

i2b2 – NCBO Collaboration to Provide i2b2 Ontology Services

Shawn Murphy MD, Ph.D.
Lori Phillips MS

What are the pressing problems in i2b2 ontologies? – what we have heard

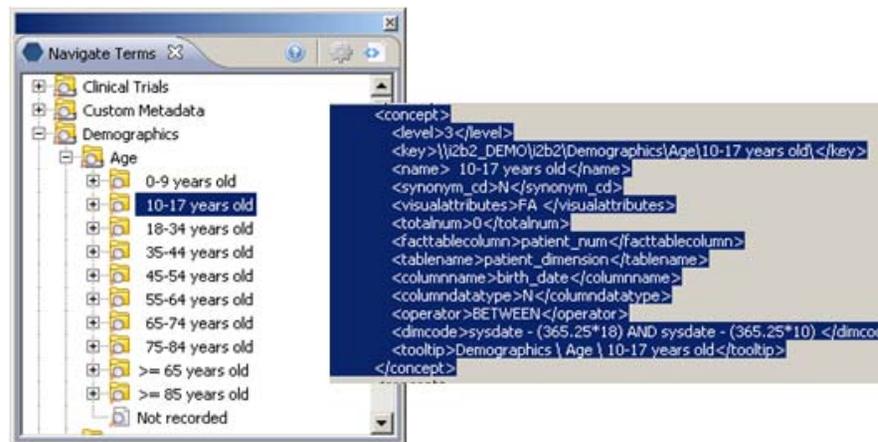
- Creation of i2b2 ontologies that represent the standard ontologies
- Support the creation of custom (local) ontologies
- Versioning of ontologies
 - Necessary for backward compatibility of queries
- Standards across institutions
 - Mapping two ontologies
 - [patient/visit/fact table choices]
 - [concept vs. modifier]
 - Creating new global ontologies

how are we solving some of these problems?

- Creation of i2b2 ontologies that represent the standard ontologies – NCBO-linked tools
- Support the creation of custom (local) ontologies – Ontology editor in i2b2 Workbench
- Versioning of ontologies – Version awareness when using NCBO web services, Guidance
- Standards across institutions
 - Mapping two ontologies – Mapping Cell that will be discussed in SMART presentation, SHRINE tools
 - [patient/visit/fact table choices] – Guidance, community discussion
 - [concept vs. modifier] – Guidance, community discussion
 - Creating new global ontologies – Out of Scope

Patient and Visit dimension table

- Patient and Visit table can contain data, what data should be in these tables rather than the observation_fact table?
 - Guidance
 - Data that will be used in a computation (birth date, visit date) – because rules can be created in the ontology tables
 - Patient dimension table can hold data with no valid date of observation



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

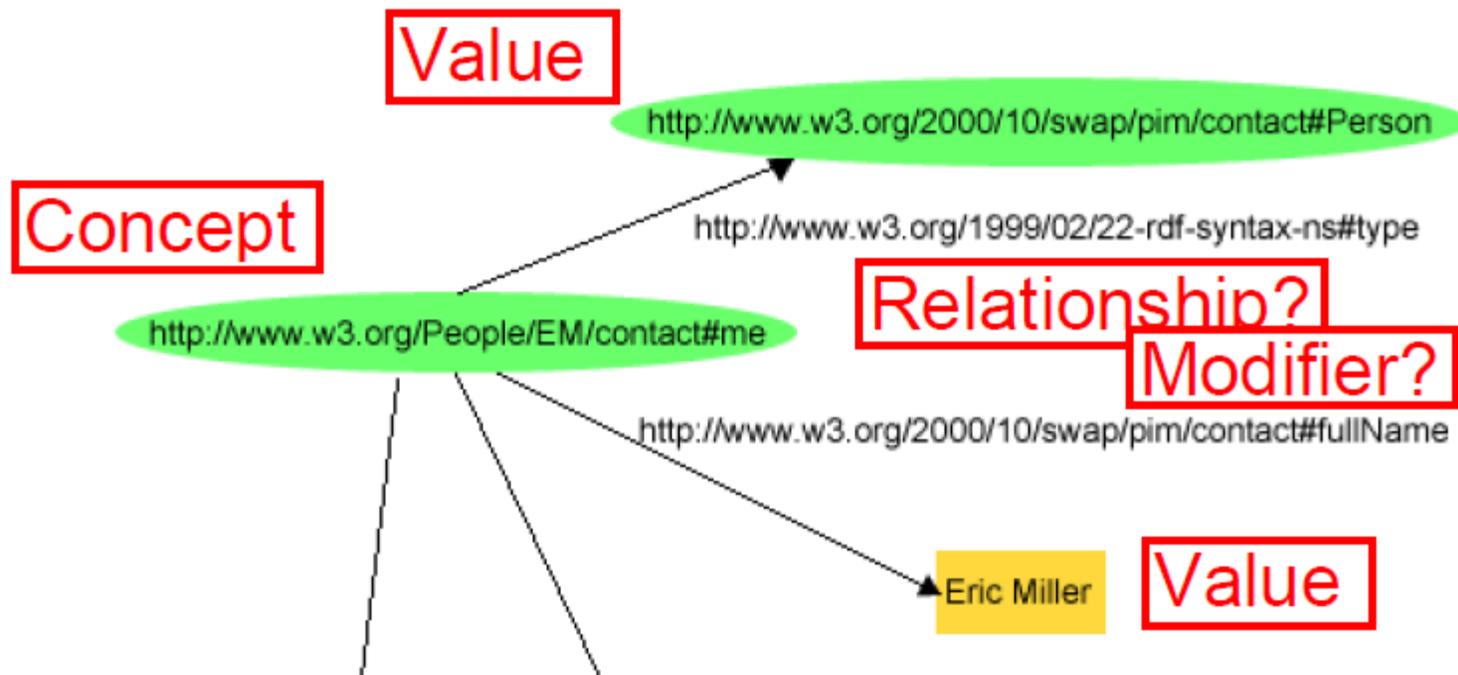
Concept vs. Modifier

- Occasionally it is not clear what is the concept and what is the modifier
 - Staph pneumonia
- Guidance
 - Pick what is closest to the patient = pneumonia

RDF Underpinnings

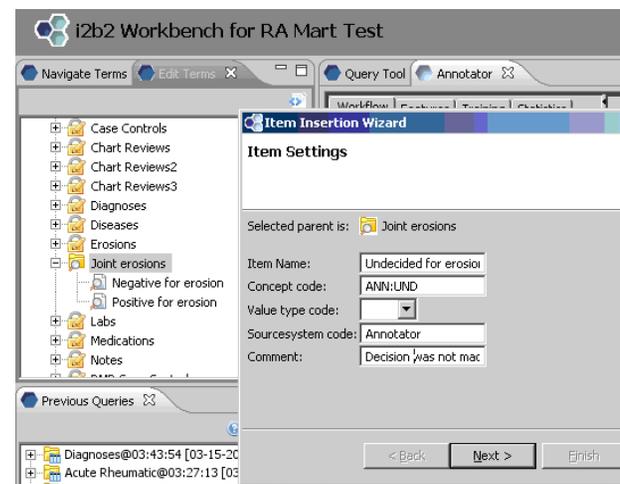
■ RDF Graph from W3C document

- Note that metadata relationships and data relationships are not distinguished
- I2b2 does distinguish, metadata relationships in terminology hierarchies, data relationships with modifiers



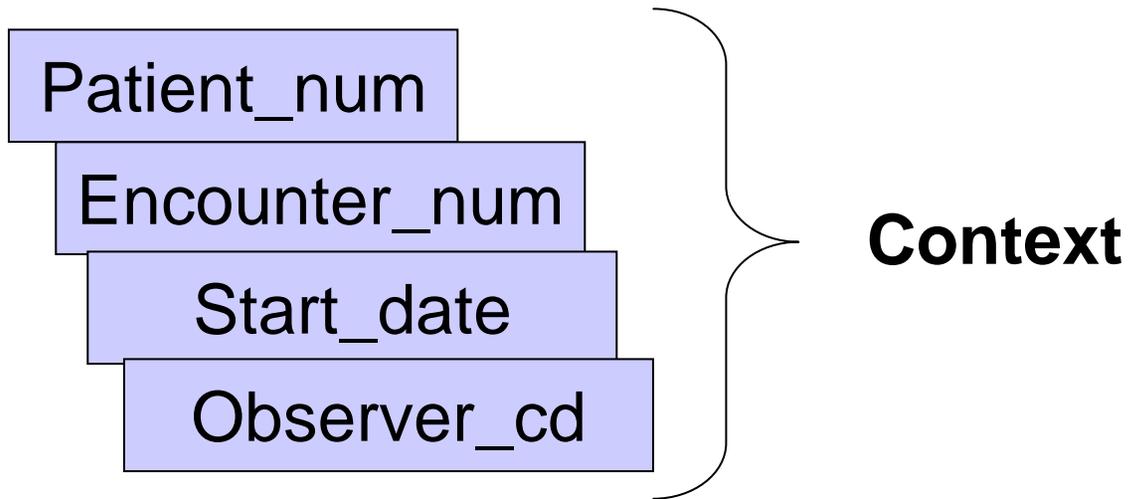
I2b2 metadata relationships – expressed in hierarchies

- ISA – most common relationship in diagnoses, medication, procedure, laboratory test hierarchies
 - Other hierarchies could be expressed in separate hierarchies, but usually aren't
 - Could add as part of Big Parent title – Diagnoses (ISA relationships), Anatomy (PART OF relationships), but would likely confuse more than clarify, relationships are implied.
- ORGANIZEDBY – organization of your desktop, very popular with small, custom terminologies



I2b2 concept relationships – expressed in modifiers

observation_fact	
PK	<u>Patient_Num</u>
PK	<u>Encounter_Num</u>
PK	<u>Concept_CD</u>
PK	<u>Observer_CD</u>
PK	<u>Start_Date</u>
PK	<u>Modifier_CD</u>
PK	<u>Instance_Num</u>
	ValType_CD TVal_Char NVal_Num ValueFlag_CD Observation_Blob



