, General Counsel, Sr. Research Officers, Sr. Management, PR offices, Key Faculty, Curmudgeons, Skeptics – Think “Shuttle Diplomacy” – SINGLE MOST IMPORTANT STEP

What Kind of Data to Share?

- Aggregate Totals
- Limited Data Sets
- Identified Data

Access to Data

- Who can use the system to access data?
- What about secondary faculty appointees?
- What about non-clinically based investigators within your system?
- What about investigators who wish to collaborate with industry? With other academic health centers?
- Should query topics be reviewed and approved?

Limitations on Data Access and Use

- Should any type of data be reserved?
- Same rules for everyone?
- Should access be limited in time?
- Small samples?
- Harmonize available time spans?
- At what level do you authenticate users (institution vs central)?

Insuring Integrity of Data Use

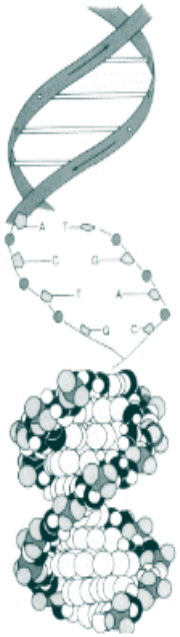
- Should you seek personal assurances of integrity?
- Should you review requests for “ethical” appropriateness?
- Should you archive and review all queries?
- Should each institution review use of name in publications?
- What are the penalties for violating rules?
- Reporting?
- How do you manage new policy issues?
- How will you handle patient concerns?

Etc.

- Intellectual Property
- Linking to Biospecimens
- Consent
- Governance for discipline-specific Networks
- Scaling from Local to National Networks

i2b2/SHRINE Team

- Zak Kohane, Andy McMurry, Griffin Weber, Shawn Murphy, Susanne Churchill
- www.i2b2.org
- schurchill@partners.org



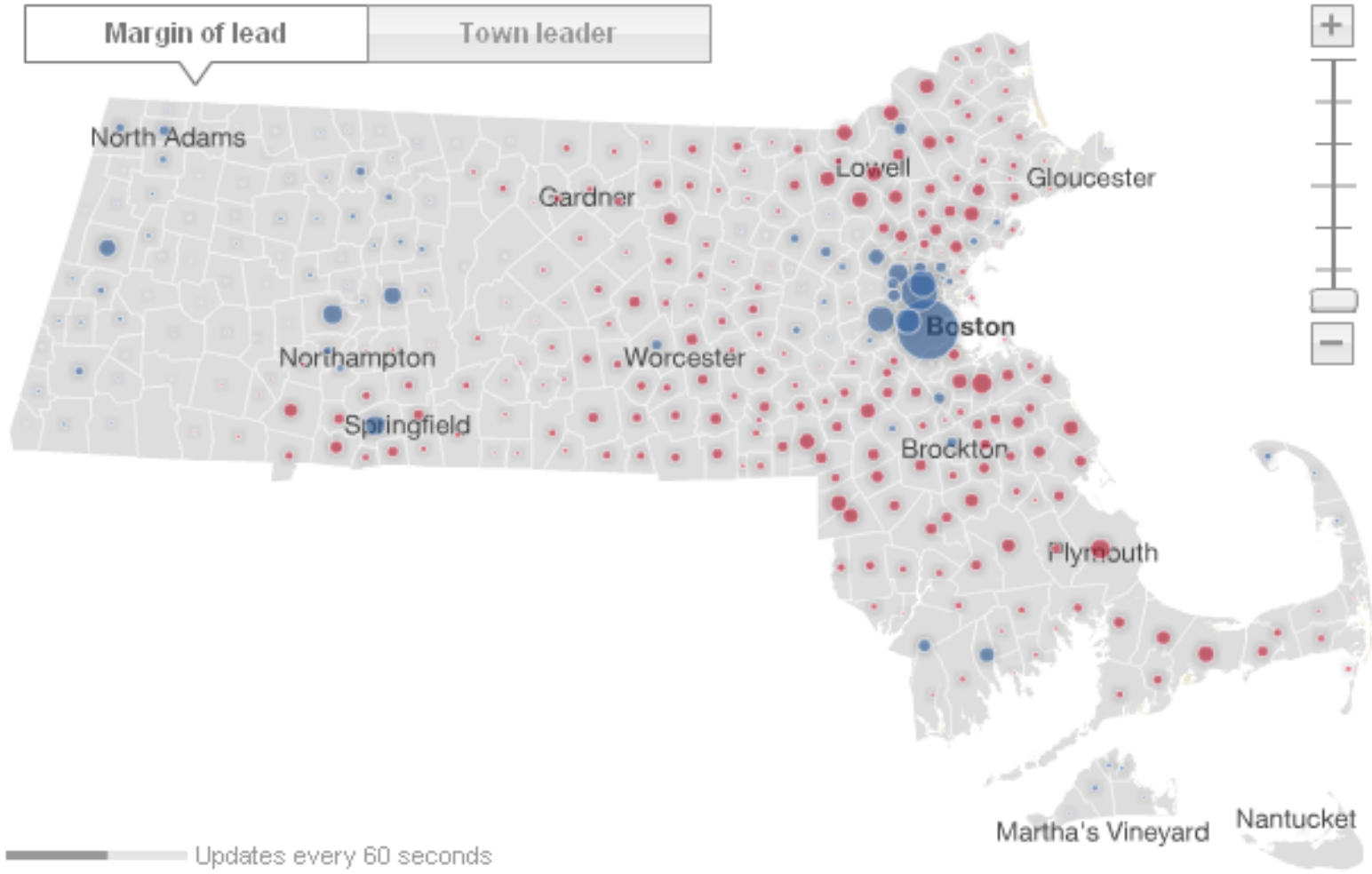
Shared Health Research Information Network

Andrew McMurry, MS
Informatics Team Lead and Architect,
Harvard Medical School Center for BioMedical Informatics

SHRINE: Obtaining Samples for Research

- Sufficient cohort sizes
  - Often cited concern for biomedical research is obtaining cohorts of sufficient size. Often requires aggregating across locations where patients receive care or where investigators do research.
- Reproducible study results
  - Many papers raise alarms that even in studies with sufficient sample sizes they may still have strong “population sampling” biases.

