



T02: Ontology services for translational research in the i2b2 Workbench

Shawn Murphy MD, Ph.D.

Nigam Shah MBBS, Ph.D.

Lori Phillips MS



Tutorial Guiding Use Case

- A hospital has converted from ICD9 to SNOMED-CT – how can the i2b2 ontology adapt to handle this situation?
 - SNOMED-CT ontology needs to be created in an i2b2 format
 - We will learn about the format and requirements of i2b2 ontologies and how to build new ones
 - Patient data needs to be returned as SNOMED-CT codes
 - We will look at NCBO's Bioportal web services that can provide mappings from one ontology to another
 - Old ICD9 codes need to be integrated into SNOMED-CT hierarchy
 - When two or more coding systems are mixed together, one will assume the master position of establishing the hierarchy, and the others are mapped into it at various levels.



Introduction to i2b2 Ontology

Shawn Murphy

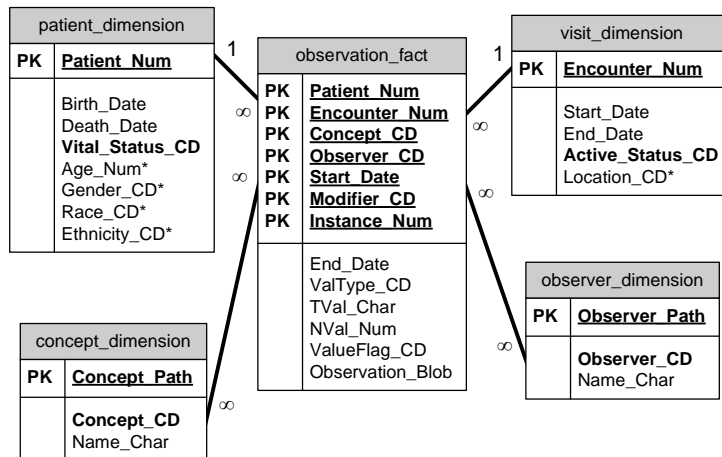
Introduction to i2b2 Ontology

- I2b2 Data model
 - Concept representation
 - Value representation
 - Modifier representation
- I2b2 Metadata
 - Defining concepts
 - Defining values
- Building queries with i2b2 metadata
- Building advanced queries with i2b2 metadata

Star Schema

One fact table surrounded radially by numerous dimension tables.

i2b2 Star Schema



Data Model: Captures data from many different ontologies

- **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.

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

Observation (fact table) Columns

Patient_num Distinct number for every patient

Encounter_num Distinct number for every visit

Concept_cd Distinct code for every concept

Observer_cd Distinct code for every observer

Start_date Date-time observation began

Modifier_cd Code to modify concept_cd

Instance_num Grouper for concept modifiers

Value Columns

Valtype_cd The type of value object
For example, 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 numeric value goes here

Valueflag_cd Flag (for high or low values, for example)

Values in i2b2 data model

concept_...	name_char	valtype_cd	tval_char	rval_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

Modifiers in i2b2 data model

```
P_num|E_num|Instance_num|C_CD|S_d|Modifier_CD|ValType_CD|TVal|NVal
-----
123|107|1|cpt:59622|20060304|@|<null>|<null>|<null>
123|107|1|cpt:59622|20060304|cptmod:62|<null>|<null>|<null>
123|107|1|cpt:59622|20060304|cptmod:AA|<null>|<null>|<null>
123|107|1|cpt:59622|20060304|cptmod:TH|<null>|<null>|<null>
123|567|1|med:aspirin|20100404|@|<null>|<null>|<null>
123|567|1|med:aspirin|20100404|MED:DOSE|N|E|325
123|567|1|med:aspirin|20100404|MED:FREQ|T|QD|<null>
123|567|1|med:aspirin|20100404|MED:ROUTE|T|PO|<null>
123|567|2|med:aspirin|20100404|@|<null>|<null>|<null>
123|567|2|med:aspirin|20100404|MED:DOSE|N|E|83
123|567|2|med:aspirin|20100404|MED:FREQ|T|BID|<null>
123|567|2|med:aspirin|20100404|MED:ROUTE|T|PO|<null>
```

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.

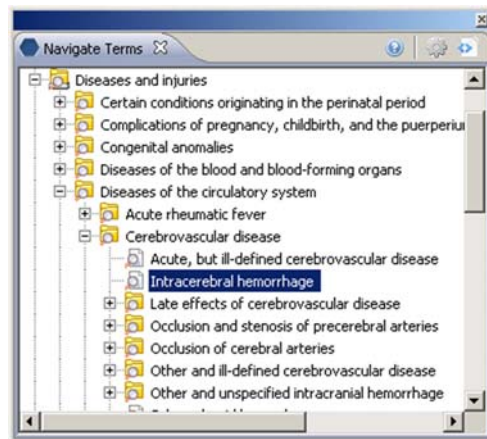
Performing Queries

The screenshot displays the I2b2 Query & Analysis Tool interface. The main window is titled "I2b2 Query & Analysis Tool" and contains several panes. On the left, there is a "Navigate Terms" pane with a tree view of medical categories such as Demographics, Diagnoses, Laboratory Tests, Chemistry, Coagulation, Hematology, Immunology, Pulmonary Function Reports, Urology, Medications, and Procedures. A yellow callout box says "drag an item from here". The central "Query Tool" pane shows a query named "45-54-Circu-Male@18:59:32" with three groups of criteria: Group 1 (45-54 years old, 55-64 years old), Group 2 (Circulatory system), and Group 3 (Male). The groups are connected by "AND" operators. Below the query tool, there are buttons for "Run Query" and "New Query". At the bottom, a "Query Status" window shows "Executing query", "Elapsed time (seconds): 0.9", "Query Finished", and "Matching patients: 5". A "Previous Queries" list is visible at the bottom left, showing a list of query names and dates.

Typical i2b2 Metadata Categories

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

i2b2 Hierarchies

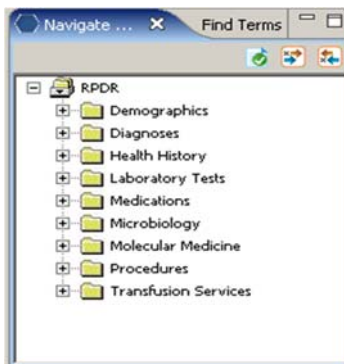


Diseases and injuries \ Diseases of the circulatory system \
Cerebrovascular disease \ Intracerebral hemorrhage

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

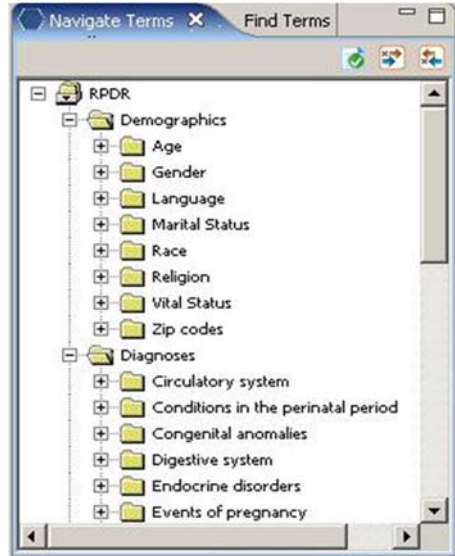
c_hlevel = 1



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

	c_hlevel	c_fullname	c_name	c_visualattributes
1	1	\\RPDR\Microbiology	Microbiology	FA
2	1	\\RPDR\Procedures	Procedures	FA
3	1	\\RPDR\Lablests	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

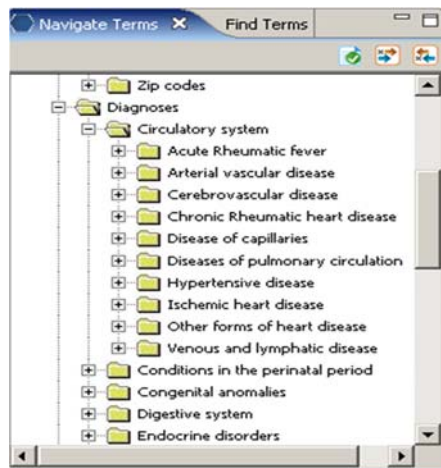
c_hlevel = 2



c_hlevel = 2

c_hlevel = 2

c_hlevel = 3



c_hlevel=3

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_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 of^\(714-0) Rheumatoid arthritis	Rheumatoid arthritis
\RPDR\Diagnoses\Musculoskeletal and connective tissue (710-739)\Arthropathies (710-719)\((714) Rheumatoid arthritis and of^\(714-0) Rheumatoid arthritis	Atrophic arthritis
\RPDR\Diagnoses\Musculoskeletal and connective tissue (710-739)\Arthropathies (710-719)\((714) Rheumatoid arthritis and of^\(714-0) Rheumatoid arthritis	RA [Rheumatoid arthritis]
\RPDR\Diagnoses\Musculoskeletal and connective tissue (710-739)\Arthropathies (710-719)\((714) Rheumatoid arthritis and of^\(714-0) Rheumatoid arthritis	Chronic rheumatic arthritis

c_basecode

The basecode is the coded value for the term.