Subject

Predicate

Object

Concept_cd	Modifer_cd	[VALUE]
------------	------------	---------

Expression of 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>	

Modifiers in UI

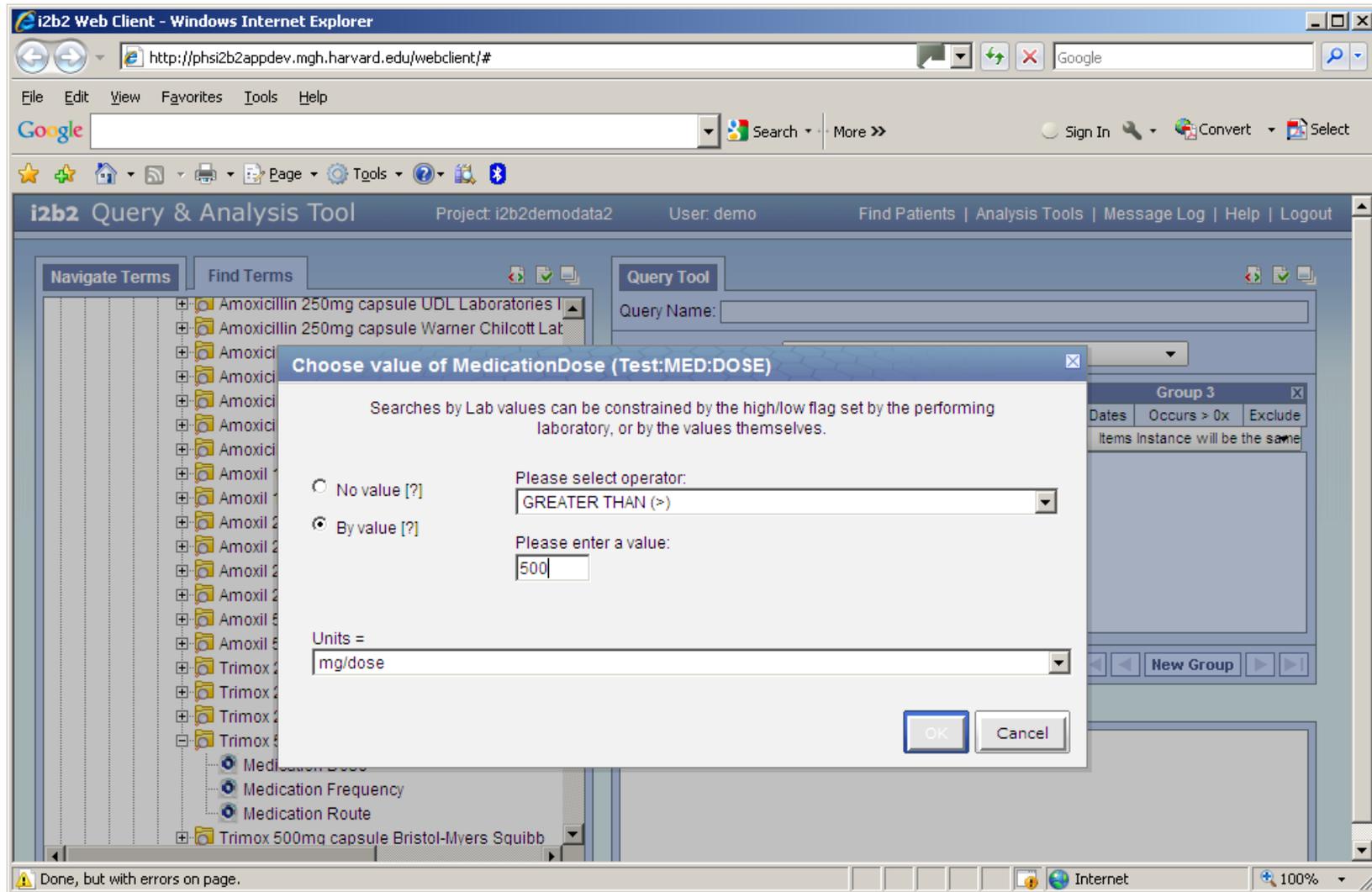
The screenshot displays the i2b2 Query & Analysis Tool interface within a Windows Internet Explorer browser. The browser's address bar shows the URL `http://phsi2b2appdev.mgh.harvard.edu/webclient/#`. The tool's header includes the text "i2b2 Query & Analysis Tool", "Project: i2b2demodata2", "User: demo", and navigation links for "Find Patients", "Analysis Tools", "Message Log", "Help", and "Logout".

The interface is divided into two main sections:

- Navigate Terms / Find Terms:** A list of medication terms is displayed, including various Amoxicillin and Amoxil formulations (e.g., "Amoxicillin 250mg capsule UDL Laboratories", "Amoxil 125 mg/5ml powder for reconstitution S"). A sub-menu is open, showing "Medication Dose", "Medication Frequency", and "Medication Route".
- Query Tool:** This section allows for building a query. It includes a "Query Name" field, a "Temporal Constraint" dropdown set to "Items Instance will be the same", and three groups (Group 1, Group 2, Group 3) for defining query criteria. Each group has columns for "Dates", "Occurs > 0x", and "Exclude". A context menu is open over the term "rimox 500mg capsule Bristol-Myers Squibb" in Group 1, with options for "Delete", "Set Modifier Value", and "drop a term on here".

At the bottom of the Query Tool, there are buttons for "Run Query", "Clear", "Print Query", and "New Group". The "Query Status" section is currently empty.

Modifiers in UI



Modifiers in UI

The screenshot shows the i2b2 Query & Analysis Tool interface within a Windows Internet Explorer browser. The browser's address bar shows the URL `http://phsi2b2appdev.mgh.harvard.edu/webclient/#`. The tool's header includes the text "i2b2 Query & Analysis Tool" and "Project: i2b2demodata2 User: demo".

The interface is divided into two main sections:

- Left Panel (Find Terms):** A tree view of medication terms. The first few items are:
 - Amoxicillin 250mg capsule UDL Laboratories
 - Amoxicillin 250mg capsule Warner Chilcott Lat
 - Amoxicillin 250mg tablet chewable Zenith Gold
 - Amoxicillin 500mg capsule Novopharm USA In
 - Amoxicillin 500mg capsule Teva Pharmaceutic
 - Amoxicillin 500mg capsule Watson/Rugby Lab
 - Amoxicillin 500mg capsule Zenith Goldline Pha
 - Amoxil 125 mg/5ml powder for reconstitution S
 - Amoxil 125 mg/5ml powder for reconstitution S
 - Amoxil 250 mg/5ml powder for reconstitution S
 - Amoxil 250 mg/5ml powder for reconstitution S
 - Amoxil 250mg capsule SmithKline Beecham
 - Amoxil 250mg tablet chewable SmithKline Bee
 - Amoxil 500mg capsule SmithKline Beecham
 - Amoxil 500mg capsule SmithKline Beecham
 - Trimox 250 mg/5ml powder for reconstitution B
 - Trimox 250mg capsule Bristol-Myers Squibb
 - Trimox 250mg capsule Bristol-Myers Squibb
 - Trimox 500mg capsule Bristol-Myers Squibb
 - Medication Dose
 - Medication Frequency
 - Medication Route
 - Trimox 500mg capsule Bristol-Myers Squibb
- Right Panel (Query Tool):** A query builder interface. It includes a "Query Name:" field, a "Temporal Constraint:" dropdown set to "Items Instance will be the same", and three groups for building a query:
 - Group 1:** Contains the text "quibb [Medication Dose > 500]" and a green box with the text "one or more of these".
 - Group 2:** Contains a blue box with the text "AND".
 - Group 3:** Contains a yellow box with the text "drop a term on here".Below the groups are buttons for "Run Query", "Clear", "Print Query", and "New Group".

The status bar at the bottom of the browser shows a warning icon and the text "Done, but with errors on page." and the system tray includes "Internet" and "100%" zoom.

Modifiers in UI

i2b2 Query & Analysis Tool Project: i2b2demodata2 User: demo Find Patients | Analysis Tools | Message Log | Help | Logout

Navigate Terms **Find Terms**

Search by Names Search by Codes

Containing Find Any Category

- Trimox 250 mg/5ml powder for reconstitution Bristol-Myers Squibb
- Trimox 250mg capsule Bristol-Myers Squibb
- Trimox 250mg capsule Bristol-Myers Squibb
- Trimox 500mg capsule Bristol-Myers Squibb
- Trimox 500mg capsule Bristol-Myers Squibb

Find Modifiers for Trimox 500mg capsule Bristol-Myers Squibb

Containing

Find Name Find Code Get All

- Medication Dose
- Medication Frequency
- Medication Route

Query Tool

Query Name:

Temporal Constraint:

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude
Squibb [Medication Dose > 500]			Medication Route = ("By Mouth")			Items Instance will be the same		