Sampling Bias Example 1



http://www.boston.com/news/special/politics/2010/senate/results.html

i2b2/shrine National Participation (growing rapidly)

SHRINE: Current Status

- Can we capitalize on the critical mass of i2b2 deployments? YES

- Simply being able to search for patient populations across hospitals is a major first step
  - Demographics : HITSP C32 / HL7
  - Diagnoses : ICD9-CM
  - Medications : RxNorm
  - Lab Procedures : LOINC

  Selecting national standards that reflect the reality of data collected at the point of care

East Coast SHRINE: Harvard Deployment

- 5.6M+ Patients
- 890M+ Facts
- 3 Ontology Categories, 18k+ Terms
- 7,500+ Potential Users
- 4 IRBs
- 4 major competing hospitals, 3 sites
- Partners Health Care (BWH, MGH)
- Children's Hospital Boston
- Beth Israel Deaconess Medical Center

West Coast SHRINE: Cross-Institutional Clinical Translational Research

- **3 CTSAs in demonstration network**
  - University of Washington
  - UC Davis
  - UC San Francisco
- **Primary focus is diabetes**
- **IRB approvals secured for all sites**
- ***Nick Anderson et all***
  with support from Recombinant Corp

National SHRINE: Pediatric Registry

- 60 sites!
- **Grand Opportunity (GO) Stimulus Award**
- **National registry of rare pediatric disorders**
- **Childhood Arthritis & Rheumatology Research Alliance**
  http://www.carragroup.org/

How to implement SHRINE

- **OBTAIN Institutional buy-in**
- **ETL into your locally controlled i2b2 Data Repository**
- **MAP your local terminology codes (software provided)**
- **LINK your local user authentication system**
- **DEPLOY shrine on a locally controlled server**

How to implement SHRINE

- ETL into your locally controlled i2b2 Data Repository **(Or use your existing i2b2 instance)**
  - Review the STAR schema
    - Patient (very basic)
    - Concept (ICD, Medication, etc)
    - Encounter (hospital visit)
    - Observation Fact (any recorded fact and value)
    - “Dimension” tables

How to implement SHRINE

- MAP your local terminology codes (software provided)
  - SHRINE concept = Local concept
    - Think “Male = M” instead of complex XML/RDF models
  - A single SHRINE concept can relate to multiple local concepts
  - Demographics mappings can be done in a day
  - Diagnosis are easy (ICD9-CM)
  - Medications are harder (ingredients)
  - Labs are hardest due to lab values

How to implement SHRINE

- LINK your local user authentication system
  - Local user auth stays OFF the public internet
  - Your webclient asks “username and password valid” and is returned a list of URLs or invalid logon
  - Probably easiest to “wrap” an existing service at your hospital than doing a daily copy

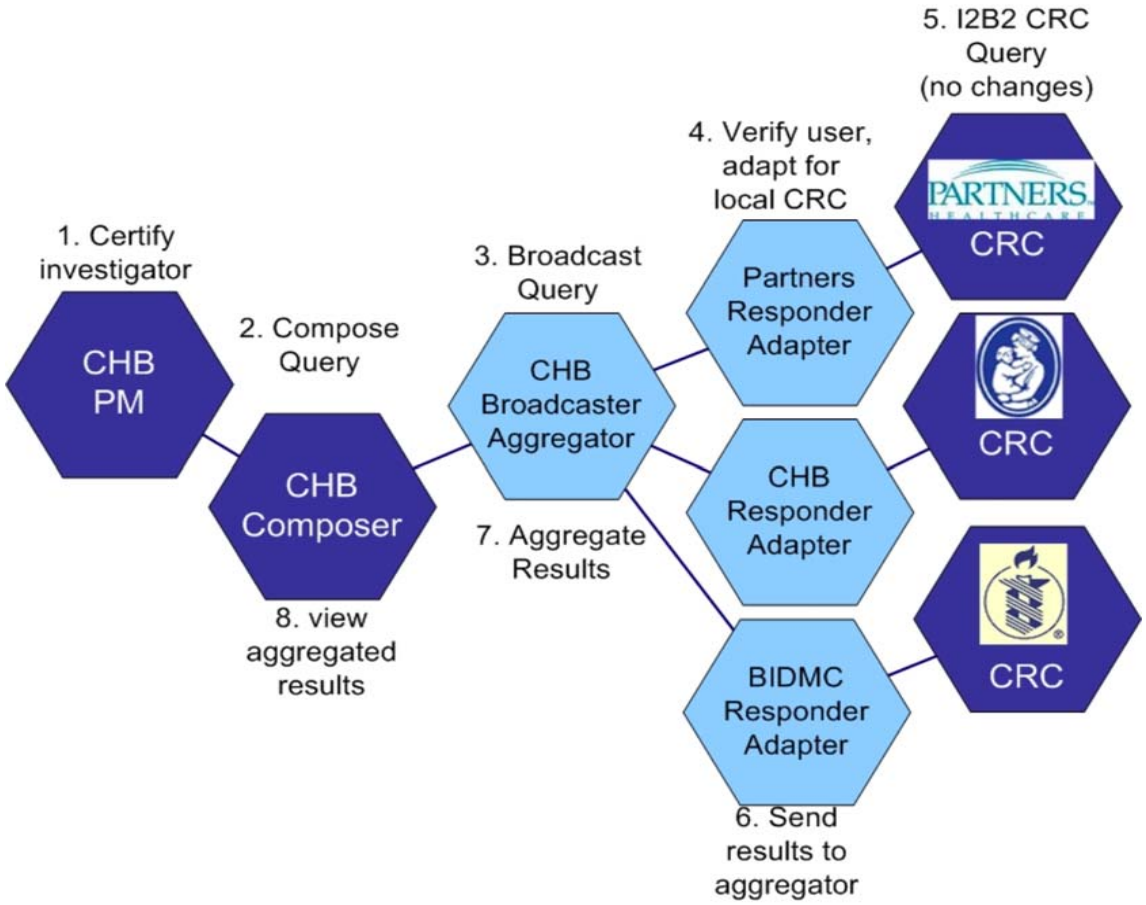
How to implement SHRINE

- DEPLOY shrine on a locally controlled server
  - Can deploy the broadcaster-aggregator and adapter on the same machine or on separate machines. Your preference.
  - SHRINE components served up from cheap commodity hardware