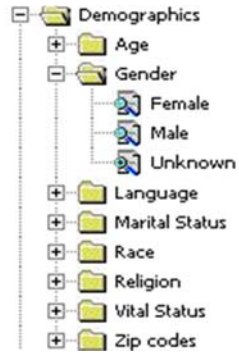
c_level	c_fullname	c_name	c_visualattributes	c_basecode	c_operator
2	\\RPDR\Demographics\Gender	Gender	FA	NULL	LIKE
3	\\RPDR\Demographics\Gender\Female	Female	LA	DEMISEX:f	LIKE
3	\\RPDR\Demographics\Gender\Male	Male	LA	DEMISEX:m	LIKE
3	\\RPDR\Demographics\Gender\Unknown	Unknown	MA	NULL	LIKE
4	\\RPDR\Demographics\Gender\Unknown\Unknown-@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

It maps to the concept_cd in the star schema tables.

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_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)

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

XML to direct value queries

[Also includes ETL directives]

```
<ValueMetadata>
  <DataType/>
  <Flagstouse/>
  <Oktousevalues/>
  <UnitValues>
    <NormalUnits/>
    <EqualUnits/>
    <ConvertingUnits/>
      <Units/>
      <MultiplyingFactor/>
    <ConvertingUnits/>
    <ExcludingUnits/>
  </UnitValues>
  <EnumValues>
    <Val description=""/>
  </EnumValues>
  <MaxStringLength/>
  <LowofLowReference/>
  <HighofLowReference/>
  <LowofHighReference/>
  <HighofHighReference/>
</ValueMetadata>
```

```
<ValueMetadata>
  <Version>3.02</Version>
  <CreationDateTime>
  <TestID/>
  <TestName/>
  <DataType></DataType>
  <Flagstouse/>
  <Oktousevalues/>
  <UnitValues>
    <NormalUnits/>
    <EqualUnits/>
    <ConvertingUnits/>
      <Units/>
      <MultiplyingFactor/>
    <ConvertingUnits/>
    <ExcludingUnits/>
  </UnitValues>
  <EnumValues>
    <Val description=""/>
    <ExcludingVal description=""/>
  </EnumValues>
  <MaxStringLength/>
  <CommentsDeterminingExclusion>
    <Com></Com>
  </CommentsDeterminingExclusion >
  <LowofLowReference/>
  <HighofLowReference/>
  <LowofHighReference/>
  <HighofHighReference/>
  <LowofToxicReference/>
  <HighofToxicReference/>
  <Analysis>
    <Enums/>
    <Counts/>
    <New/>
  </Analysis>
</ValueMetadata>
```

c_metadaxml

DataType will contain the code for what kind of data to expect for this test. Possible values are:

- PosInteger – domain of all positive integers
- Integer – domain of all integers
- PosFloat – domain of all positive real numbers
- Float – domain of all real numbers
- Enum – domain of enumerated values

String – domain of free text, NOT enumerated text values, which would be the Enum data type.

Flagstouse is a string of concatenated flags that are valid for this concept_cd. For example, for most PosIntegers it would be "LNH" Some acceptable values are L (low), N (normal), H (high), A (abnormal), T (toxic)

Oktousevalues will contain a "Y" or a message, which indicates why it is not OK to use values. Nothing indicates that values cannot be used and the user may only specify values using flags.

PosInteger, PosFloat, Integer, Float XML

UnitValues is the parent tag of a set of possibly repeating tags. It contains data when the datatype is PosInteger, Integer, Float, and PosFloat. All units are always LOWER CASE.

NormalUnits can exist only once, it contains a string that a user would recognize which represents the units of the value as we have it in the data warehouse.

EqualUnits can repeat, it contains other strings that are numerically equal to the NormalUnits string.

ConvertingUnits can repeat, it contains other strings, Units, and the factor, MultiplyingFactor, such that values in these units need to be multiplied by the multiplyingfactor to convert them into values of NormalUnits. For example, if NormalUnits was in feet, and ConvertingUnits was in yards, the MultiplyingFactor would be 0.333.

ExcludingUnits can repeat, it contains units that will cause the test with these units to be excluded from the query (and in versions 1.5 and prior, should not be included in the data load). These are units that can not be converted with a simple multiplier to NormalUnits. These concepts will need a new code, or if grouped will need to go into their own group, in order to be queried by value.

PosInteger, PosFloat, Integer, Float XML

LowofLowReference specifies the lowest of the low range values for PosInteger, Integer, Float, and PosFloat datatypes.

HighofLowReference specifies the highest of the low range values for PosInteger, Integer, Float, and PosFloat datatypes.

LowofHighReference specifies the lowest of the high range values for PosInteger, Integer, Float, and PosFloat datatypes.

HighofHighReference specifies the highest of the high range values for PosInteger, Integer, Float, and PosFloat datatypes.

LowofToxicReference specifies the lowest of the toxic range values for PosInteger, Integer, Float, and PosFloat datatypes.

HighofToxicReference specifies the highest of the toxic range values for PosInteger, Integer, Float, and PosFloat datatypes (rarely used).

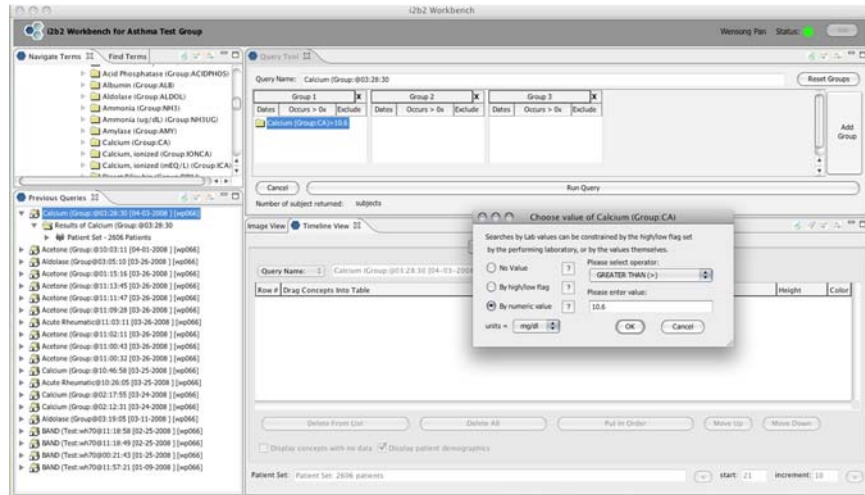
Enum XML

EnumValues is used to specify the list of acceptable enumerated values, each enclosed in the Val tag. Enumerated values that indicate an invalid test result (for the enum datatype) can be enclosed in the ExcludingVal tag. The “descriptions” parameter exists to allow a humanly readable value to be presented for the enumerated value in choice boxes of user interfaces. ExcludingVal is directed to loading processes, and specifies values not to lead into the database or display in a user interface (for example, “pending”).

String XML

- MaxStringLength will contain a positive integer, representing the longest acceptable string length, if the datatype is “string”.