one or more of these AND one or more of these AND drop a term on here

Run Query Clear Print Query 2 Groups New Group

Query Status

Finished Query: "Medicat-Medicat@19:32:26" [6.8 secs]
Compute Time: 3.9 secs

Number of patients for "Medicat-Medicat@19:32:26"
patient_count: 6

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

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

[Also includes ETL directives]

```
<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”.

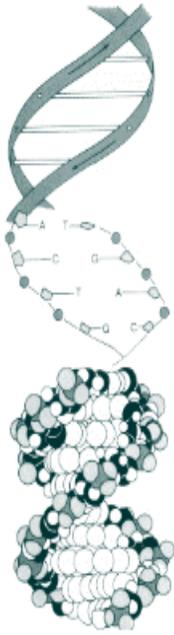
Query by values

The screenshot displays the i2b2 Workbench interface for an Asthma Test Group. The main window is titled "i2b2 Workbench" and shows a "Query Tool" window. The query name is "Calcium (Group:@03:28:30)". The query tool has three groups (Group 1, Group 2, Group 3) with columns for "Dates", "Occurs > 0x", and "Exclude". A query is entered in Group 1: "Calcium (Group:CA) > 10.6".

A dialog box titled "Choose value of Calcium (Group:CA)" is open, showing options for constraining the search by lab values. The dialog includes a "Please select operator:" dropdown set to "GREATER THAN (>)" and a "Please enter value:" text box containing "10.6". The units are set to "mg/dl".

The "Previous Queries" window on the left shows a list of queries, including "Calcium (Group:@03:28:30 [04-03-2008] [wp066])" and "Acetone (Group:@10:03:11 [04-01-2008] [wp066])". The "Results of Calcium (Group:@03:28:30)" window shows a patient set of 2606 patients.

The "Image View" and "Timeline View" tabs are visible at the bottom of the main window. The "Patient Set" is "Patient Set: 2606 patients" and the "start" and "increment" values are 21 and 10, respectively.



i2b2 Ontology Services

Lori Phillips MS

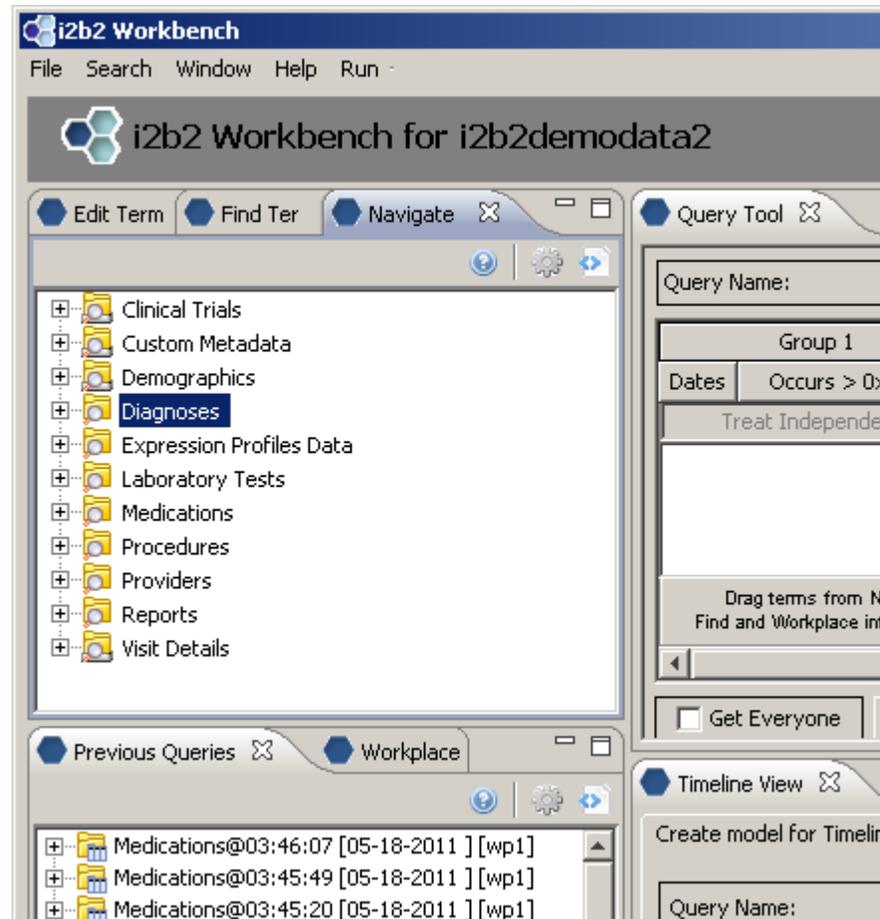
i2b2 AUG

June 29, 2011

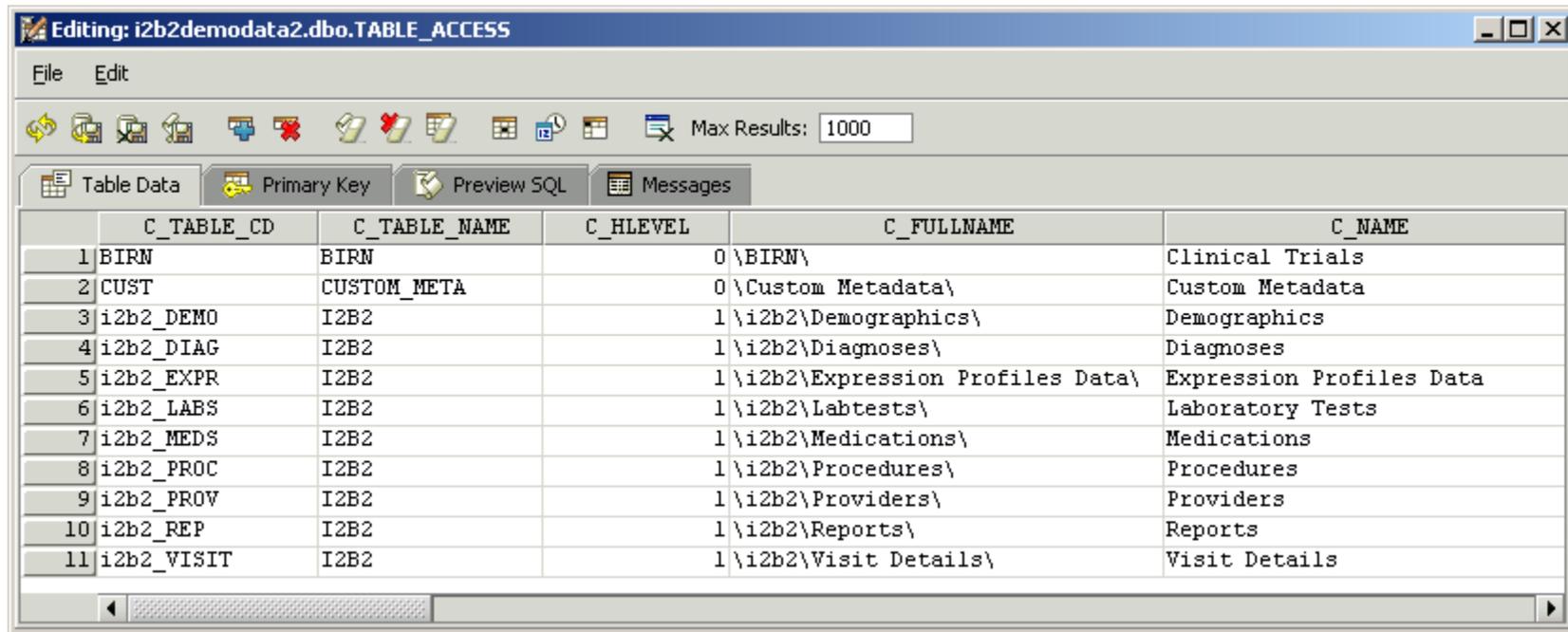
Overview

- Basic Ontology features
 - UI
 - Organization
- Advanced topics
 - Support for modifiers
 - NCBO Ontology extraction tool
- Query building
 - Concept_dimension based queries
 - Non-concept_dimension based queries

i2b2 Navigate Terms View



Creating the root level nodes



	C_TABLE_CD	C_TABLE_NAME	C_HLEVEL	C_FULLNAME	C_NAME
1	BIRN	BIRN	0	\BIRN\	Clinical Trials
2	CUST	CUSTOM_META	0	\Custom Metadata\	Custom Metadata
3	i2b2_DEMO	I2B2	1	\i2b2\Demographics\	Demographics
4	i2b2_DIAG	I2B2	1	\i2b2\Diagnoses\	Diagnoses
5	i2b2_EXPR	I2B2	1	\i2b2\Expression Profiles Data\	Expression Profiles Data
6	i2b2_LABS	I2B2	1	\i2b2\Labtests\	Laboratory Tests
7	i2b2_MEDS	I2B2	1	\i2b2\Medications\	Medications
8	i2b2_PROC	I2B2	1	\i2b2\Procedures\	Procedures
9	i2b2_PROV	I2B2	1	\i2b2\Providers\	Providers
10	i2b2_REP	I2B2	1	\i2b2\Reports\	Reports
11	i2b2_VISIT	I2B2	1	\i2b2\Visit Details\	Visit Details

- One to one mapping between the entries of the table_access table and the root level nodes for a given project.

Building local metadata

- Concept key == \\c_table_cd\c_fullname\

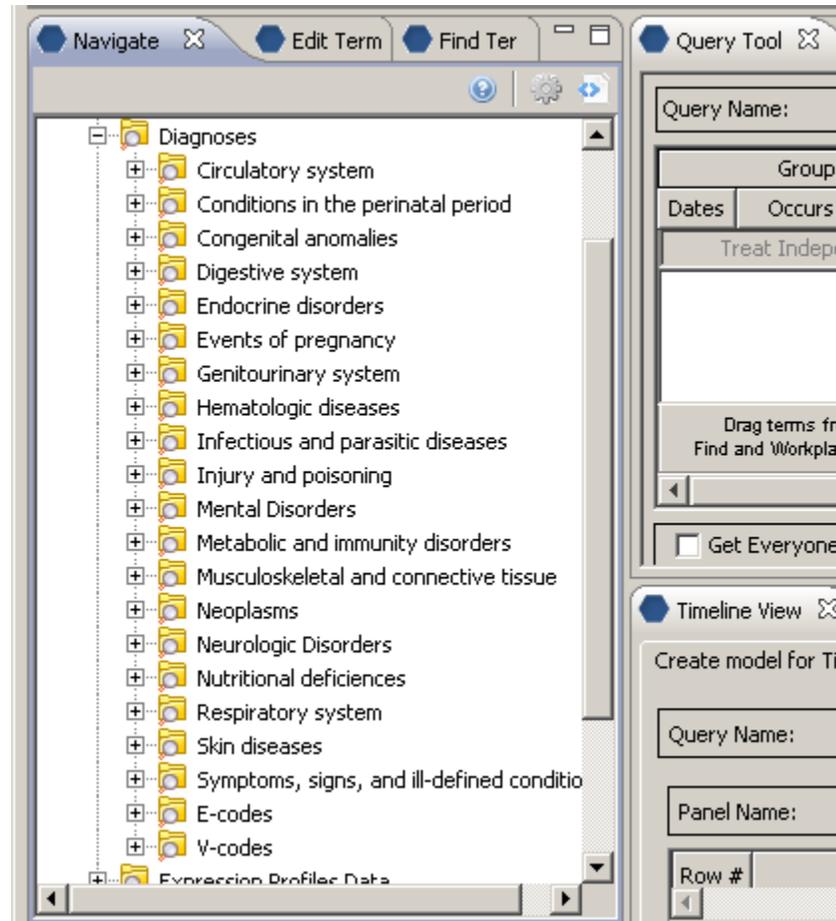
```
<concept>
  <level>1</level>
  <key>\\i2b2_DIAG\i2b2\Diagnoses\</key>
  <name>Diagnoses</name>
  ...
</concept>
```

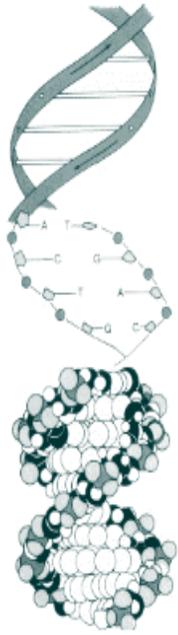
	C_TABLE_CD	C_TABLE_NAME	C_HLEVEL	C_FULLNAME	C_NAME
1	BIRN	BIRN	0	\BIRN\	Clinical Trials
2	CUST	CUSTOM_META	0	\Custom Metadata\	Custom Metadata
3	i2b2_DEMO	I2B2	1	\\i2b2\Demographics\	Demographics
4	i2b2_DIAG	I2B2	1	\\i2b2\Diagnoses\	Diagnoses
5	i2b2_EXPR	I2B2	1	\\i2b2\Expression Profiles Data\	Expression Profiles Data
6	i2b2_LABS	I2B2	1	\\i2b2\Labtests\	Laboratory Tests
7	i2b2_MEDS	I2B2	1	\\i2b2\Medications\	Medications

Children of concept Diagnoses [\i2b2\Diagnoses\]

C_HLEVEL	C_FULLNAME	C_NAME
2	\i2b2\Diagnoses\Circulatory system (390-459)\	Circulatory system
2	\i2b2\Diagnoses\Conditions in the perinatal period (760-779)\	Conditions in the perinatal period
2	\i2b2\Diagnoses\Congenital anomalies (740-759)\	Congenital anomalies
2	\i2b2\Diagnoses\Digestive system (520-579)\	Digestive system
2	\i2b2\Diagnoses\Endocrine disorders (240-259)\	Endocrine disorders
2	\i2b2\Diagnoses\Events of pregnancy (630-677)\	Events of pregnancy
2	\i2b2\Diagnoses\Genitourinary system (580-629)\	Genitourinary system
2	\i2b2\Diagnoses\Hematologic diseases (280-289)\	Hematologic diseases
2	\i2b2\Diagnoses\Infectious and parasitic diseases (001-139)\	Infectious and parasitic diseases
2	\i2b2\Diagnoses\Injury and poisoning (800-999)\	Injury and poisoning
2	\i2b2\Diagnoses\Mental Disorders (290-319)\	Mental Disorders
2	\i2b2\Diagnoses\Metabolic and immunity disorders (270-279)\	Metabolic and immunity disorders
2	\i2b2\Diagnoses\Musculoskeletal and connective tissue (710-739)\	Musculoskeletal and connective tissue
2	\i2b2\Diagnoses\Neoplasms (140-239)\	Neoplasms
2	\i2b2\Diagnoses\Neurologic Disorders (320-389)\	Neurologic Disorders
2	\i2b2\Diagnoses\Nutritional deficiencies (260-269)\	Nutritional deficiencies
2	\i2b2\Diagnoses\Respiratory system (460-519)\	Respiratory system
2	\i2b2\Diagnoses\Skin diseases (680-709)\	Skin diseases
2	\i2b2\Diagnoses\Symptoms, signs, and ill-defined conditions (780-799)\	Symptoms, signs, and ill-defined conditions
2	\i2b2\Diagnoses\zz E-codes\	E-codes
2	\i2b2\Diagnoses\zz V-codes\	V-codes

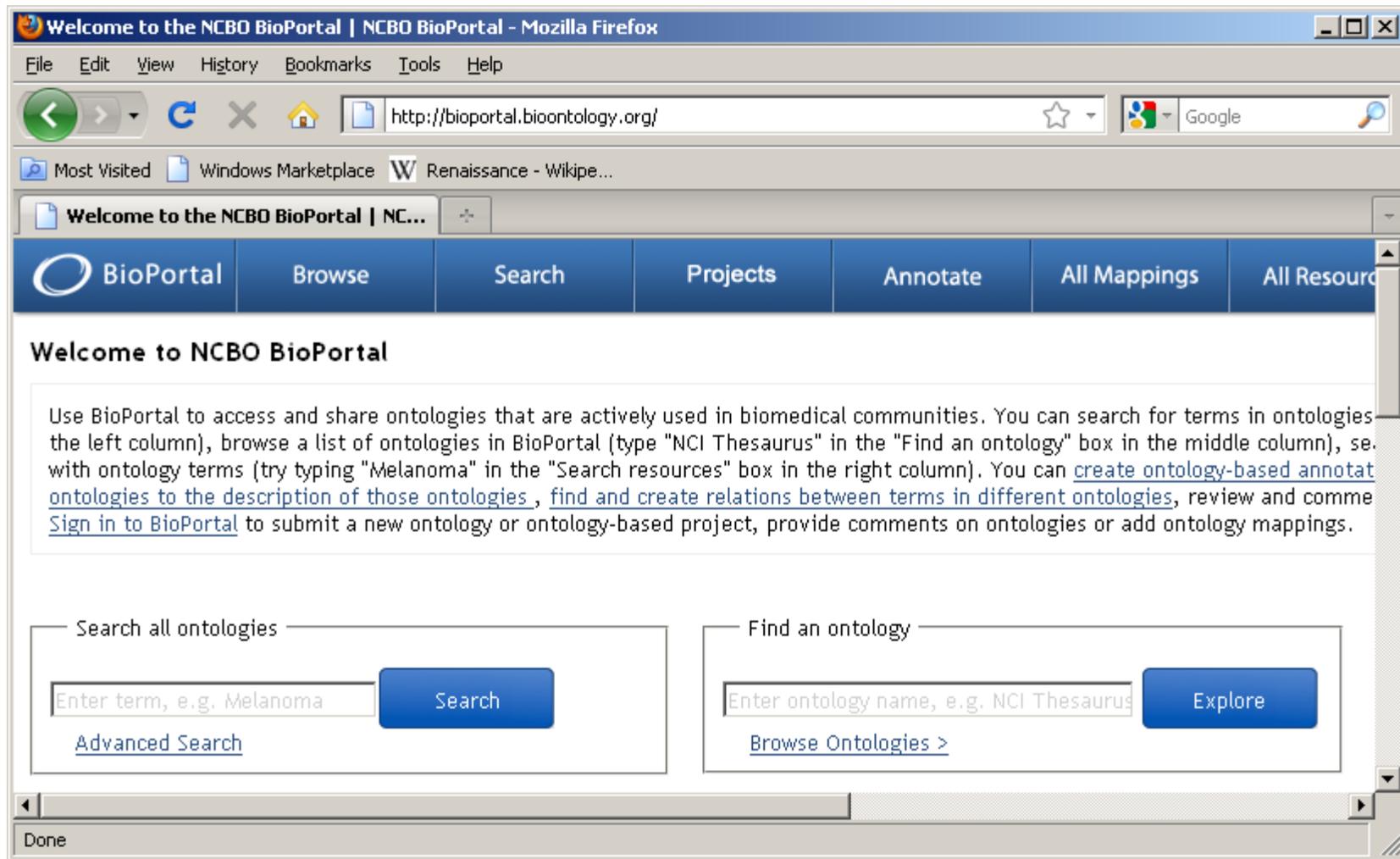
Children of concept Diagnoses display





NCBO Ontology Tools

Creating metadata for standard ontologies using BioPortal



The screenshot shows a Mozilla Firefox browser window displaying the NCBO BioPortal website. The address bar shows the URL <http://bioportal.bioontology.org/>. The browser's menu bar includes File, Edit, View, History, Bookmarks, Tools, and Help. The page title is "Welcome to the NCBO BioPortal | NCBO BioPortal - Mozilla Firefox".

The website's navigation bar features the BioPortal logo and several menu items: Browse, Search, Projects, Annotate, All Mappings, and All Resources. Below the navigation bar, the main heading reads "Welcome to NCBO BioPortal".