How SHRINE works

1. Broadcast queries across a network of i2b2 clinical databases
2. Aggregate query results across all sites in near-real time
3. Each site maintains an autonomous database
4. Patients remain de-identified

SHRINE : how it works



SHRINE plugs into an existing i2b2 hive

Previous work leading up to SHRINE

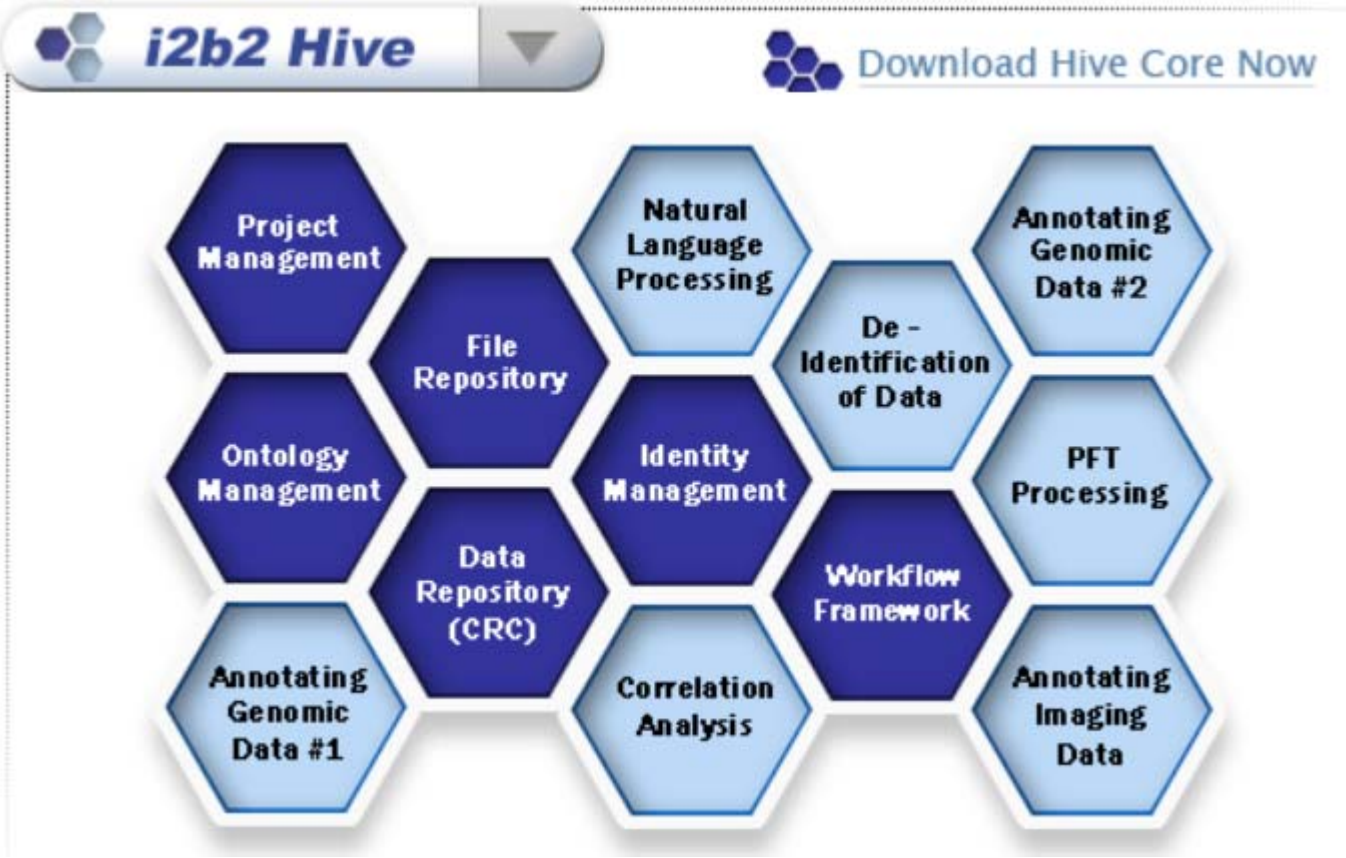
- i2b2 → Clinical Data Repository
  - Zak Kohane, Shawn Murphy et all
  - Turning everyday clinical encounters into huge research cohorts
  - Star-schema database with associated ontology

- SPIN → cross hospital query model
  - Federated query of autonomous/independent hospitals
  - Abstracted away the differences of each data repository enabling *on the fly query translation*

Overly simplified : i2b2 clinical database + spin federated query = shrine

I2b2 features leveraged for SHRINE

Webclient, Project Management, Ontology, Data Repository



I2b2 features leveraged for SHRINE: **Webclient**

1. Webclient speaks to the broadcaster aggregator as if it were another i2b2 Clinical Data Repository

2. Minor enhancements to support additional auditing/monitoring of SHRINE queries

Griffin Weber, Nicholas Benik

I2b2 features leveraged for SHRINE: Project Management Cell

1. Authenticate local investigators

2. Pointers to “Cells in the Hive”
   - Location for the **Ontology**
   - Location of the **Data Repository**
   - URLs for all i2b2 cells

I2b2 features leveraged for SHRINE: Ontology Cell

Handles deep “nested” hierarchies like
Diagnoses\Neoplasms\Cancer of breast

1. Demographics
   - Age
   - Gender - HL7 Administrative Gender
   - Language - ISO 639-1
   - Marital Status - HL7 Marital Status
   - Religion - HL7 Religious Affiliation
   - Race and Ethnicity - CDC Race & Ethnicity Code Sets
2. Diagnoses - ICD-9-CM and CCS hierarchy
3. Medications - RxNorm and NDF-RT hierarchy
4. Lab Tests (demo) - LOINC

I2b2 features leveraged for SHRINE: Data Repository Cell

✓Star Schema design concurrently supports multiple ontologies (local ontology and SHRINE ontology)

✓SHRINE works with CRC v1.3 and v1.4

✓Use your existing i2b2 deployment, no code or data changes required!