Query by values



**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

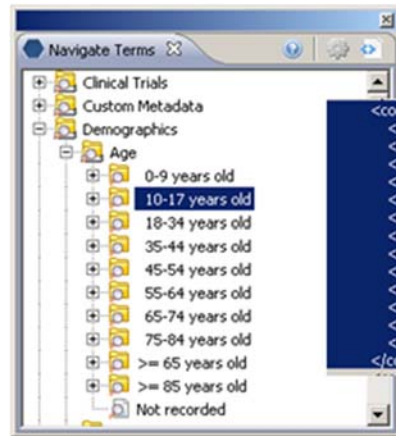
Select * from observation_fact where c_facttablecolumn in
(select concept_cd from c_tablename where c_columnname c_operator 'c_dimcode%')

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

Multiple paths

```
select distinct(patient_num) 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 -\38241) 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-\%')
```

Ontology terms can define more types of queries



Patient_Dimension

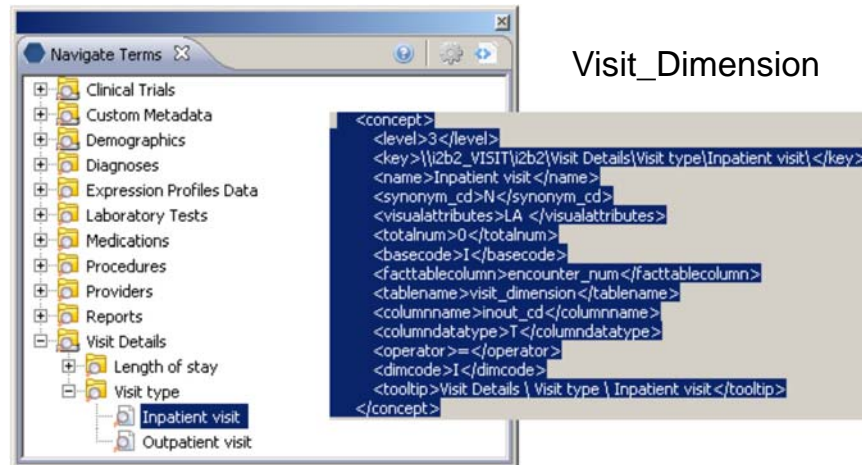
```
<concept>
  <level>3</level>
  <key>||2b2_DEMO||2b2|Demographics|Age|10-17 years old|</key>
  <name> 10-17 years old</name>
  <synonym_cd>N</synonym_cd>
  <visualattributes>FA </visualattributes>
  <totalnum>0</totalnum>
  <facttablecolumn>patient_num</facttablecolumn>
  <tablename>patient_dimension</tablename>
  <columnname>birth_date</columnname>
  <columndatatype>N</columndatatype>
  <operator>BETWEEN</operator>
  <dimcode>sysdate - (365.25*18) AND sysdate - (365.25*10) </dimcode>
  <tooltip>Demographics | Age | 10-17 years old</tooltip>
</concept>
```

Ontology terms can define more types of queries

```
select patient_num from [c_facttablename] where
[c_facttablecolumnname] IN
(select [c_facttablecolumnname] from
[c_tablename] where [c_columnname]
[c_operator] [c_dimcode])
```

```
select patient_num from patient_dimension
where birth_date BETWEEN
sysdate – (365.25*18) AND sysdate –
(365.25*10)
```

Ontology terms can define more types of queries



Visit_Dimension

```

<concept>
  <level>3</level>
  <key>\\2b2_VISIT\\2b2\\Visit Details\\Visit type\\Inpatient visit</key>
  <name>Inpatient visit</name>
  <synonym_cd>N</synonym_cd>
  <visualattributes>LA</visualattributes>
  <totalnum>0</totalnum>
  <basecode>I</basecode>
  <facttablecolumn>encounter_num</facttablecolumn>
  <tablename>visit_dimension</tablename>
  <columnname>inout_cd</columnname>
  <columndatatype>T</columndatatype>
  <operator>=</operator>
  <dimcode>I</dimcode>
  <tooltip>Visit Details \\ Visit type \\ Inpatient visit</tooltip>
</concept>

```

Encounters for all patients that had a visit when they were 40-45 years old:
(SQL Server)

```

select encounter_num from visit_dimension
where start_date BETWEEN
((select birth_date from patient_dimension
  where patient_num = visit_dimension.patient_num)
+ (365.25 * 40))
AND
((select birth_date from patient_dimension
  where patient_num = visit_dimension.patient_num)
+ (365.25 * 45))

```

(Ontology table entries)

```

[c_name] = patients that had a visit when they were 40-45 years old
[c_fullpath] = \\encounter detail\\ age at visit\\ 40 – 45 y/o
[c_facttablecolumnname] = encounter_num
[c_tablename] = visit_dimension
[c_columnname] = start_date
[c_operator] = BETWEEN
[c_dimcode] =
((select birth_date from patient_dimension
  where patient_num = visit_dimension.patient_num)
+ (365.25 * 40))
AND
((select birth_date from patient_dimension
  where patient_num = visit_dimension.patient_num)
+ (365.25 * 45))

```

Ontology services from NCBO

Nigam Shah
nigam@stanford.edu



NCBO: Key activities



- We **create and maintain a library** of biomedical ontologies.
- We **build tools and Web services** to enable the use of ontologies and their derivatives.
- We **collaborate with scientific communities** that develop and use ontologies.

THE NATIONAL CENTER FOR BIOMEDICAL ONTOLOGY

www.bioontology.org

National Center for Biomedical Ontology

Community Technology

Learning About Ontologies Dissemination & Training

Ontology Library Data Annotation

Go to BioPortal Go to Annotator

NCBO Collaborations Forum, Blog Publications

Ontology Development Data Access Using Ontologies

Go to Resource Index

Of Current Interest

- Webinar: Webinar series takes summer break and will resume on Wednesday, Sept. 1, 10 AM (PST). Cecilia Arighi, Protein Information Resource
- Recent Publication: Jonquet, C, Musen, MA, Shah, NH (2010). Building a biomedical ontology recommender web service
- Recent Release: BioPortal 2.5 (June 7, 2010)
- Call for Panels: Panels due October 18, 2010, AMA 2011 Summit on Translational Bioinformatics
- NCBO - Public Talks - Subscribe
- More News & Events
- More Webinars
- Follow us on Twitter

NCBO User Profile

Gary An

The University of Chicago

"The development of agent-based modeling is intimately tied to the advances in the ontologies and tools provided by NCBO via BioPortal." [More >](#)

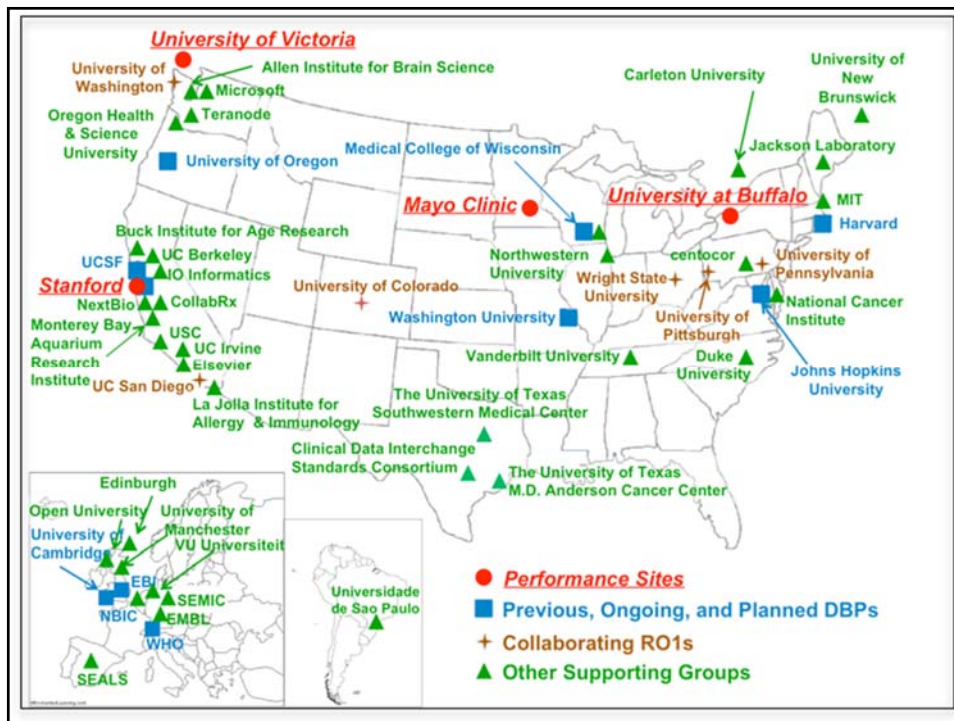
[Other profiles >](#)

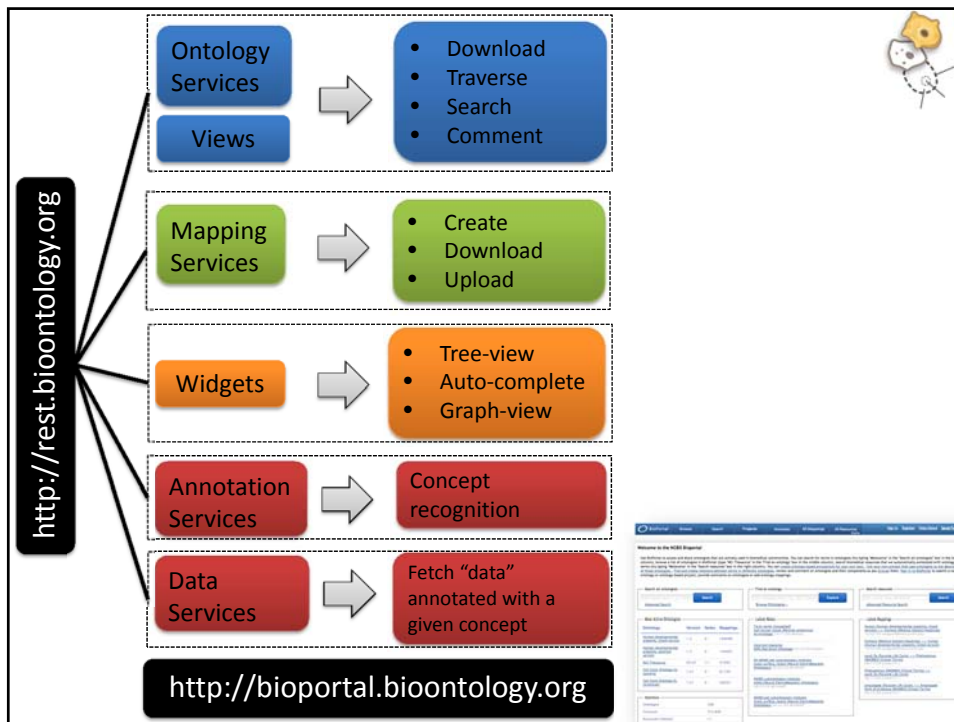
Video

Learn about Biomedical Ontologies. Watch an introductory video.