A large text block provides a welcome message and instructions: "Use BioPortal to access and share ontologies that are actively used in biomedical communities. You can search for terms in ontologies (the left column), browse a list of ontologies in BioPortal (type 'NCI Thesaurus' in the 'Find an ontology' box in the middle column), search with ontology terms (try typing 'Melanoma' in the 'Search resources' box in the right column). You can [create ontology-based annotations to the description of those ontologies](#), [find and create relations between terms in different ontologies](#), review and comment on ontologies, [Sign in to BioPortal](#) to submit a new ontology or ontology-based project, provide comments on ontologies or add ontology mappings."

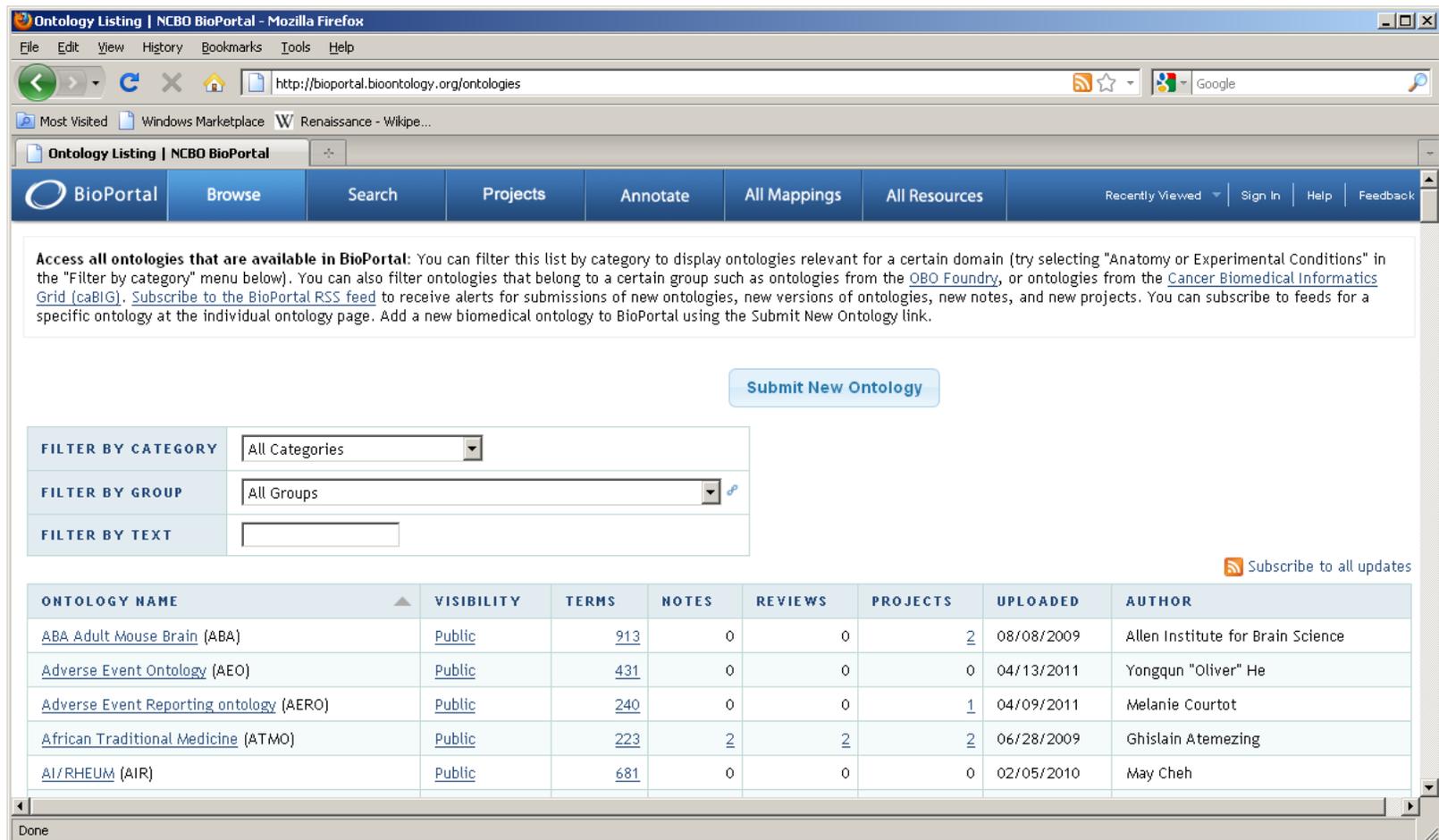
Below the text, there are two search boxes:

- Search all ontologies:** A text input field with the placeholder "Enter term, e.g. Melanoma" and a blue "Search" button. Below the input field is a link for "Advanced Search".
- Find an ontology:** A text input field with the placeholder "Enter ontology name, e.g. NCI Thesaurus" and a blue "Explore" button. Below the input field is a link for "Browse Ontologies >".

The browser's status bar at the bottom shows "Done".

NCBO BioPortal

- Hosts over 250 ontologies



The screenshot shows the NCBO BioPortal website in a Mozilla Firefox browser window. The address bar displays <http://bioportal.bioontology.org/ontologies>. The navigation menu includes links for Browse, Search, Projects, Annotate, All Mappings, and All Resources. A text box provides instructions on how to filter ontologies by category or group, and offers a link to submit a new ontology. Below this, there are three filter sections: 'FILTER BY CATEGORY' (set to 'All Categories'), 'FILTER BY GROUP' (set to 'All Groups'), and 'FILTER BY TEXT' (empty). A 'Submit New Ontology' button is located above the filters. A table lists several ontologies with columns for Ontology Name, Visibility, Terms, Notes, Reviews, Projects, Uploaded, and Author. A 'Subscribe to all updates' link is visible to the right of the table.

Access all ontologies that are available in BioPortal: You can filter this list by category to display ontologies relevant for a certain domain (try selecting "Anatomy or Experimental Conditions" in the "Filter by category" menu below). You can also filter ontologies that belong to a certain group such as ontologies from the [OBO Foundry](#), or ontologies from the [Cancer Biomedical Informatics Grid \(caBIG\)](#). [Subscribe to the BioPortal RSS feed](#) to receive alerts for submissions of new ontologies, new versions of ontologies, new notes, and new projects. You can subscribe to feeds for a specific ontology at the individual ontology page. Add a new biomedical ontology to BioPortal using the Submit New Ontology link.

[Submit New Ontology](#)

ONTOLGY NAME	VISIBILITY	TERMS	NOTES	REVIEWS	PROJECTS	UPLOADED	AUTHOR
ABA Adult Mouse Brain (ABA)	Public	913	0	0	2	08/08/2009	Allen Institute for Brain Science
Adverse Event Ontology (AEO)	Public	431	0	0	0	04/13/2011	Yongqun "Oliver" He
Adverse Event Reporting ontology (AERO)	Public	240	0	0	1	04/09/2011	Melanie Courtot
African Traditional Medicine (ATMO)	Public	223	2	2	2	06/28/2009	Ghislain Atemezing
AI/RHEUM (AIR)	Public	681	0	0	0	02/05/2010	May Cheh

[Subscribe to all updates](#)

Details summary page

International Classification of Diseases - Summary | NCBO BioPortal - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://bioportal.bioontology.org/ontologies/1101

Most Visited Windows Marketplace Renaissance - Wikipe...

International Classification of Diseas...

BioPortal Browse Search Projects Annotate All Mappings All Resources Recently Viewed Sign In Help

International Classification of Diseases Summary

Details

ONTOLOGY ID:	1101
BIOPORTAL PURL:	http://purl.bioontology.org/ontology/ICD-9
STATUS:	
FORMAT:	RRF
CATEGORIES:	
GROUPS:	Unified Medical Language System WHO Family of International Classifications
CONTACT:	The World Health Organization, info@who.int
HOME PAGE:	http://www.who.int/classifications/icd/en/
PUBLICATIONS PAGE:	
DOCUMENTATION PAGE:	
DESCRIPTION:	The ICD is the international standard diagnostic classification for all general epidemiological, many health management purposes and clinical use.

Metrics

NUMBER OF CLASSES:	21669
NUMBER OF INDIVIDUALS:	0
NUMBER OF PROPERTIES:	10
MAXIMUM DEPTH:	6
MAXIMUM NUMBER OF SIBLINGS:	25
AVERAGE NUMBER OF SIBLINGS:	1
CLASSES WITH A SINGLE SUBCLASS:	23
CLASSES WITH MORE THAN 25 SUBCLASSES:	0
CLASSES WITH NO AUTHOR:	0
CLASSES WITH NO DEFINITION:	0

Done

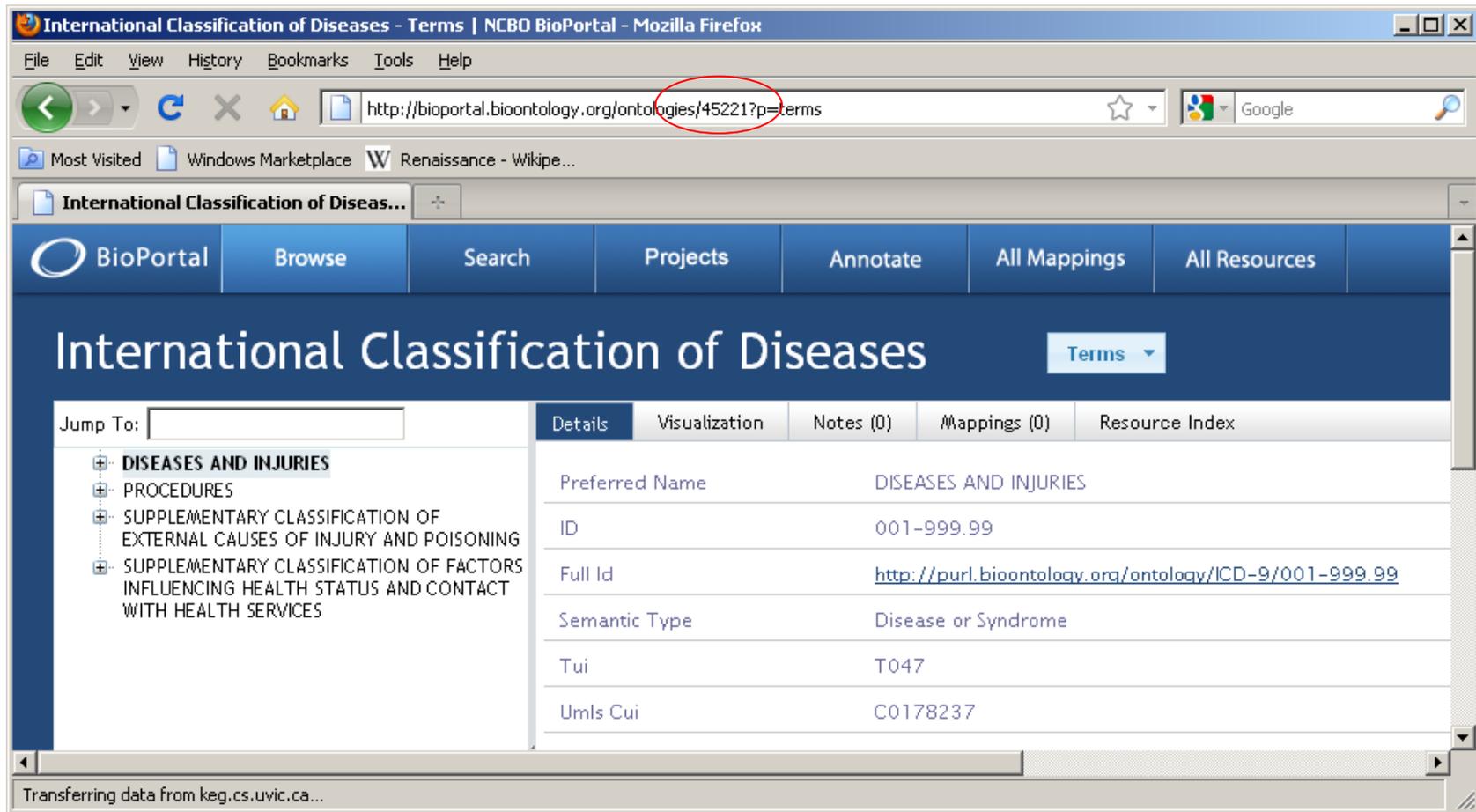
Locate ontology id for version of interest [45221]

The screenshot shows a Mozilla Firefox browser window with the title "International Classification of Diseases - Summary | NCBO BioPortal - Mozilla Firefox". The address bar contains the URL "http://bioportal.bioontology.org/ontologies/1101". The main content area displays the "Versions" section for the International Classification of Diseases ontology, featuring a table with the following data:

VERSION	RELEASE DATE	DOWNLOADS
9	05/16/2008	Ontology
9	05/16/2008	Ontology

The address bar at the bottom of the browser window shows the URL "http://bioportal.bioontology.org/ontologies/45221?p=terms", with the number "45221" circled in red.

BioPortal – ICD9



International Classification of Diseases - Terms | NCBO BioPortal - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://bioportal.bioontology.org/ontologies/45221?p=terms

Most Visited Windows Marketplace Renaissance - Wikiped...

International Classification of Diseas...

BioPortal Browse Search Projects Annotate All Mappings All Resources

International Classification of Diseases

Terms

Jump To:

- ⊕ **DISEASES AND INJURIES**
- ⊕ PROCEDURES
- ⊕ SUPPLEMENTARY CLASSIFICATION OF EXTERNAL CAUSES OF INJURY AND POISONING
- ⊕ SUPPLEMENTARY CLASSIFICATION OF FACTORS INFLUENCING HEALTH STATUS AND CONTACT WITH HEALTH SERVICES

Details	Visualization	Notes (0)	Mappings (0)	Resource Index
Preferred Name	DISEASES AND INJURIES			
ID	001-999.99			
Full Id	http://purl.bioontology.org/ontology/ICD-9/001-999.99			
Semantic Type	Disease or Syndrome			
Tui	T047			
Umls Cui	C0178237			

Transferring data from keg.cs.uvic.ca...

NCBO Ontology Extraction Strategy

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