Shawn Murphy et all.

SPIN features leveraged for SHRINE

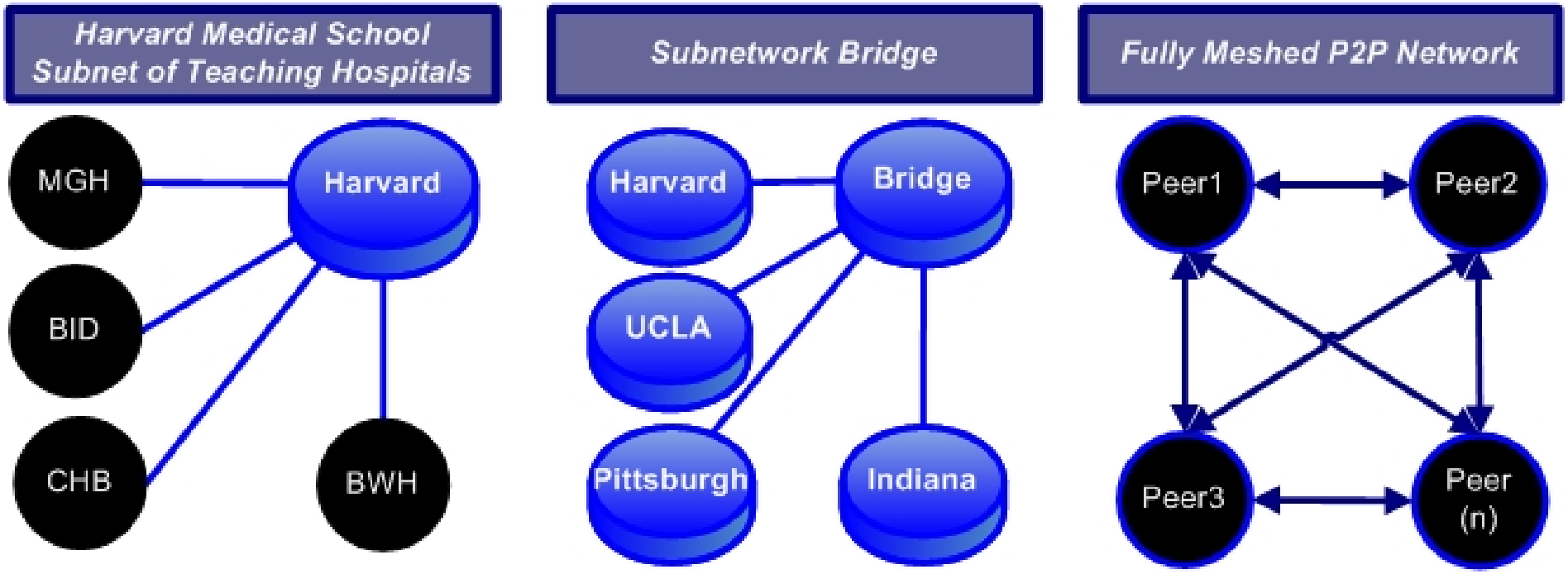
- Local database control to engender participation
- Query Broadcaster-Aggregator
- p2p Trust Model with De-Identification
- Coming in 2010: Limited Data Sets
- Coming in 2010: Search for Biospecimens

SPIN features leveraged for SHRINE
Model of Autonomy

- **SPIN had already proven successful for linking up de-identified pathology reports across independent HIPAA covered entities**

Brigham & Women's Hospital*
Beth Israel Deaconess Medical Center*
Cedars-Sinai Medical Center
Dana-Farber Cancer Institute*
Children's Hospital Boston*
Harvard Medical School*
Massachusetts General Hospital*
National Institutes of Health
National Cancer Institute
Olive View Medical Center
Regenstrief Institute
University of California at Los Angeles Medical Center
University of Pittsburgh Medical Center
VA Greater LA Healthcare System

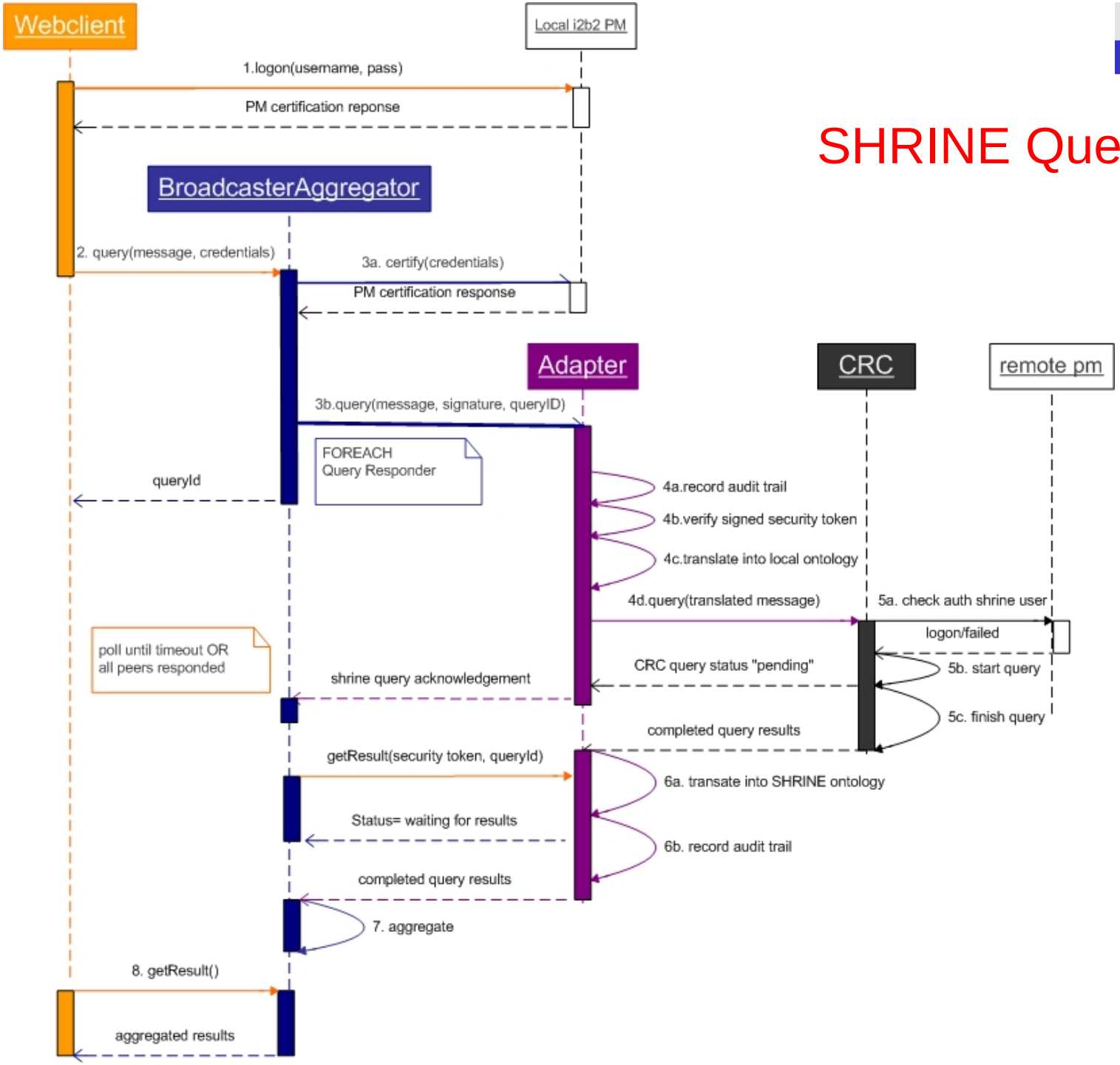
SPIN features leveraged for SHRINE: Query Broadcaster-Aggregator