Browse ontologies in BioPortal!

BioPortal allows users to browse, search and visualize ontologies.





ONTOLOGY SERVICES

Accessing, browsing, searching and traversing ontologies in *Your* application

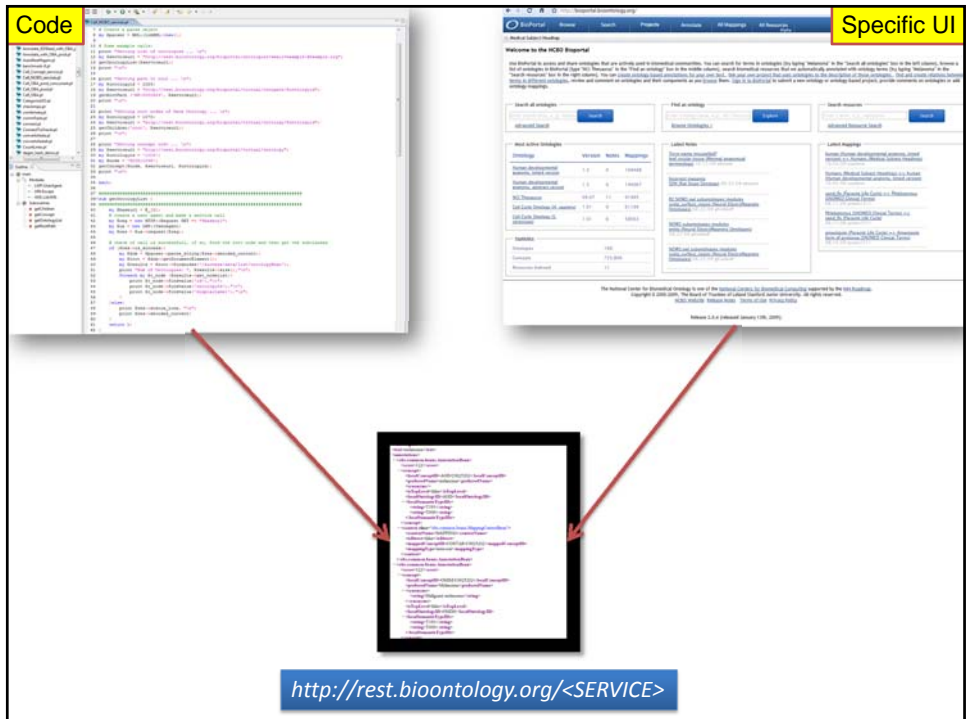
www.bioontology.org/wiki/index.php/NCBO_REST_services



Contents [hide]

- 1 Announce List
- 2 Services to access ontologies and ontology versions
 - 2.1 List all the latest version of ontologies
 - 2.2 Get a specific ontology based on a version id
 - 2.3 Download an ontology file
 - 2.4 Download the latest ontology file
 - 2.5 Get all versions of an ontology from a virtual ontology id
 - 2.6 Get latest version of an ontology id
 - 2.7 List all ontology categories
 - 2.8 List all ontology groups
- 3 Services to access ontology views and ontology view versions
 - 3.1 List all the latest version of views
 - 3.2 Get all versions of views from a virtual ontology id
- 4 Concept services
 - 4.1 Get concept
 - 4.2 Get all root concepts for an ontology id
 - 4.3 Get concept for latest ontology version id
 - 4.4 Get all concepts
- 5 Search services
 - 5.1 Search BioPortal
- 6 Hierarchy Services
 - 6.1 Get parents/children of a given concept in a specific ontology version
 - 6.2 Get parents/children of a given concept in the latest version of a given ontology
 - 6.3 Get paths to roots/leaves from a concept in a specific ontology version
 - 6.4 Get paths to root/leaves from a concept in the latest version of a given ontology
 - 6.5 Get siblings of a given concept in a specific ontology version
 - 6.6 Get siblings of a given concept in the latest version of a given ontology
- 7 Notes Service
- 8 Mapping Service
- 9 Annotator Service
- 10 Usage Logging
- 11 Overview - Using NCBO Technology in Your Project

49



```

<success>
  <data>
    <list>
      <ontologyBean>
        <id>40401</id>
        <ontologyId>1351</ontologyId>
        <displayLabel>Medical Subject Headings</displayLabel>
        <description>Medical Subject Headings (MeSH);National I
        <abbreviation>MSH</abbreviation>
        <format>RRF</format>
        <internalVersionNumber>1</internalVersionNumber>
        <versionNumber>2009_2009_02_13</versionNumber>
        <contactName>Stuart Nelson, M.D.</contactName>
        <contactEmail>nelson@nlm.nih.gov</contactEmail>
        <statusId>3</statusId>
        <categoryIds>
          <int>5058</int>
        </categoryIds>
        <groupIds>
          <int>6008</int>
        </groupIds>
        <isFoundry>0</isFoundry>
        <dateCreated>2009-07-31 06:01:06.0 PDT</dateCreated>
        <versionNumber>1</versionNumber>
        <contactName>anatomy JAX</contactName>
        <contactEmail>anatomy@informatics.jax.org</contactEmail>
        <statusId>5</statusId>
        <categoryIds>
          <int>2812</int>
          <int>2811</int>
          <int>2810</int>
          <int>2817</int>
        </categoryIds>
      </ontologyBean>
    </list>
  </data>
</success>

```

http://rest.bioontology.org/bioportal/ontologies

```

<success>
  <accessedResources>/bioportal/search/melanoma/</accessedResource>
  <accessDate>2009-10-14 15:17:57.410 EDT</accessDate>
  <data>
    <page>
      <page>
        <numPages>1</numPages>
        <pageSize>15</pageSize>
        <numResultsPage>15</numResultsPage>
      </page>
      <searchBean>
        <ontologyVersionId>40401</ontologyVersionId>
        <ontologyId>1351</ontologyId>
        <ontologyDisplayLabel>Medical Subject Headings</ontologyDisplayLabel>
        <recordType>RECORD_TYPE_PREFERRED_NAME</recordType>
        <conceptId>D008545</conceptId>
        <conceptIdShort>D008545</conceptIdShort>
        <preferredName>Melanoma</preferredName>
        <contents>Melanoma</contents>
      </searchBean>
      <ontologyDisplayLabel>Medical Subject Headings</ontologyDisplayLabel>
      <recordType>RECORD_TYPE_PREFERRED_NAME</recordType>
      <conceptId>D008545</conceptId>
      <conceptIdShort>D008545</conceptIdShort>
      <preferredName>Melanoma, Experimental</preferredName>
      <contents>Melanoma, Experimental</contents>
    </searchBean>
    <searchBean>
      <ontologyVersionId>40401</ontologyVersionId>
      <ontologyId>1351</ontologyId>
      <ontologyDisplayLabel>Medical Subject Headings</ontologyDisplayLabel>
      <recordType>RECORD_TYPE_SYNONYM</recordType>
      <conceptId>C092706</conceptId>
      <conceptIdShort>C092706</conceptIdShort>
      <preferredName>BAGE protein, human</preferredName>
      <contents>B melanoma antigen, human</contents>
    </searchBean>
  </data>
</success>

```

http://rest.bioontology.org/bioportal/search/melanoma/?ontologyids=1351



VIEWS

Custom subsets of large ontologies



Views and Value Sets

- Users can contribute their derivatives of BioPortal ontologies,
 - which become first-class objects in BioPortal and can be used as all other ontologies are (e.g., as value sets)
- Recently added: **a view-extractor service**
 - Enables users to extract a subtree of an ontology in OWL