```

<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://ihtsdo.org/snomedct/clinicalFinding#10000006</id>
      <fullId>http://ihtsdo.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>
      </entry>
    </classBean>
  </classBeanResultList>
</contents>
  </data>

```

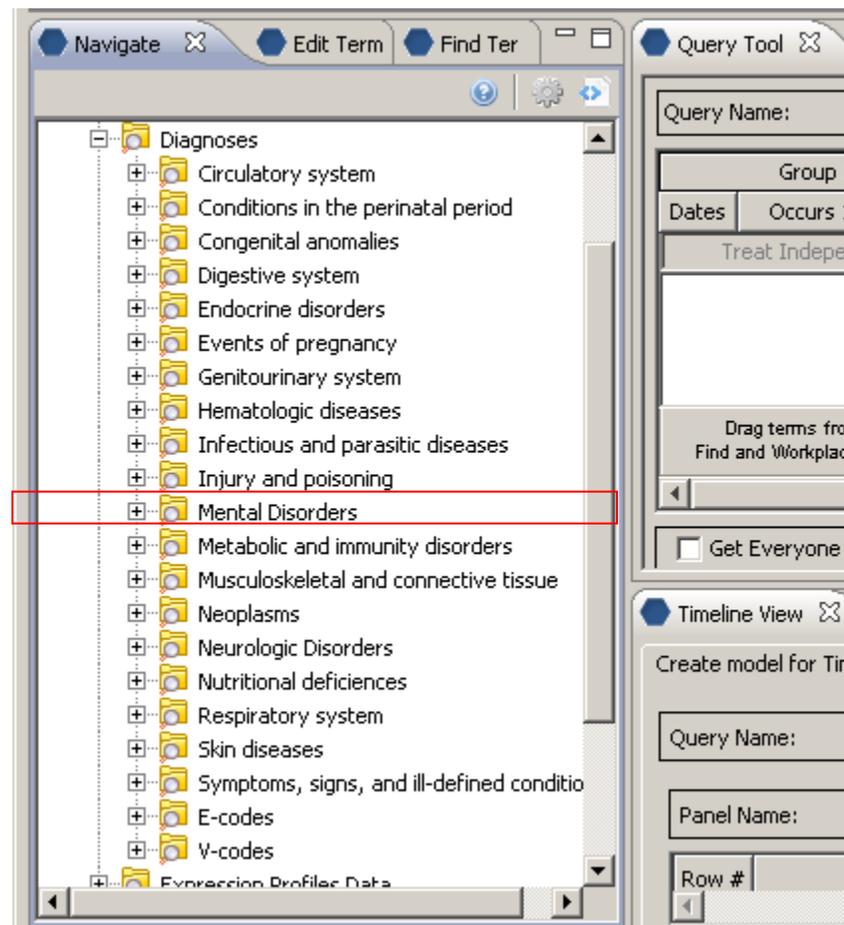


METADATA COLUMN	DATA TYPE (ORACLE)	DATA TYPE (SQL)
C_HLEVEL	INT	INT
C_FULLNAME	VARCHAR2(700)	VARCHAR(700)
C_NAME	VARCHAR2(2000)	VARCHAR(2000)
C_SYNONYM_CD	CHAR(1)	CHAR(1)
C_VISUALATTRIBUTES	CHAR(3)	CHAR(3)
C_BASECODE	VARCHAR2(50)	VARCHAR(50)
C_METADATAXML	CLOB	TEXT
C_FACTTABLECOLUMN	VARCHAR2(50)	VARCHAR(50)
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

\ Diagnoses \
Mental Disorders \

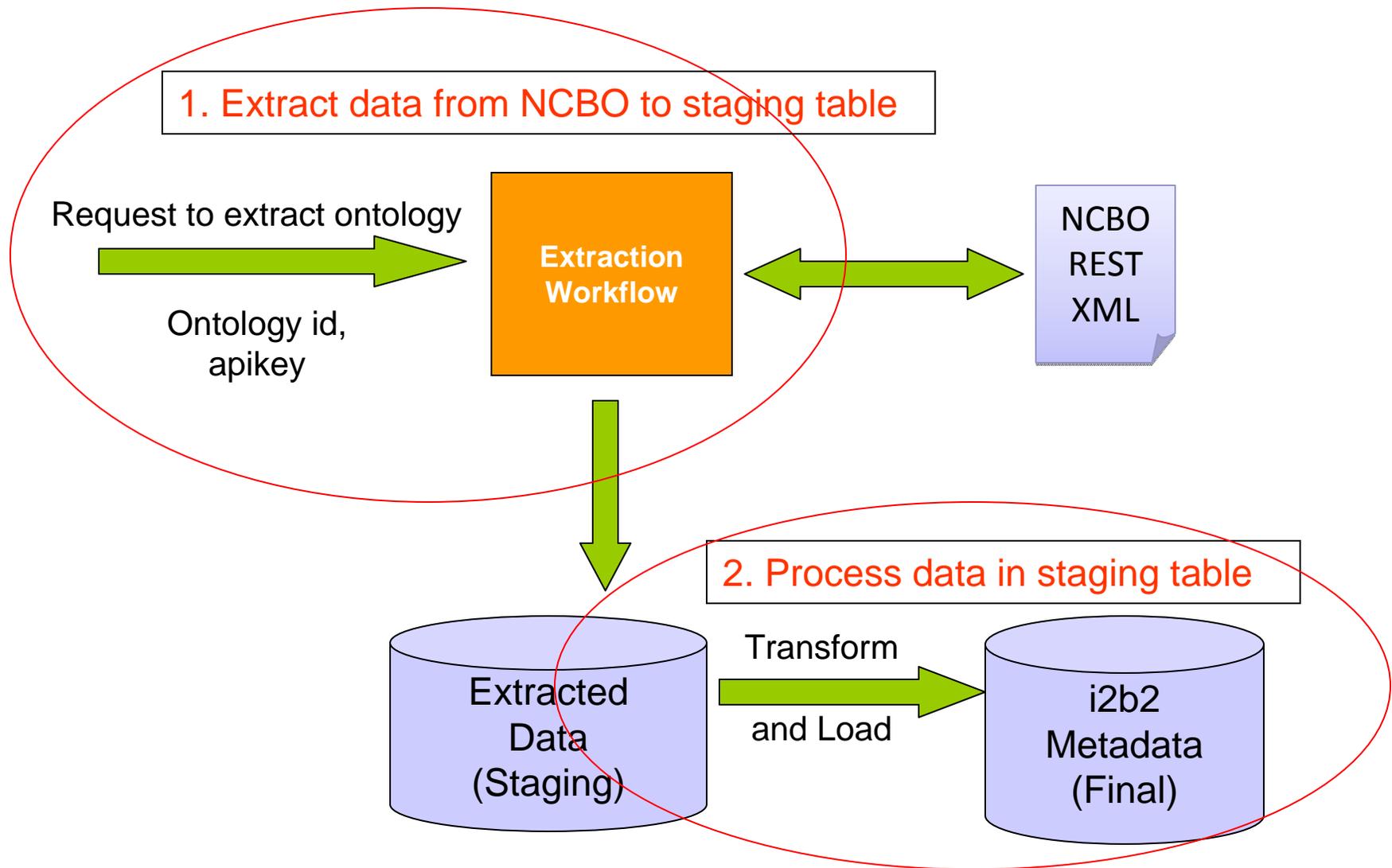


Challenges..

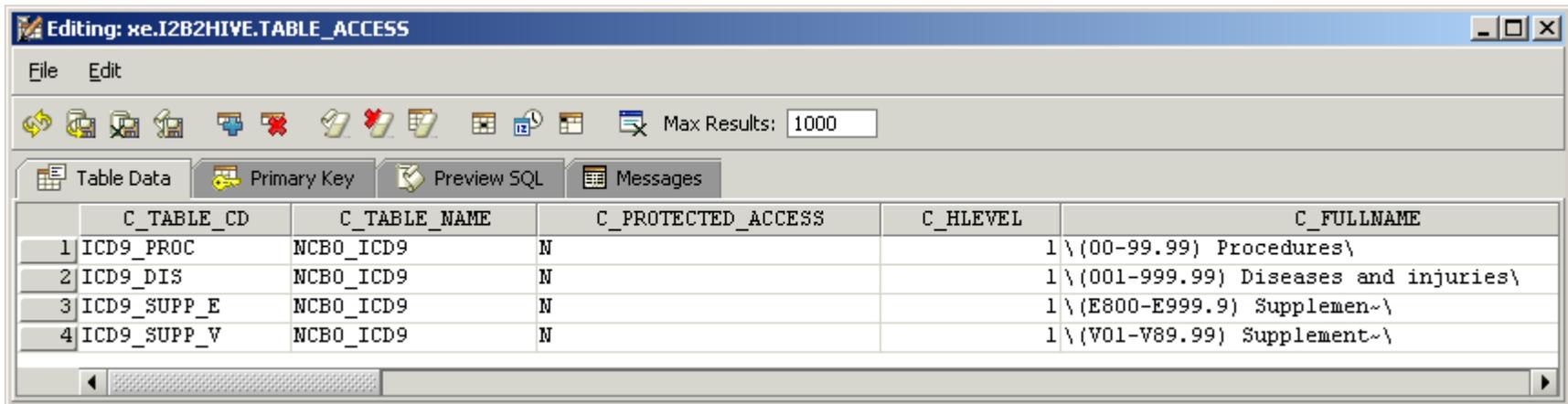
- BioPortal REST service that enables pull of concepts includes immediate parent/child info only
 - **i2b2 c_fullname , c_tooltip must be computed**