SPIN features leveraged for SHRINE: p2p Trust Model with De-Identification

- W3c standard security with patient de-identification
  - Well beyond SSL alone
  - W3C official libraries for digital signatures
  - Encryption of result concents
  - IP restricted firewalls
  - Anonymize patient counts

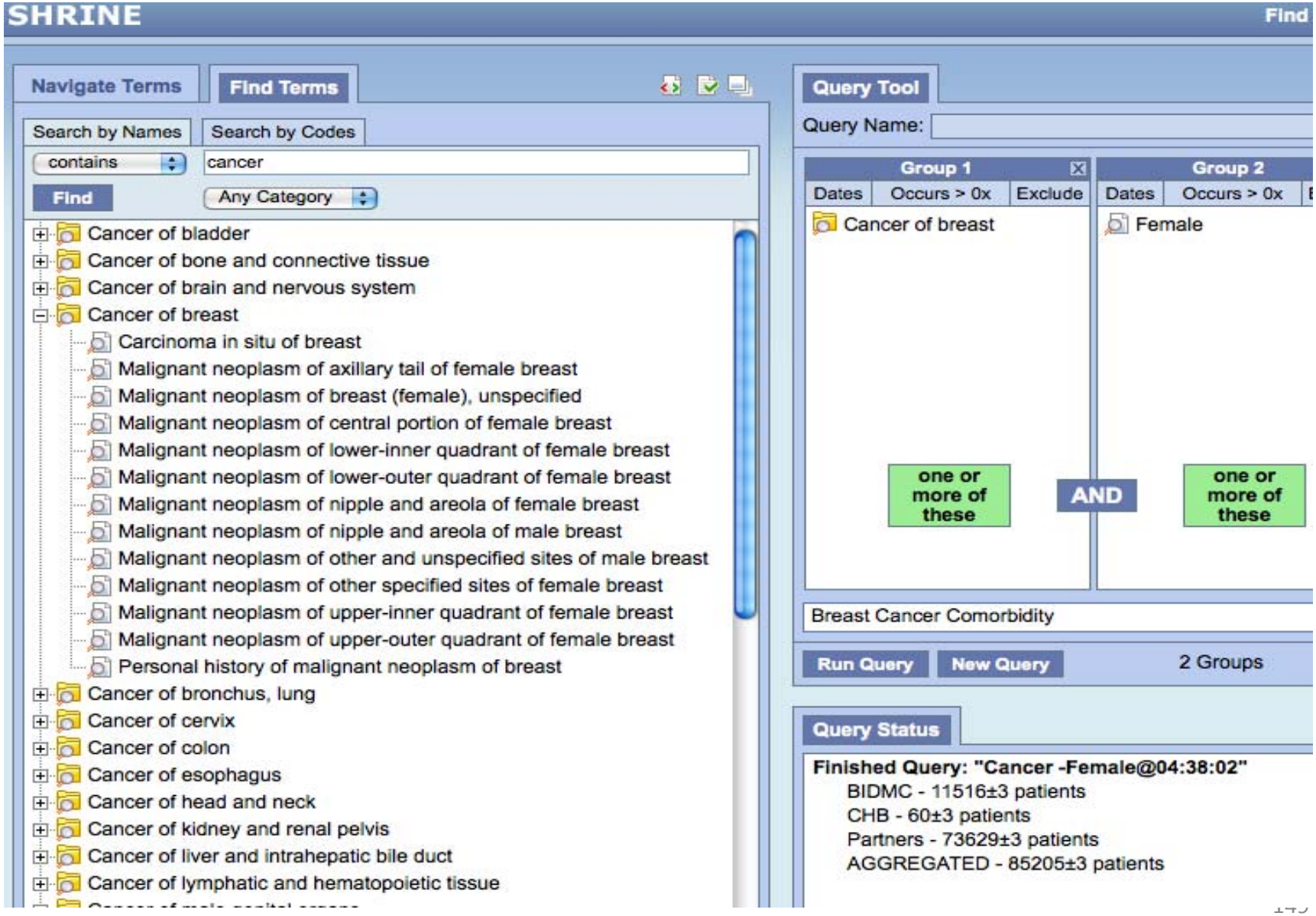