Views in BioPortal



BioPortal | Browse | Search | Projects | Annotate | All Mappings | All Resources Alpha | Logged In As: ratasha | My Accounts | Log Out

SNOMED Clinical Terms [Subscribe to updates](#)

[Submit New Version](#) | [Edit Ontology Information](#)

Metadata | **Views (8)** | Projects (1) | Reviews (8) | Metrics | Ontology Widgets

Create New View

Expand All | Collapse All

- CORE Subset of SNOMED CT** [@Edit Metadata](#)
 - Description: The CORE Problem List Subset of SNOMED CT
 - Ontology ID: 2040
 - Definition Language: Manual

VERSION	BASE VERSION	CREATED	CREATED BY	ONTOLOGY FILE	VISUALIZE
CORE_2009_01_31	2009_07_31	01/31/2009	Vivian A. Auld, auld@nlm.nih.gov	Download View	Explore

[Submit New View Version](#)

- SNOMED Anatomy** [@Edit Metadata](#)
- SNOMED Clinical Findings** [@Edit Metadata](#)
- SNOMED Organism** [@Edit Metadata](#)
- SNOMED Terminos Clinicos** [@Edit Metadata](#)
 - Description: SNOMED Terminos Clinicos (SNOMED CT), Edicion en Espanol, Distribucion Internacional, Octubre de 2008, 2009_04_30
 - Ontology ID: 1416

VERSION	BASE VERSION	CREATED	CREATED BY	ONTOLOGY FILE	VISUALIZE
2009_04_30	2009_07_31	04/30/2009	custserv@nlm.nih.gov	Download View	Explore
2009_04_30	2009_07_31	04/30/2009	custserv@nlm.nih.gov	Download View	Explore

[Submit New View Version](#)

- SNOMED Test Findings** [@Edit Metadata](#)

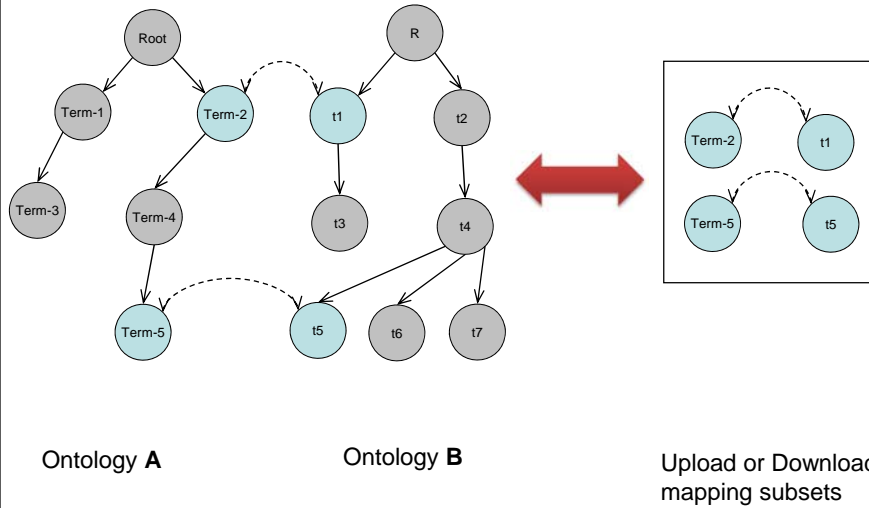
57

MAPPINGS



Using NCBO technology to integrate terminologies and ontologies

Mappings

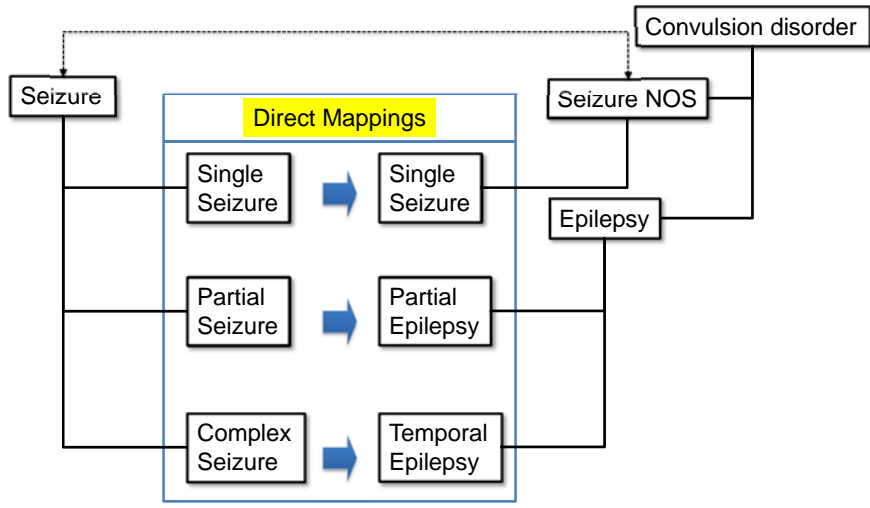


Using Mappings for query federation



FROM (site #1)

TO (site #2)





WIDGETS

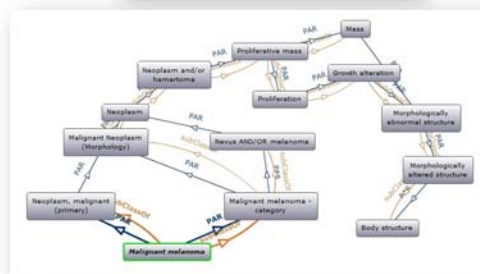
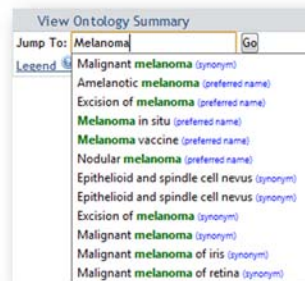
Using NCBO technology on your web pages

Ontology Widgets



UI components with “BioPortal inside”:

- **term-selection** widget for a specific ontology
- **form fields** with auto-complete from a specific BioPortal ontology
- **RSS feed** for an ontology
- **Visualization** widget
- **Tree** widget





ANNOTATOR SERVICE

Using Ontologies to Annotate *Your* Data

Annotation as a Web service



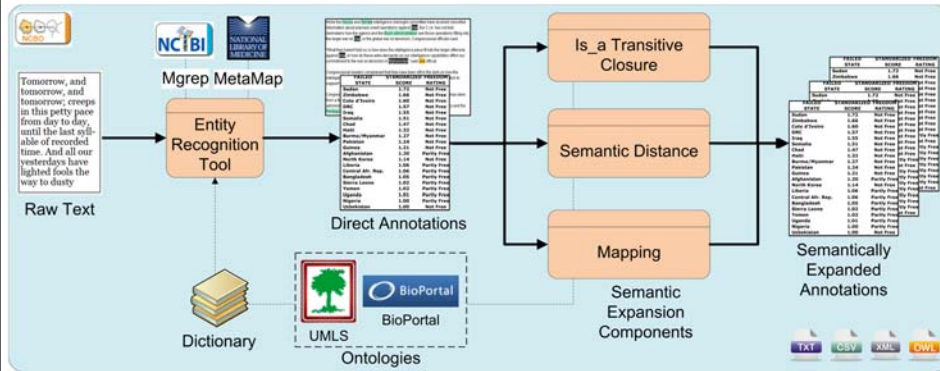
Process textual metadata to automatically tag text with as many ontology terms as possible.

GDS Summary		
Accession:	GDS986 View Expression (GEO profiles)	
Title:	Bladder smooth muscle cell response to mechanical stretch	
Data Set type:	Gene expression array-based (GSA) (in silico oligonucleotide)	
Summary:	Expression profiling of cultured bladder smooth muscle cells subjected to repetitive mechanical stimulation for 4 hours. Chronic overdistension results in bladder wall thickening, associated with loss of muscle contractility. Results identify genes whose expression is altered by mechanical stimuli.	
Platform:	GPL14476 (Affymetrix Human Genome U133A Array) (see chip)	
Citations:	Adam RM, Eaton SH, Estrada C, Rangasankar A et al. Mechanical stretch is a highly selective regulator of gene expression in human bladder smooth muscle cells. <i>Physiol Genomics</i> 2004 Dec; 15:20(1):26-44. PMID: 15467014	
Sample organism:	homo sapiens	Platform organism:
Feature count:	22263	Value type:
Series:	GSE1398	Series published:
Last GDS update:	12/09/2004	