```
<data>
  <classBean>
    <id>001-999.99 </id>
    <label>Diseases and injuries</label>
    <relations>
      <entry>
        <string>SuperClass</string>
        <list>
          <classBean>
            <id>290-319.99</id>
            <label>Mental disorders</label>
          </classBean>
        </list>
      </entry>
    </relations>
  </classBean>
</data>
```

NCBO Ontology Extraction Workflow

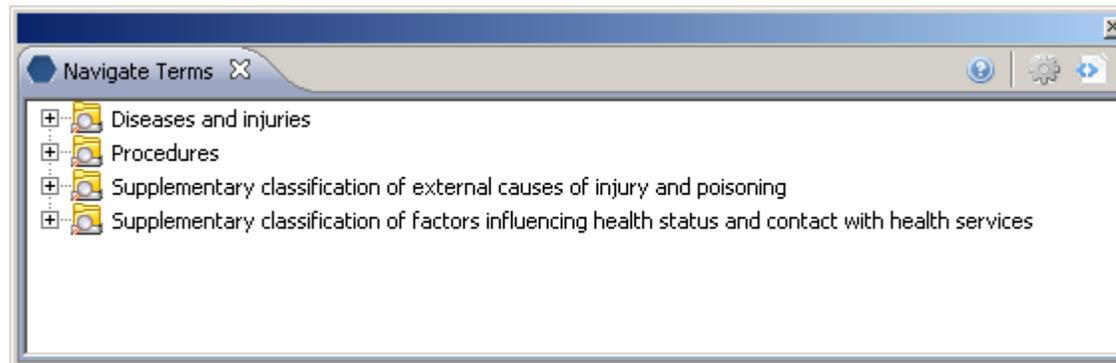


Display extracted ontology



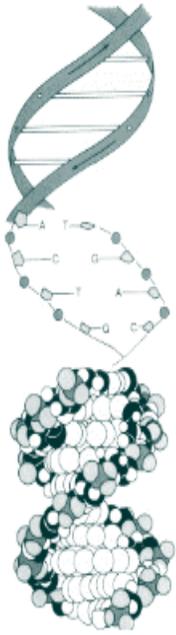
The screenshot shows a database editor window titled "Editing: xe.I2B2HIVE.TABLE_ACCESS". The window contains a table with the following columns: C_TABLE_CD, C_TABLE_NAME, C_PROTECTED_ACCESS, C_HLEVEL, and C_FULLNAME. The table contains four rows of data.

	C_TABLE_CD	C_TABLE_NAME	C_PROTECTED_ACCESS	C_HLEVEL	C_FULLNAME
1	ICD9_PROC	NCBO_ICD9	N	1	\{(00-99.99) Procedures\}
2	ICD9_DIS	NCBO_ICD9	N	1	\{(001-999.99) Diseases and injuries\}
3	ICD9_SUPP_E	NCBO_ICD9	N	1	\{(E800-E999.9) Supplemen~\}
4	ICD9_SUPP_V	NCBO_ICD9	N	1	\{(V01-V89.99) Supplemen~\}



The screenshot shows a "Navigate Terms" window with a list of ontology terms. Each term is preceded by a plus sign icon, indicating it is expandable.

- + Diseases and injuries
- + Procedures
- + Supplementary classification of external causes of injury and poisoning
- + Supplementary classification of factors influencing health status and contact with health services



Ontology Support for Modifiers

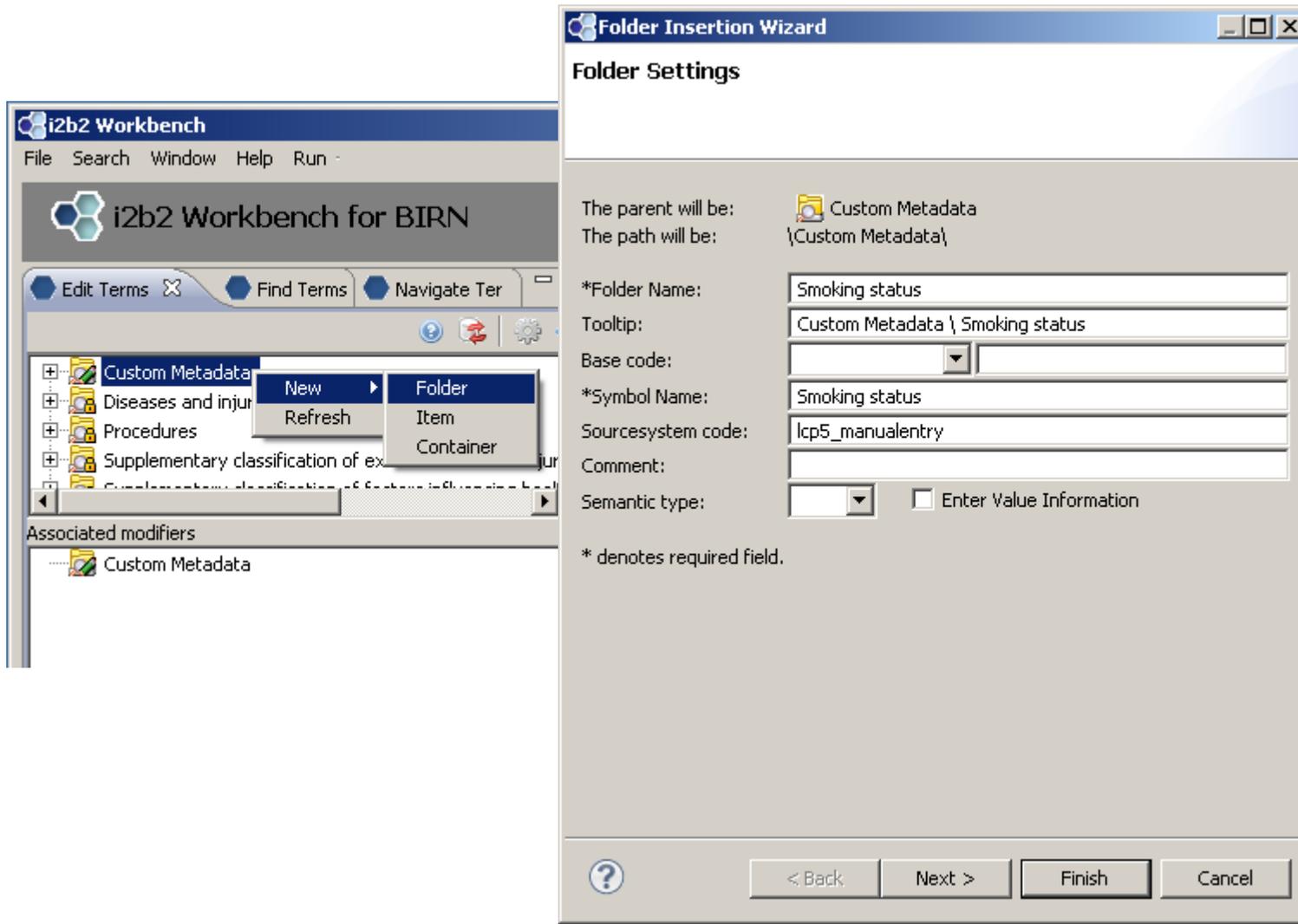