SHRINE Query Sequence



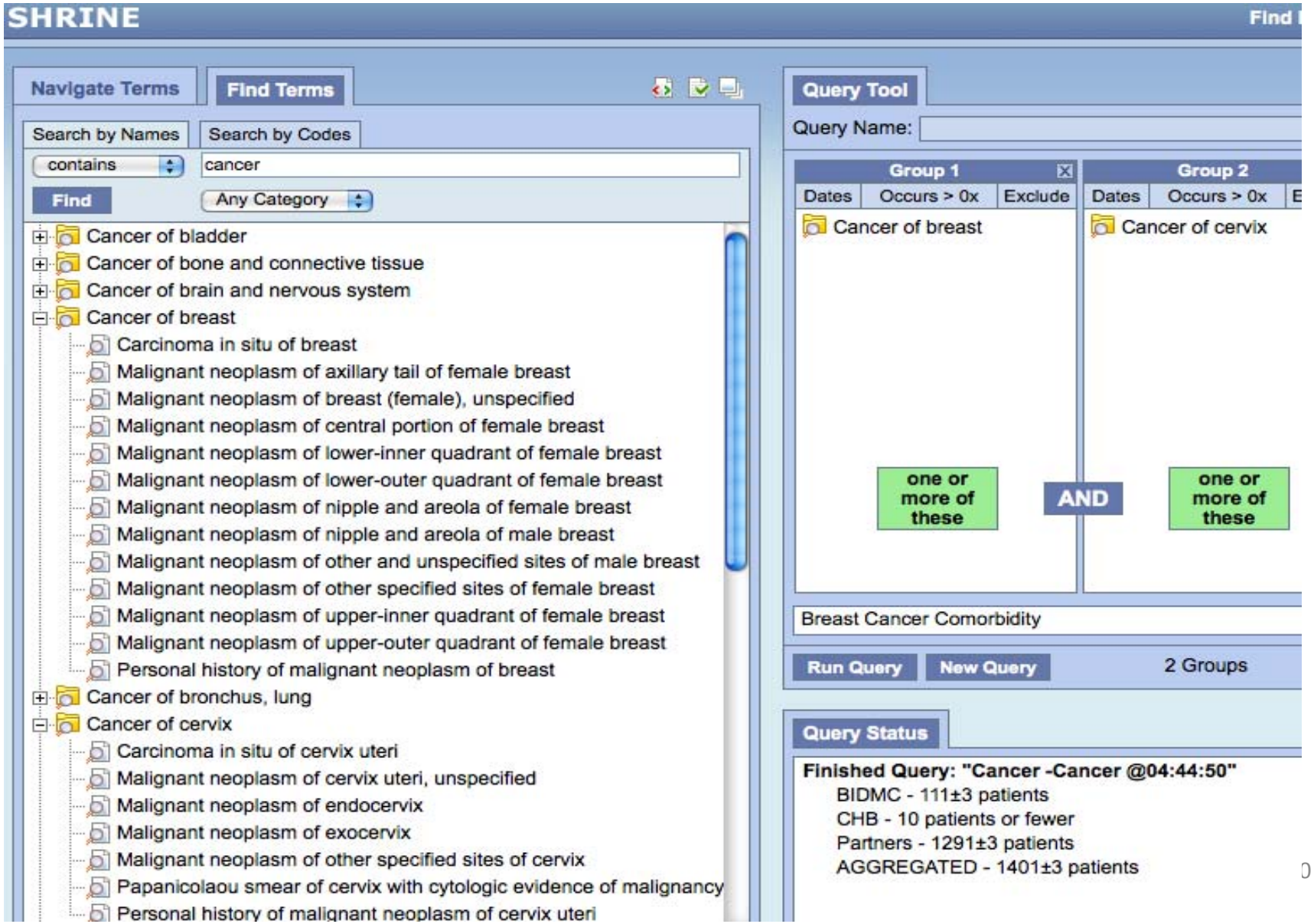
SHRINE queries (Actual Patient Counts)

- **Comorbidity of Breast and Cervical Cancers**
- **Acute Myocardial Infarction with Gender Breakdown**
- **Pediatric Rhematoid Arthritis (rare)**
- ***Countless more secondary uses of electronic medical records…***