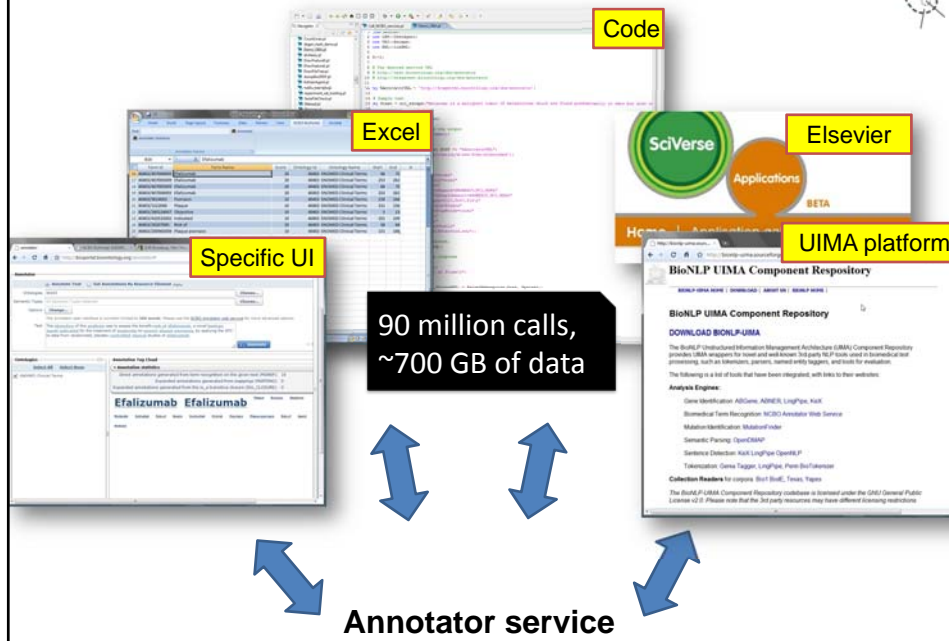
Expression, Expression of bladder, bladder, smooth, bladder muscle, muscle, smooth muscle, cells, mechanical, mechanical stimulation, stimulation, Chronic, results, bladder overdistension, associated, associated with, with, loss, genes, altered

Annotator: workflow



- "Melanoma is a malignant tumor of melanocytes which are found predominantly in skin but also in the bowel and the eye".
 - 39228/DOID:1909, *Melanoma* in Human Disease
- Transitive closure
 - 39228/DOID:191, *Melanocytic neoplasm*, direct parent of *Melanoma* in Human Disease
 - 39228/DOID:0000818, *cell proliferation disease*, grand parent of *Melanoma* in Human Disease

Multiple ways to access

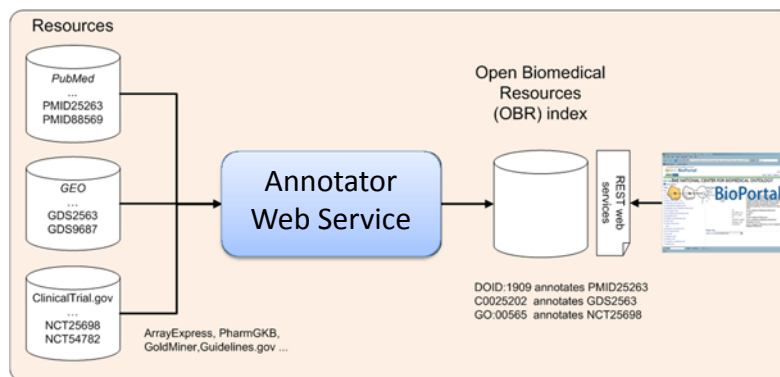




DATA SERVICE

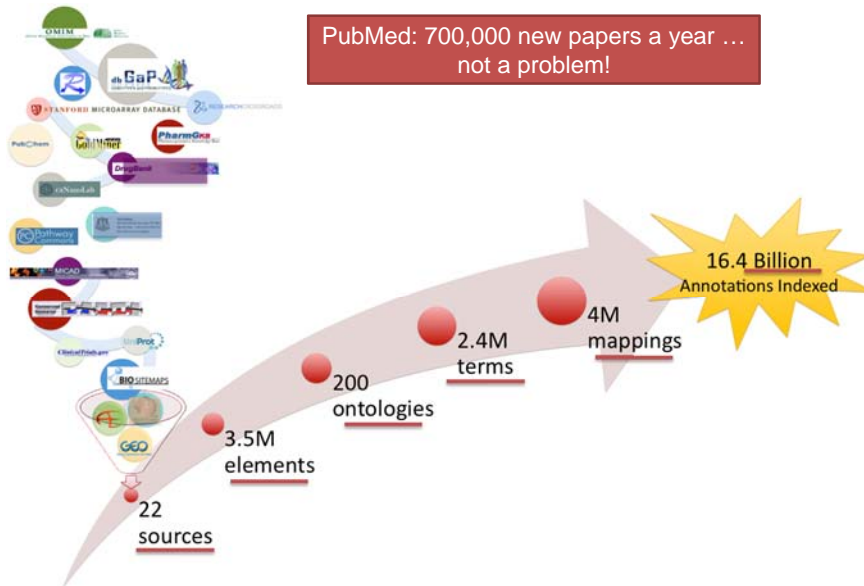
Using Ontologies to Access Public Data

Resource index: The Basic Idea



- The index can be used for:
 - Search
 - Data mining

Indexed resources (target = 100 in 5 years)



<http://bioportal.bioontology.org/resources>



NCBO is building a system for automated ontology-based annotation and indexing of biomedical data. We process the textual metadata of diverse elements of biomedical resources such as gene expression data sets, descriptions of radiology images, clinical-trial reports, and PubMed abstracts to annotate and index them with terms from appropriate ontologies. Use the interface below to search the resulting index of annotations and to identify biomedical data resources annotated with particular ontology terms. Try searching for "Melanoma" and use NCI Thesaurus as your ontology.

Won 1st prize at the 2010 Semantic Web Challenge @ ISWC

BioPortal Resource Search [View Demo!](#)

Type in a search term

- | | |
|--|--|
| 1155 ARRS GoldMiner | 1172881 Adverse Event Reporting System Data |
| 15190 ArrayExpress | 1653 Biositemaps |
| 98304 ClinicalTrials.gov | 40809 Conserved Domain Database (CDD) |
| 251 Database of Genotypes and Phenotypes | 4774 DrugBank |
| 21969 Gene Expression Omnibus DataSets | 823 MICAD |
| 21242 Online Mendelian Inheritance in Man | 923 Pathway Commons |
| 832 PharmGKB [Disease] | 1634 PharmGKB [Drug] |
| 992 PharmGKB [Gene] | 110481 PubChem |
| 87858 Reactome | 1033651 ResearchCrossroads |
| 18710 Stanford Microarray Database | 18344 UniProt KB |
| 1494 WikiPathways | 845 caNanoLab |

Multiple ways to access



- Resources annotated = 22
- Total records = 3.5 million
- Direct annotations = 517 million
- After transitive closure = 16.4 Billion

Code

Specific UI

Resource Tab



Resource Index

Use-cases for users of i2b2

Aim 1: Integrate NCBO services in i2b2



- Preliminary results:
 - Export any ontology stored in BioPortal into the format used by i2b2's ontology cell
- Future Work:
 - Make the export code available as a service
 - Embed the extraction code into the i2b2 Ontology Cell to "pull" content
 - Ensure we have the latest versions of ontologies used by i2b2 and CTSA users (ICD9, ICD10, SNOMEDCT, RXNORM, LOINC, CPT)

Aim 2: Mappings for query federation



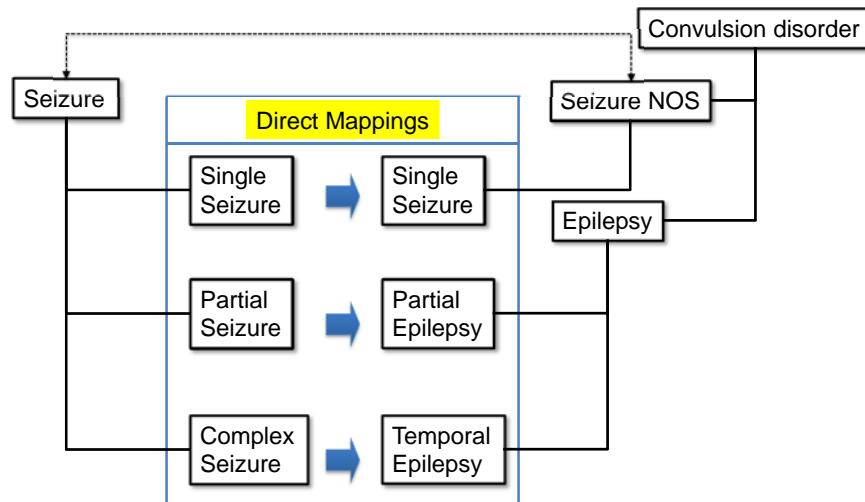
- Preliminary result:
 - Worked out the workflow for using mappings for query translation
 - Detailed discussions with the HOM and OpenMDR groups to define use-case and elicit requirements
- Future work:
 - Use BioPortal as the shared repository for inter terminology mappings
 - Tackle access, IP, performance, and institutional issues
 - Key features
 - ✓ Import outside mappings
 - ✓ Update mappings when versions change
 - ✓ Mechanism to curate mappings
 - Support proprietary curation and content

Using Mappings for query federation



FROM (site #1)

TO (site #2)



Informatics for Integrating Biology & the Bedside

A National Center for Biomedical Computing



Integration of NCBO Bioportal for building and mapping i2b2 Ontologies

Lori Phillips

Building an i2b2 Ontology with NCBO services

- To support our use case we are going to discuss extraction of SNOMED from NCBO into i2b2.