Creating project-specific (localized) metadata

- Table CUSTOM_META is provided for creating metadata within the Edit Terms View
 - Great for small, non-standard ontologies
 - Visual attribute ends in 'E' (editable)

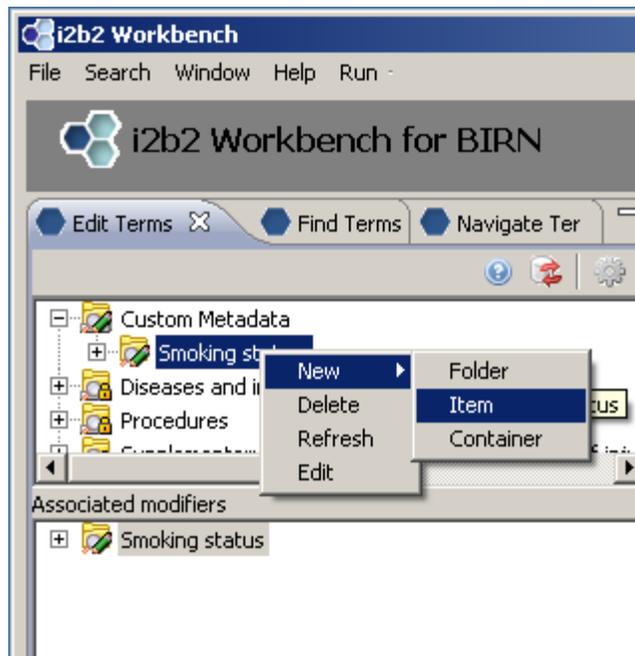
The screenshot shows a window titled "Editing: xe.I2B2HIVE.TABLE_ACCESS" with a menu bar (File, Edit) and a toolbar. Below the toolbar are tabs for "Table Data", "Primary Key", "Preview SQL", and "Edit in Window". The "Table Data" tab is active, displaying a table with the following data:

	C_TABLE_CD	C_TABLE_NAME	C_HLEVEL	C_VISUALATTRIBUTES	C_FULLNAME
1	ICD9_PROC	NCBO_ICD9	1 CA		\{00-99.99} Procedures\
2	ICD9_DIS	NCBO_ICD9	1 CA		\{001-999.99} Diseases and injuries\
3	ICD9_SUPP_E	NCBO_ICD9	1 CA		\{E800-E999.9} Supplemen~\
4	ICD9_SUPP_V	NCBO_ICD9	1 CA		\{V01-V89.99} Supplement~\
5	CUST	CUSTOM_META	0 CAE		\Custom Metadata\

Create a 'Smoking Status' concept folder



Add three concepts...



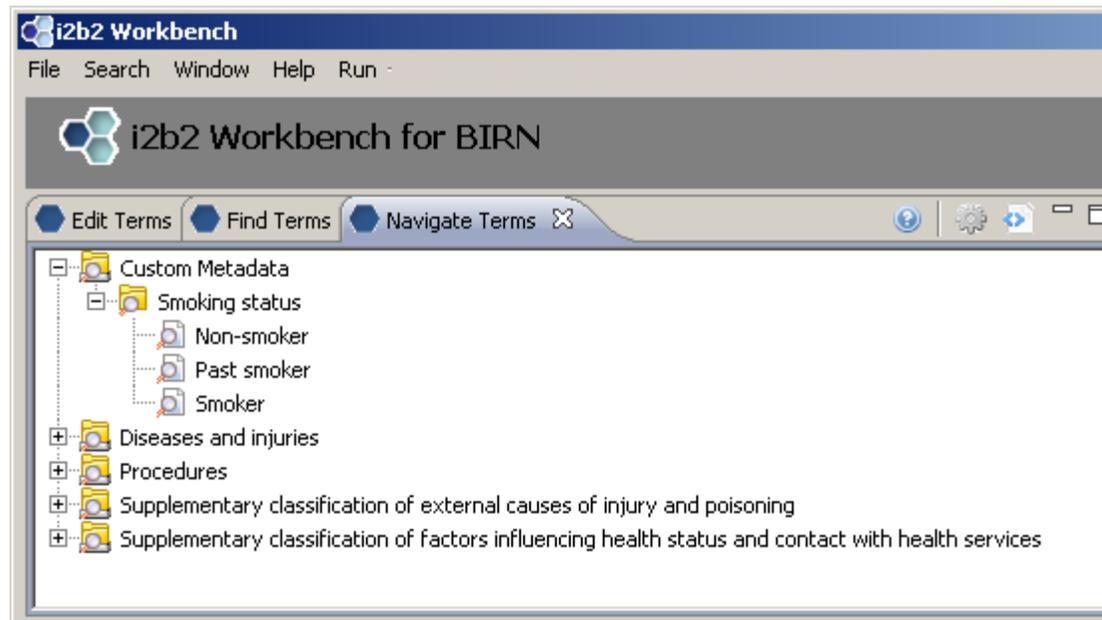
The "Item Insertion Wizard" dialog box is shown, titled "Item Settings". It contains the following fields and options:

- The parent will be: Smoking status
- The path will be: {Custom Metadata}\Smoking status\
- *Item Name: Smoker
- Tooltip: Custom Metadata \ Smoking status \ Smoker
- *Base code: LCS-I2B2 (dropdown) | smoker |
- *Symbol Name: Smoker
- Sourcesystem code: lcp5_manualentry
- Comment: (empty text box)
- Semantic type: (empty dropdown) Enter Value Information