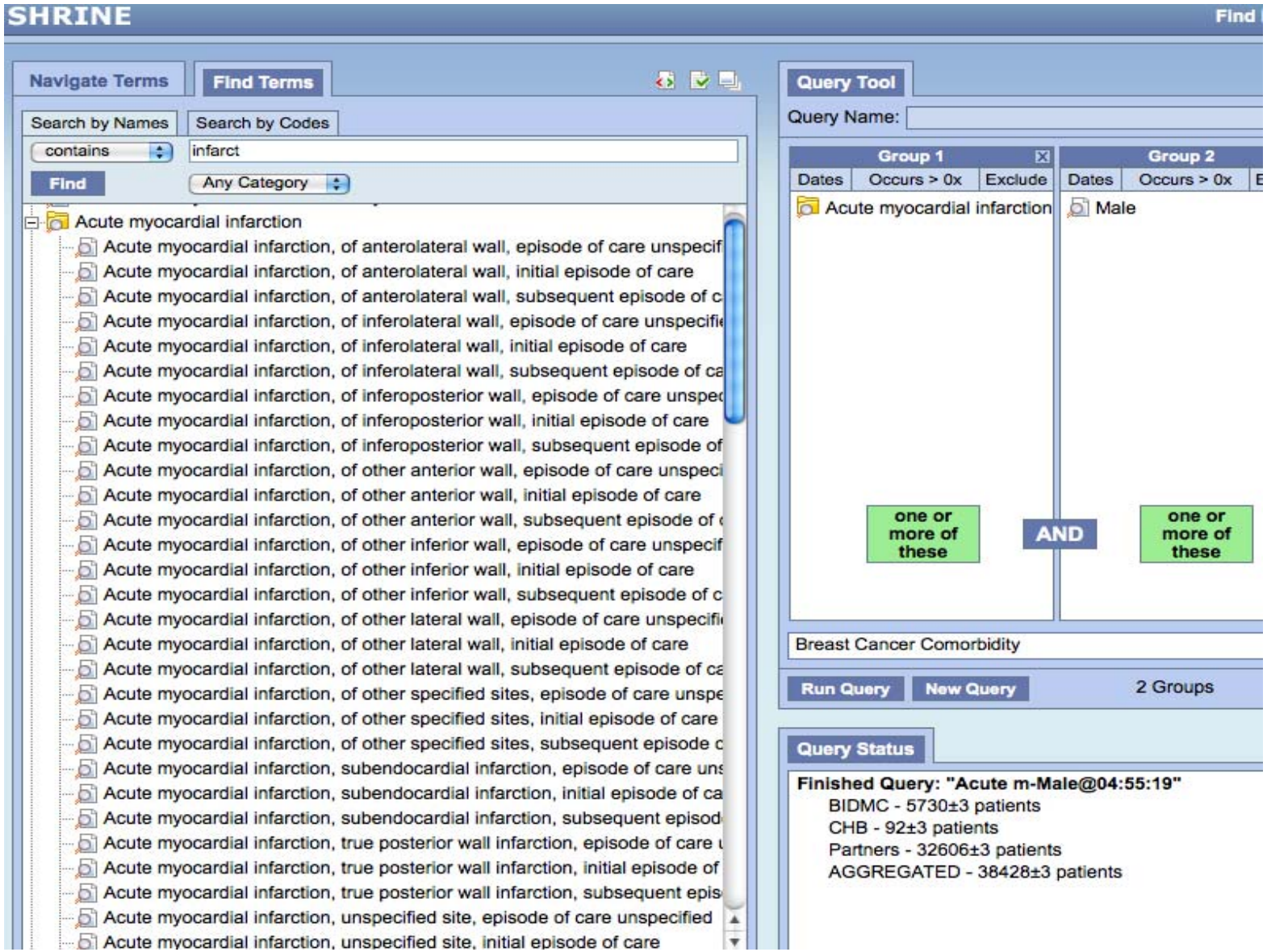
Common: Breast Cancer



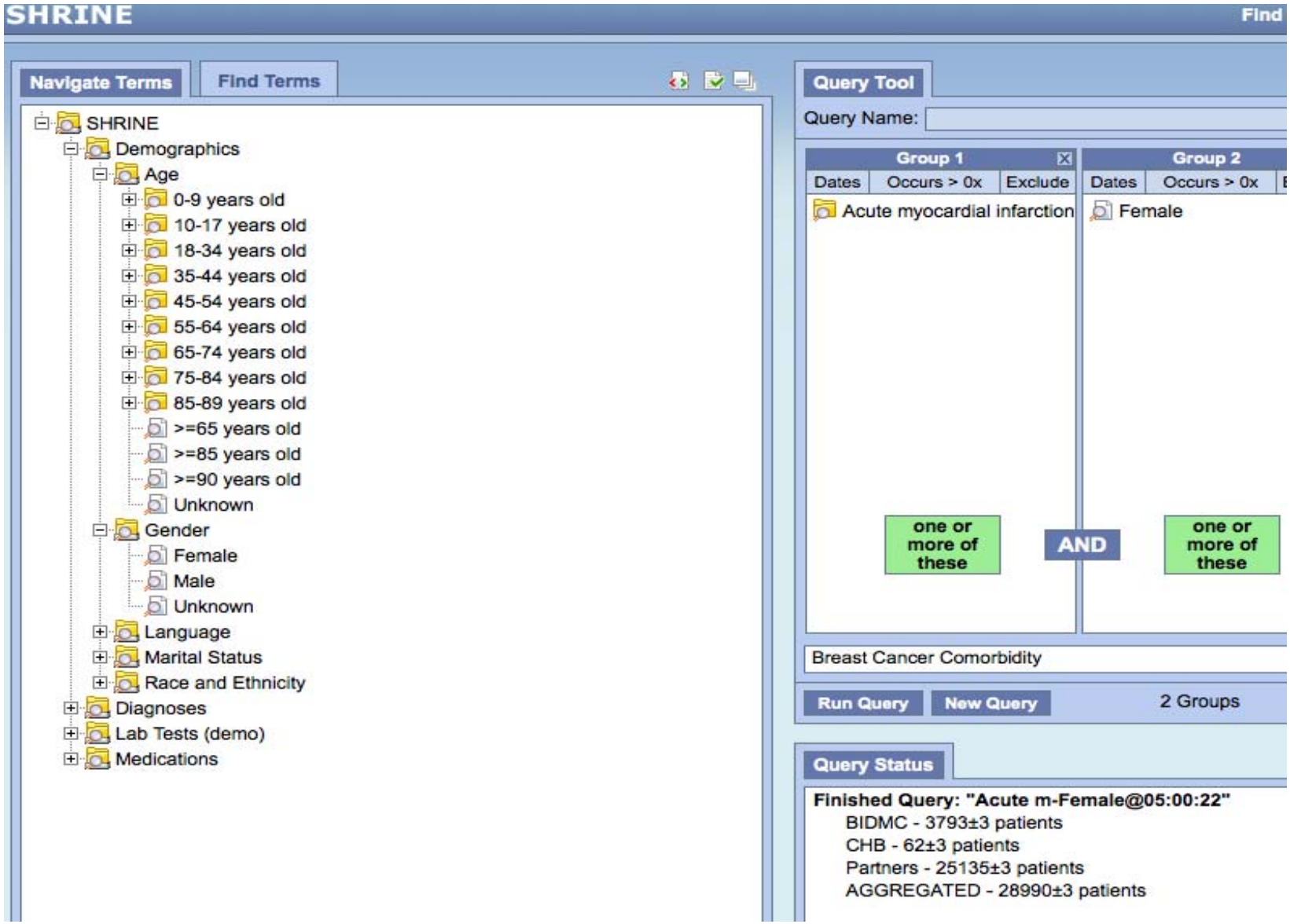
Comorbidity: Breast and Cervical Cancer



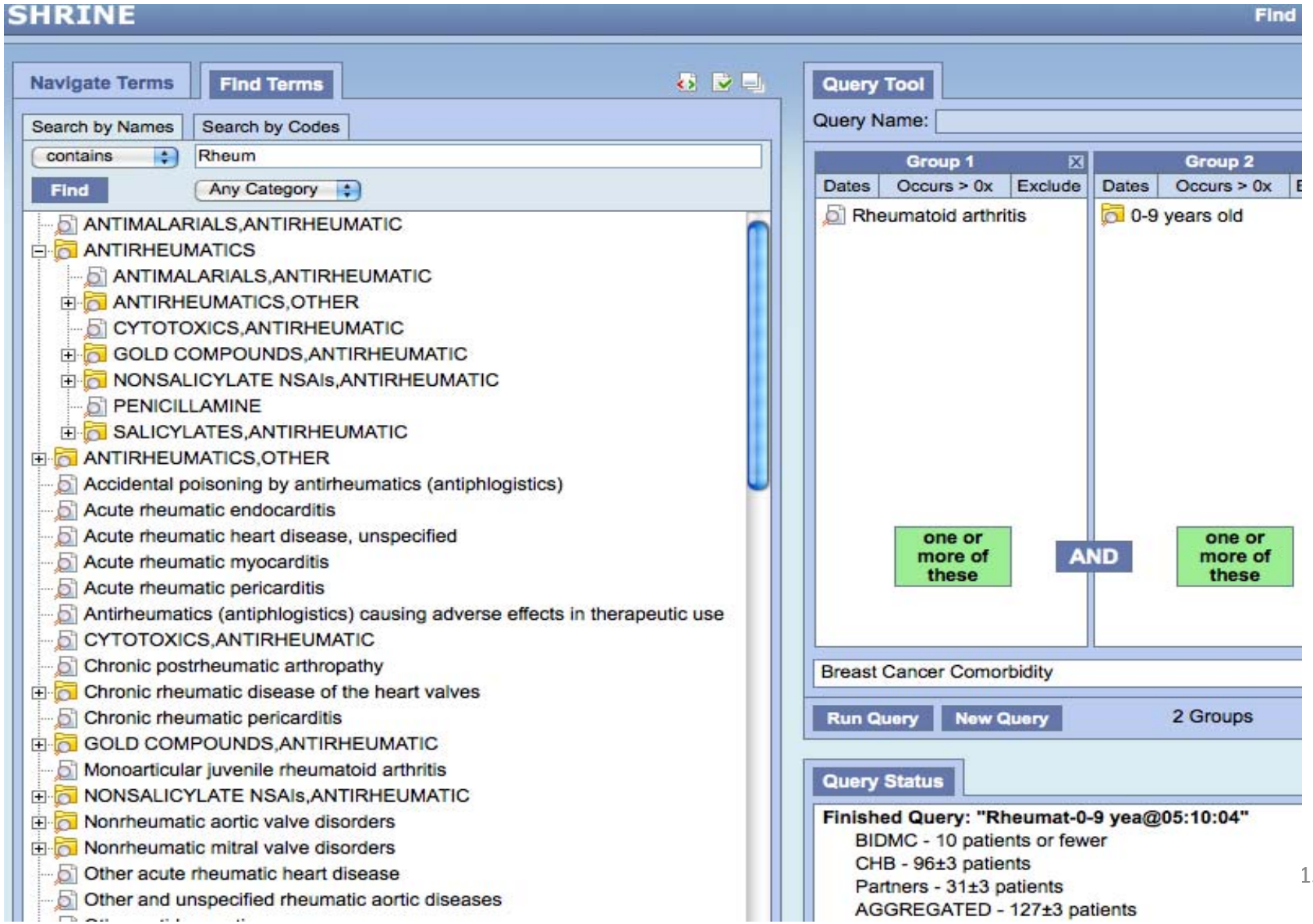
Relative counts by gender (Male Acute MI)



Relative counts by gender (Female Acute MI)



Rare: Children with Rheumatoid Arthritis



SHRINE: Current Work

- Making SHRINE rapidly deployable to 60 sites
- Training new developers
- Documentation and source code refactoring
- Launching SHRINE Open Source effort

SUMMARY

- ✓ **Rapid deployment of i2b2 technology across the country**
- ✓ **SHRINE enables federated searches of existing i2b2 deployments without modifying the software or databases**
- ✓ **SHRINE “sub-networks” exist both at Harvard and the west coast with major new efforts underway per Recovery Act funds**
- ✓ **Even simple searches can power many translational studies**
- ✓ **Current work focus = rapid deployment and open source**
- ✓ **Headed towards supporting Limited Data Set and Biospecimen exchange in 2010**

Acknowledgements: Core SHRINE team

- Zak Kohane (SHRINE PI / HMS)

- Joanna Brownstein (Project Manager, HMS)