BioPortal SNOMED_CT

The screenshot displays the BioPortal interface for the SNOMED Clinical Terms ontology. The top navigation bar includes links for BioPortal, Browse, Search, Projects, Annotate, All Mappings, and All Resources. The current page is titled 'SNOMED Clinical Terms Version 2010_07_31' and features a search bar with the term 'Body structure' and options to 'Link Here' or 'Subscribe'.

The main content area is divided into two sections. On the left, under 'View Ontology Summary', there is a 'Jump To:' field with a 'Go' button and a 'Legend' link. Below this is a list of ontology classes with expandable icons: Body structure, Clinical finding, Environment or geographical location, Event, Linkage concept, Observable entity, Organism, Pharmaceutical / biologic product, Physical force, Physical object, Procedure, Qualifier value, Record artifact, Situation with explicit context, Social context, Special concept, Specimen, Staging and scales, and Substance.

On the right, the 'Details' tab is active, showing the following information for the selected term:

ID:	123037004
Full Id:	http://purl.bioontology.org/ontology/SNOMEDCT/123037004
Synonyms:	Body structures Body structure (body structure)
Conceptstatus:	0
Ctv3:	XU5tN
Isprimitive:	1
Semantic Type:	Anatomical Structure
Snomedid:	T-D000A
Synonym Fn:	Body structure (body structure)
Synonym Sy:	Body structures
Tui:	T017
Umls Cui:	C1268086

Bioportal Clinical Finding view

SNOMED Clinical Terms [Subscribe to updates](#)

Metadata **Views (8)** Projects (1) Reviews (1) Notes (0) Metrics Ontology Widgets

Create New View

Expand All | Collapse All

- CORE Subset of SNOMED CT
- SNOMED Anatomy
- SNOMED Clinical Findings**
 - Description: The Clinical Finding subtree of SNOMED CT
 - Definition: Class subtree of ClinicalFinding
 - Ontology ID: 2018
 - Definition Language: Manual
- SNOMED Ethnic Group

VERSION	BASE VERSION	CREATED	CREATED BY	ONTOLOGY FILE	DIFF FILE	VISUALIZE
1.2	2009_07_31	07/16/2010	Tania Tudorache, tudorache@stanford.edu	Download View		Explore
1.1	2009_07_31	03/23/2010	Tania Tudorache, tudorache@stanford.edu	Download View		Explore
1.0	2009_01_31	09/09/2009	Tania Tudorache, tudorache@stanford.edu	Download View		Explore

Strategy

- Pull data from NCBO via REST services.
- Reorganize information into format used by i2b2 Ontology cell

bioportal/virtual/ontology/2018/all

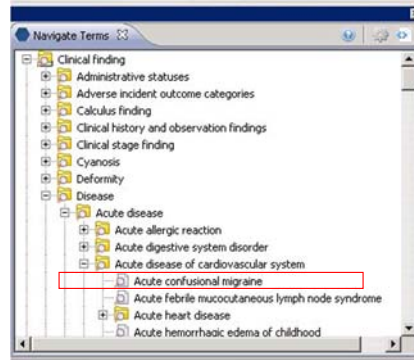
```
<data>
<pageNum>1</pageNum>
<numPages>1934</numPages>
<pageSize>50</pageSize>
<numResultsPage>50</numResultsPage>
<numResultsTotal>96677</numResultsTotal>
<contents
class="org.ncbo.stanford.bean.concept.
ClassBeanResultListBean">
<classBeanResultList>
<classBean>
<id>http://hltsdo.org/snomedct/clinicalFinding#10000006</id>
<fullId>http://hltsdo.org/snomedct/clinicalFinding#10000006</fullId>
<label>Radiating chest pain</label>
<type>class</type>
<relations>.....
<entry>
<string>CONCEPTSTATUS</string>
<list>
<string>0</string>
</list>
</entry>
<entry>
<string>SubClass</string>
<list>
<entry> .....
```



METADATA COLUMN	DATA TYPE (ORACLE)	DATA TYPE (SQL)
C_HLEVEL	INT	INT
C_FULLNAME	VARCHAR2(700)	VARCHAR(700)
C_NAME	VARCHAR2(2000)	VARCHAR(2000)
C_SYNONYM_CD	CHAR(1)	CHAR(1)
C_VISUALATTRIBUTES	CHAR(3)	CHAR(3)
C_BASECODE	VARCHAR2(50)	VARCHAR(50)
C_METADATAXML	CLOB	TEXT
C_FACTTABLECOLUMN	VARCHAR2(50)	VARCHAR(50)
C_TABLENAME	VARCHAR2(50)	VARCHAR(50)
C_COLUMNNAME	VARCHAR2(50)	VARCHAR(50)
C_COLUMNDATATYPE	VARCHAR2(50)	VARCHAR(50)
C_OPERATOR	VARCHAR2(10)	VARCHAR(10)
C_DIMCODE	VARCHAR2(700)	VARCHAR(700)
C_TOOLTIP	VARCHAR2(900)	VARCHAR(900)
UPDATE_DATE	DATE	DATETIME
DOWNLOAD_DATE	DATE	DATETIME
IMPORT_DATE	DATE	DATETIME
SOURCESYSTEM_CD	VARCHAR2(50)	VARCHAR(50)
VALUETYPE_CD	VARCHAR2(50)	VARCHAR(50)

Primary challenges

- i2b2 Ontology cell depends upon hierarchical information
 - c_fullname, c_tooltip maintain the hierarchy from root to leaves



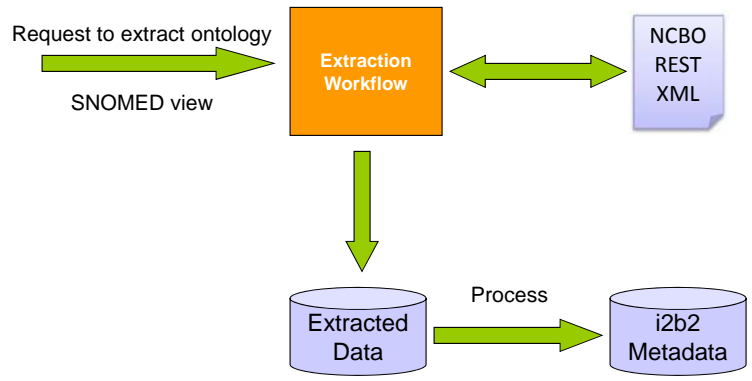
Clinical finding \ Diseases \ Acute disease \
Acute disease of cardiovascular system \ Acute confusional migraine

Challenges..

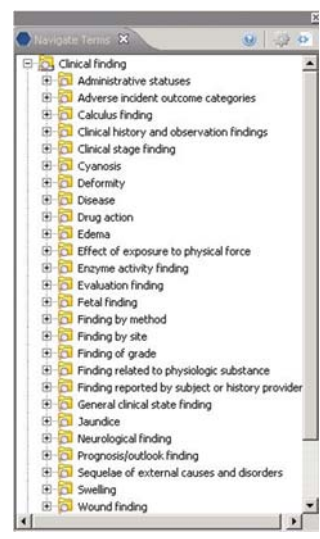
- REST service that enables pull of concepts includes immediate parent/child info only
 - i2b2 c_fullname , c_tooltip must be computed or acquired through an additional REST service (<http://rest.bioontology.org/bioportal/path>)

```
<data>
<classBean>
  <id>95653008</id>
  <label>Acute confusional migraine</label>
  <relations>
    <entry>
      <string>SuperClass</string>
      <list>
        <classBean>
          <id>128487001</id>
          <label>Acute disease of cardiovascular system</label>
        </classBean>
      </list>
    </entry>
  </relations>
</classBean>
</data>
```