- Sussane Churchill (I2B2 Executive director)
- David Hardwick (QA, HMS)
- Shaun Kelly (QA, HMS)
- Doug Macfadden (HMS CBMI IT Director)
- Charles McGow (Developer, Children's)
- Andrew McMurry (Architect / HMS)
- Mike Mendis (I2b2 /partners)
- Shawn Murphy (I2B2 CRC / partners)
- Danny Shaw (Knowledge Officer, Children's)
- Matt Sullivan (Developer, HMS)
- Phillip Trevor (Ontology / HMS)
- Nich Wattanasin (I2b2 /partners)
- Griffin Weber (HMS CTO / bidmc)

Acknowledgements: External Collaborators

- West Coast CICTR
  - Nick Anderson, PI U Washington
  - Kent Anderson, Co Director UC Davis
  - Davera Gabriel, Ontologies UC Davis
  - Michael Kamerick, Director UC San Francisco
  - Rob Wynden, Ontologies UC San Francisco
  - Many more!
- CarraNet
  - 60 institutions, special thanks to Mark Natter (PI Informatics)
- Recombinant Data
  - Andy would like to specially thank to Aaron Mandel for getting the software deployed on the west coast and Matvey Palchuck for all preliminary work on the SHRINE Core Ontology
- COUNTLESS INDIVIDUALS CONTRIBUTED TO THE MAKING OF SHRINE, WE THANK ALL OF YOU!