NCBO Extraction workflow



Extracted SNOMED terms

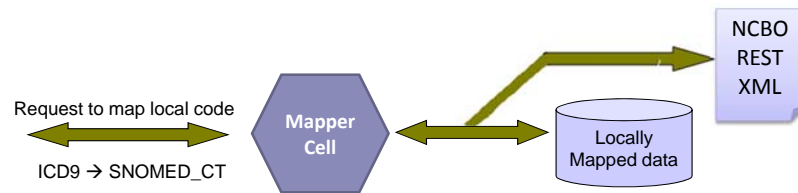


Providing mappings using NCBO services

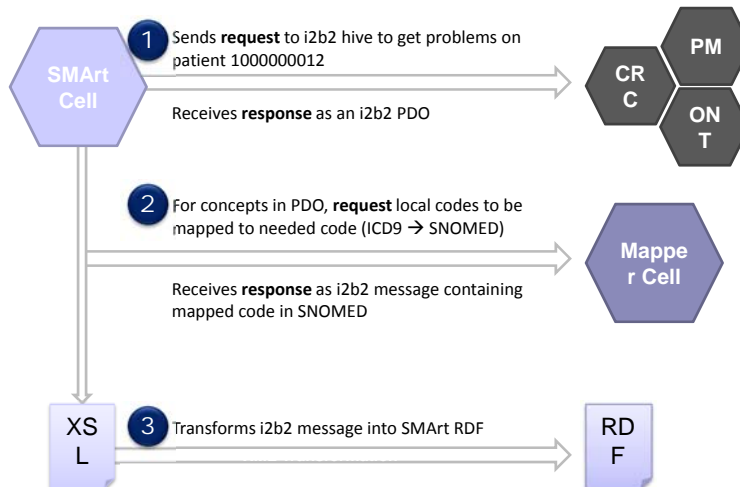
■ Mapping cell

- Need for mapping different site-specific ontologies in cross-institutional settings. (SNOMED_CT ↔ ICD-9, RxNORM ↔ NDC)
- First look for locally mapped data
- Then seek mappings through NCBO services.

<http://bioportal.bioontology.org/mappings/service/1101/>



Mapping Use Case



Example

1. An individual's problems as represented in an i2b2 XML message:

```
- <ns2:concept_set>
- <concept>
  <concept_path>\\i2b2\Diagnoses\Circulatory system (390-459)\Hypertensive disease (401-405)\(401)
  Essential hypertension\ (401-9) Unspecified essential hyp~\ </concept_path>
  <concept_cd>ICD9:401.9</concept_cd>
  <name_char>Huchard's disease</name_char>
</concept>
- <concept>
  <concept_path>\\i2b2\Diagnoses\Digestive system (520-579)\Oral cavity diseases (520-529)\(523)
  Gingival and periodontal di~\ (523-6) Accretions on teeth\ </concept_path>
  <concept_cd>ICD9:523.6</concept_cd>
  <name_char>Dental plaque</name_char>
</concept>
- <concept>
  <concept_path>\\i2b2\Diagnoses\Digestive system (520-579)\Oral cavity diseases (520-529)\(523)
  Gingival and periodontal di~\ (523-9) Unspecified gingival and ~\ </concept_path>
  <concept_cd>ICD9:523.9</concept_cd>
  <name_char>Unspecified gingival and periodontal disease</name_char>
</concept>
- <concept>
  <concept_path>\\i2b2\Diagnoses\Genitourinary system (580-629)\Nephritis, nephrotic syndrome, and
  nephrosis (580-589)\(585) Chronic renal failure\ </concept_path>
  <concept_cd>ICD9:585</concept_cd>
  <name_char>Chronic renal failure</name_char>
</concept>
```

Ontology
Cell

Case: We require diagnoses to be coded in SNOMED, but we only have ICD9 codes.

Action: Send an XML request to the Mapping Cell with the desired ICD9's (above) to be mapped to SNOMED (e.g. ICD9: 401.9)

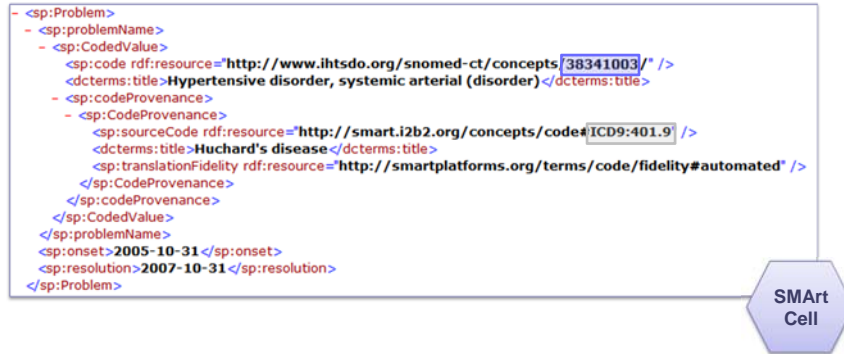
2. Response from Mapping Cell contains SNOMED codes and names:

```
- <mapped_concept_set>
- <mapping>
  <source_coding_system>ICD9</source_coding_system>
  <source_basecode>401.9</source_basecode>
  <destination_coding_system>SNO</destination_coding_system>
  <destination_basecode>38341003</destination_basecode>
  <destination_name>Hypertensive disorder, systemic arterial (disorder)</destination_name>
</mapping>
- <mapping>
  <source_coding_system>ICD9</source_coding_system>
  <source_basecode>523.6</source_basecode>
  <destination_coding_system>SNO</destination_coding_system>
  <destination_basecode>17552000</destination_basecode>
  <destination_name>Dental calculus (disorder)</destination_name>
</mapping>
- <mapping>
  <source_coding_system>ICD9</source_coding_system>
  <source_basecode>523.9</source_basecode>
  <destination_coding_system>SNO</destination_coding_system>
  <destination_basecode>2556008</destination_basecode>
  <destination_name>Periodontal disease (disorder)</destination_name>
</mapping>
- <mapping>
  <source_coding_system>ICD9</source_coding_system>
  <source_basecode>585</source_basecode>
  <destination_coding_system>SNO</destination_coding_system>
  <destination_basecode>90688005</destination_basecode>
  <destination_name>Chronic renal failure syndrome (disorder)</destination_name>
</mapping>
```

Mapping
Cell

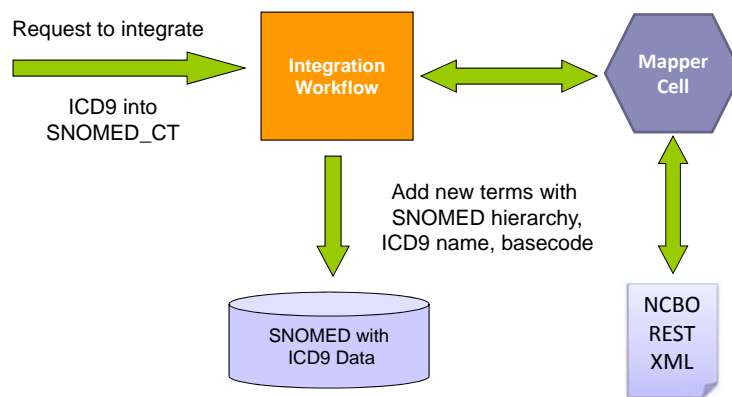
The mapping cell returned the mappings for the ICD9's sent from the previous slide (e.g. SNOMED of 38341003 mapped to ICD9: 401.9)

3. i2b2 “Mapped PDO” is transformed into RDF in SMarT Cell

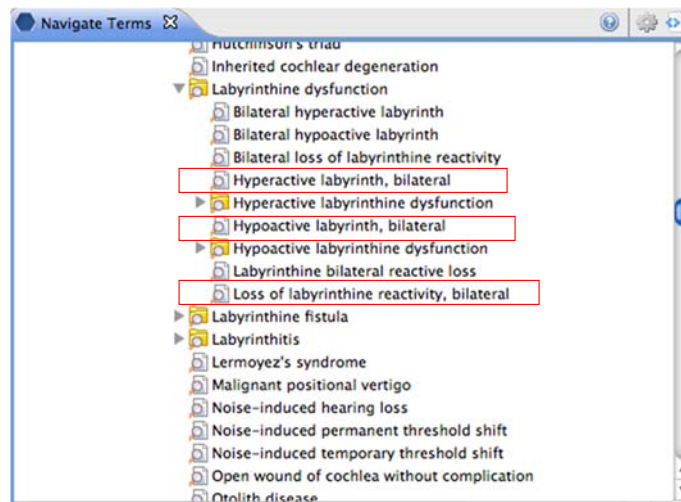


Integrating two i2b2 ontologies with NCBO services

- Integrate ICD-9 into SNOMED
 - For each SNOMED term, find a mapped ICD9 term:



SNOMED folder with ICD9 terms



THANK YOU