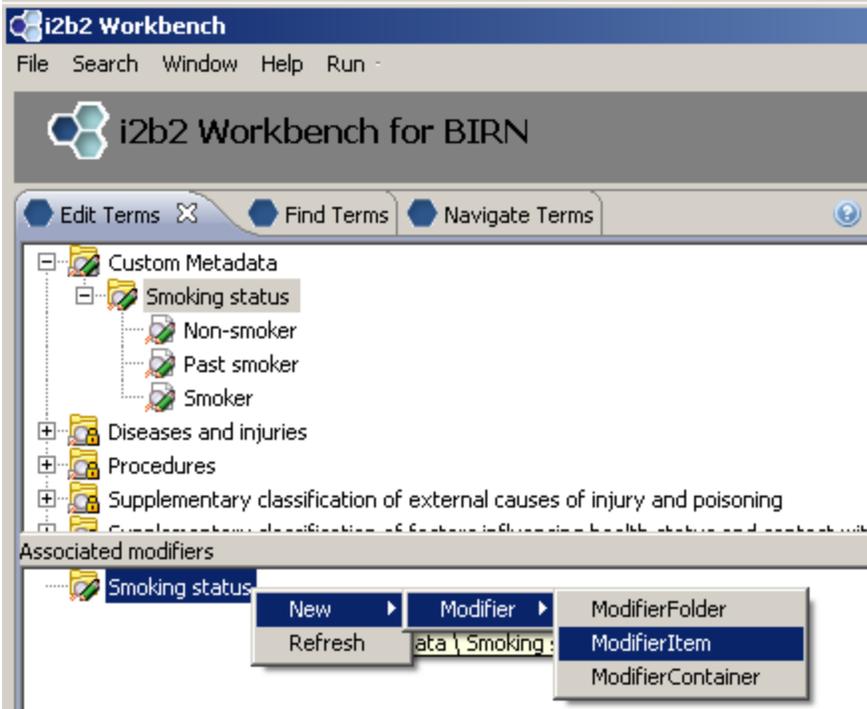
* denotes required field.

At the bottom of the dialog are buttons for "< Back", "Next >", "Finish", and "Cancel", along with a help icon (?) on the left.

New folder with three new concepts:

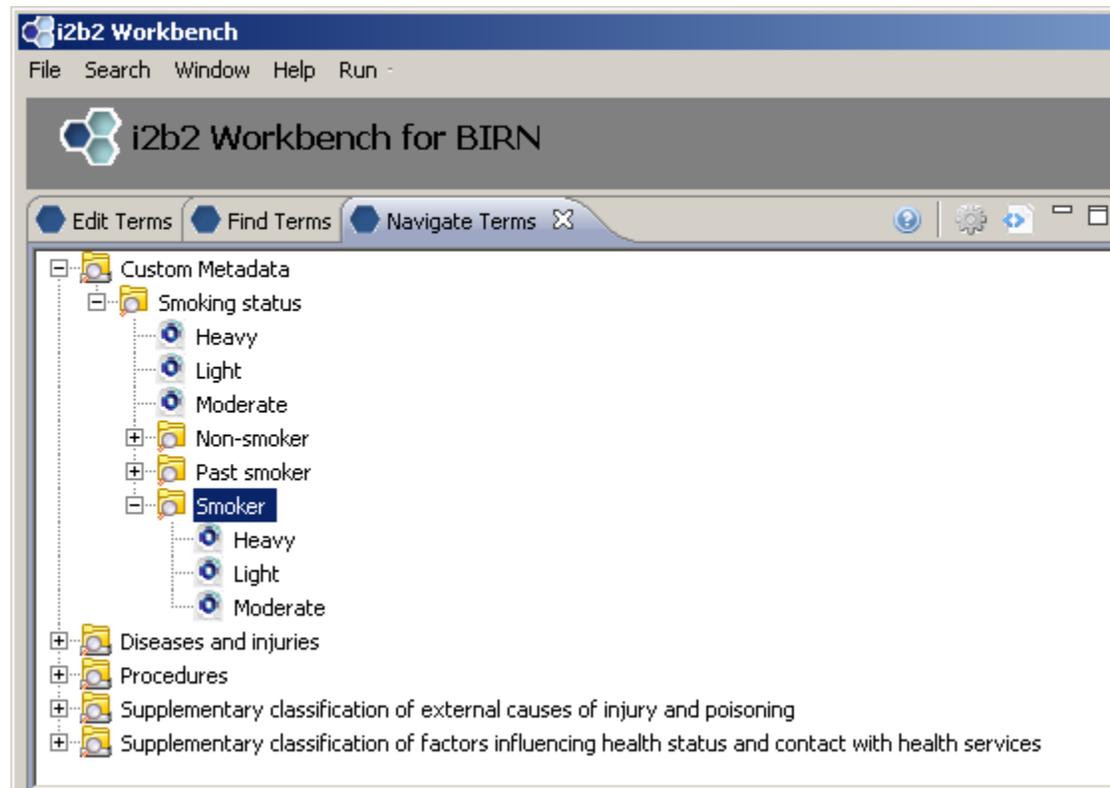


Creating modifiers for existing terms



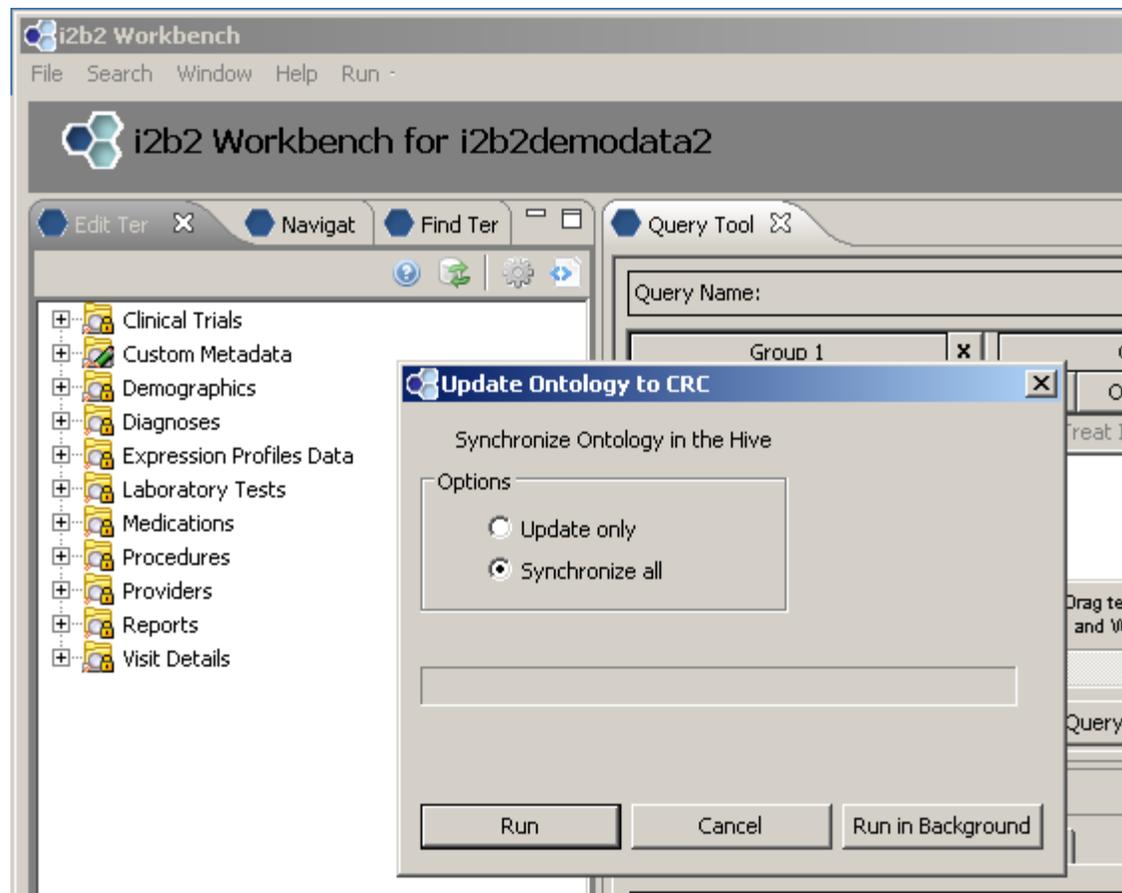
The screenshot shows the "ModifierItem Insertion Wizard" dialog box. The title bar reads "ModifierItem Insertion Wizard". The main heading is "Modifier Item Settings". The dialog contains several text fields and a checkbox. The text fields are: "The modifier's path will be:" (value: \), "*Modifier Name:" (value: Heavy), "Tooltip:" (value: Heavy), "*Modifier code:" (value: hvy), "*Symbol Name:" (value: Heavy), "Sourcesystem code:" (value: lcp5_manualentry), and "Comment:". The checkbox is labeled "Apply to all descendents of this concept" and is checked. A note at the bottom states "* denotes required field." The bottom of the dialog has a help icon, "< Back", "Next >", "Finish", and "Cancel" buttons.

Updated 'Smoker' concept (with modifiers)

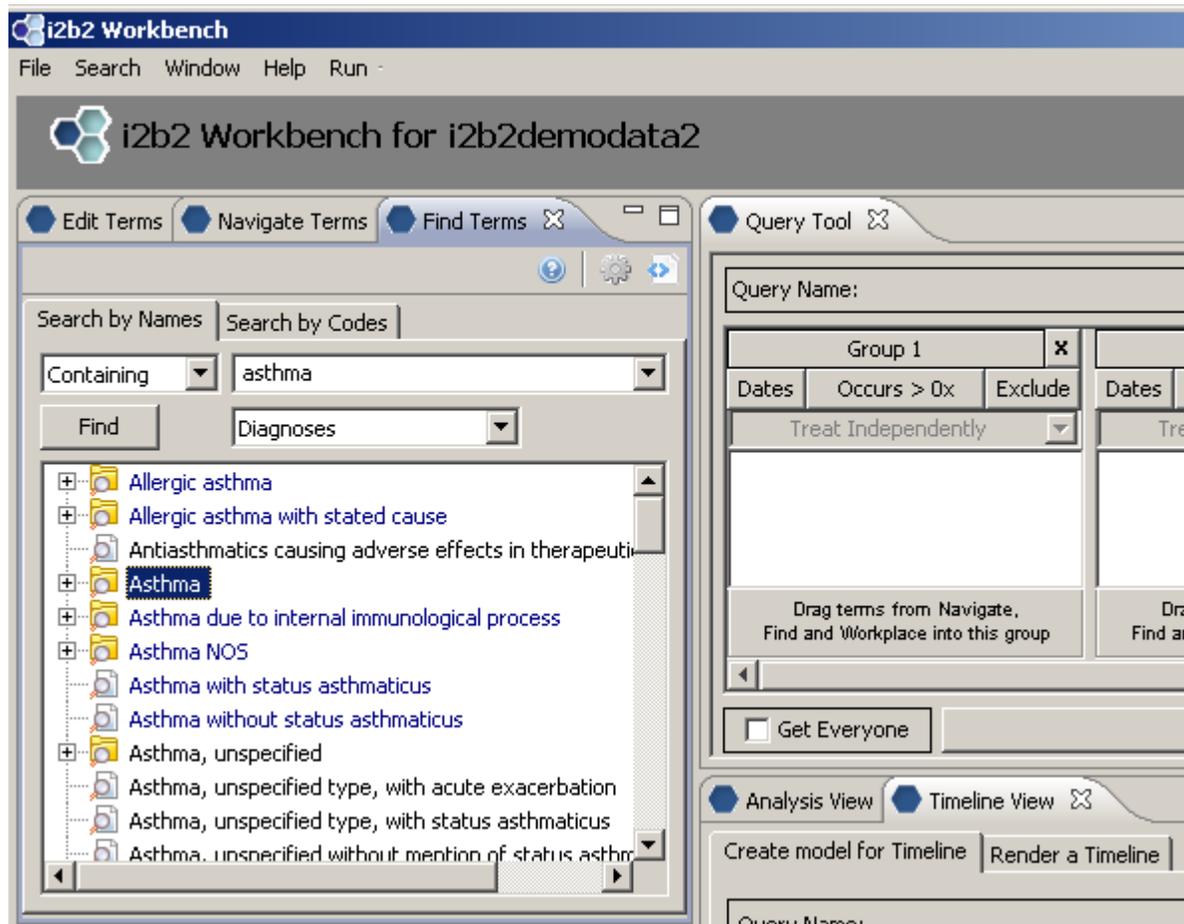


Synchronize metadata with dimension tables

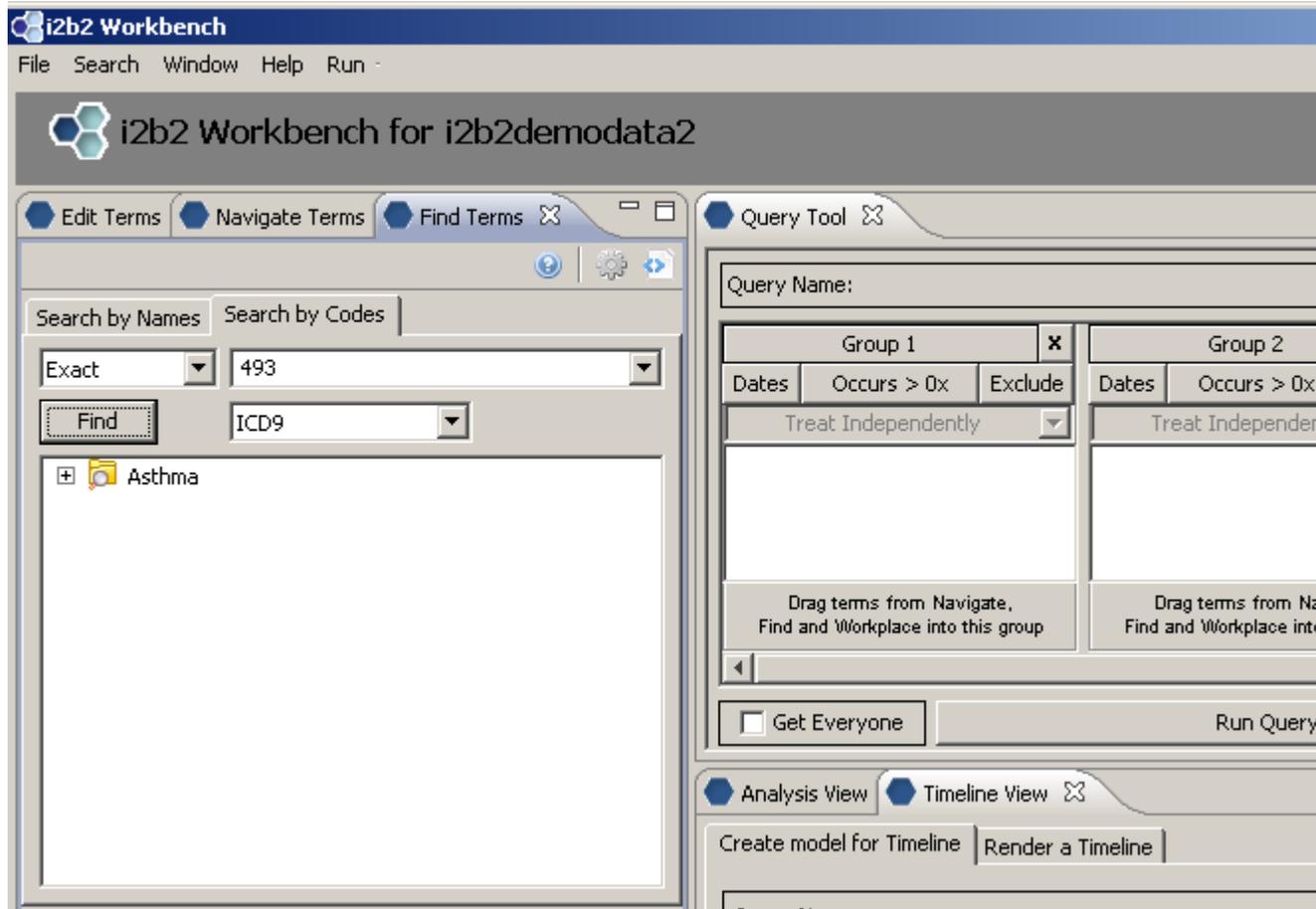
- Concept_dimension
- Modifier_dimension
- Provider_dimension



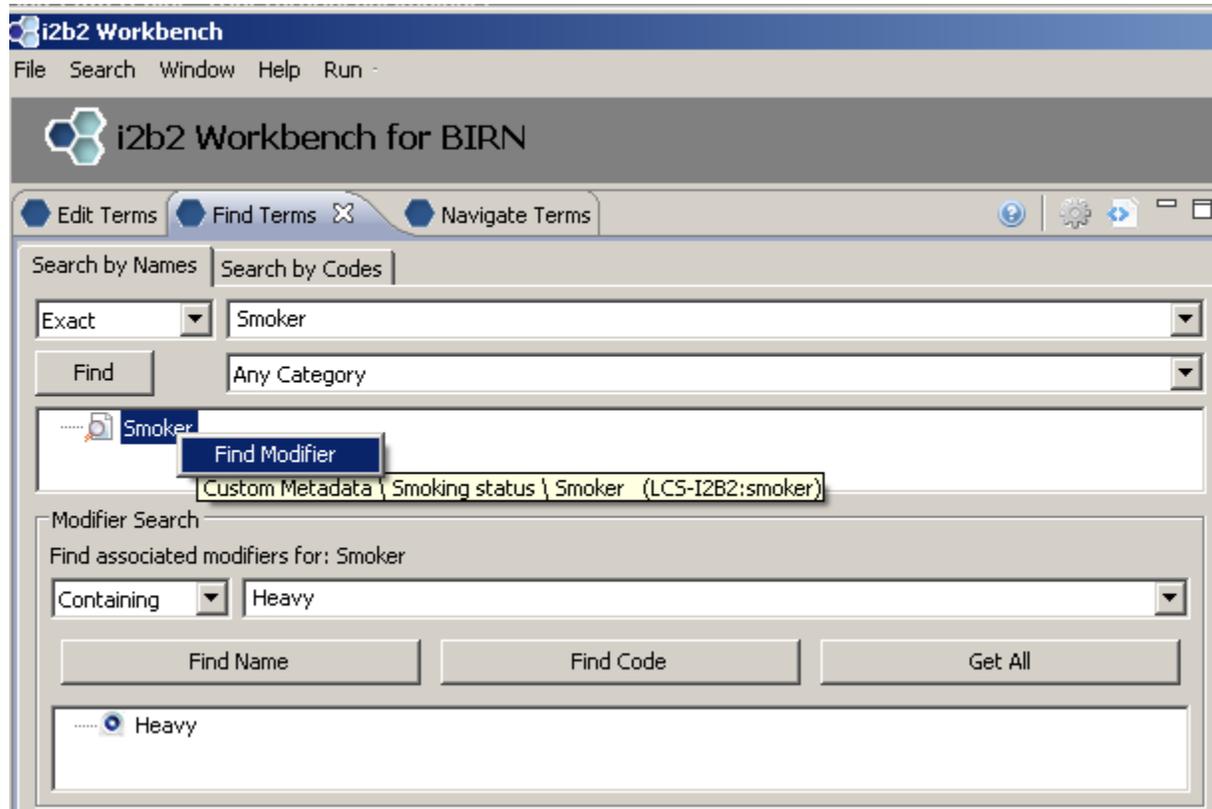
Using Find Terms Search by Names Tab



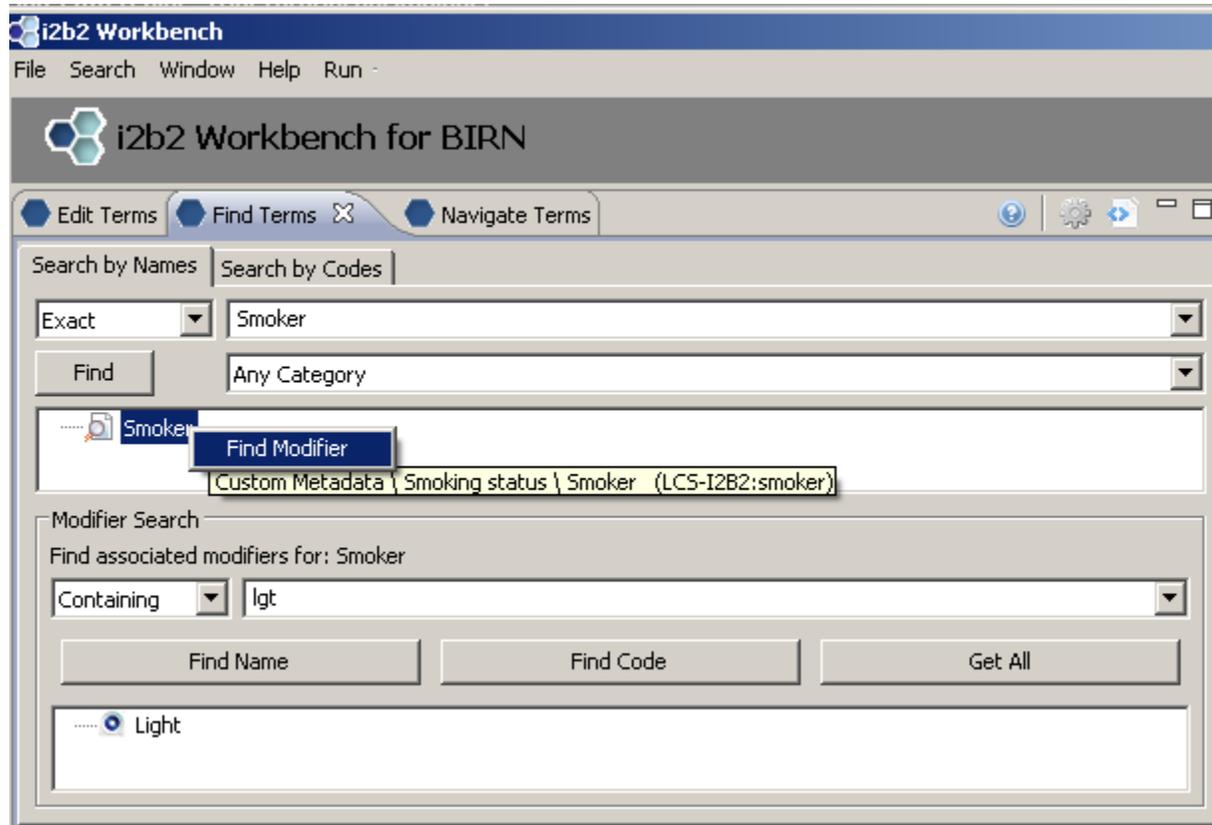
Using Find Terms Search by Codes Tab



Using Find Terms to Locate Modifier by Name

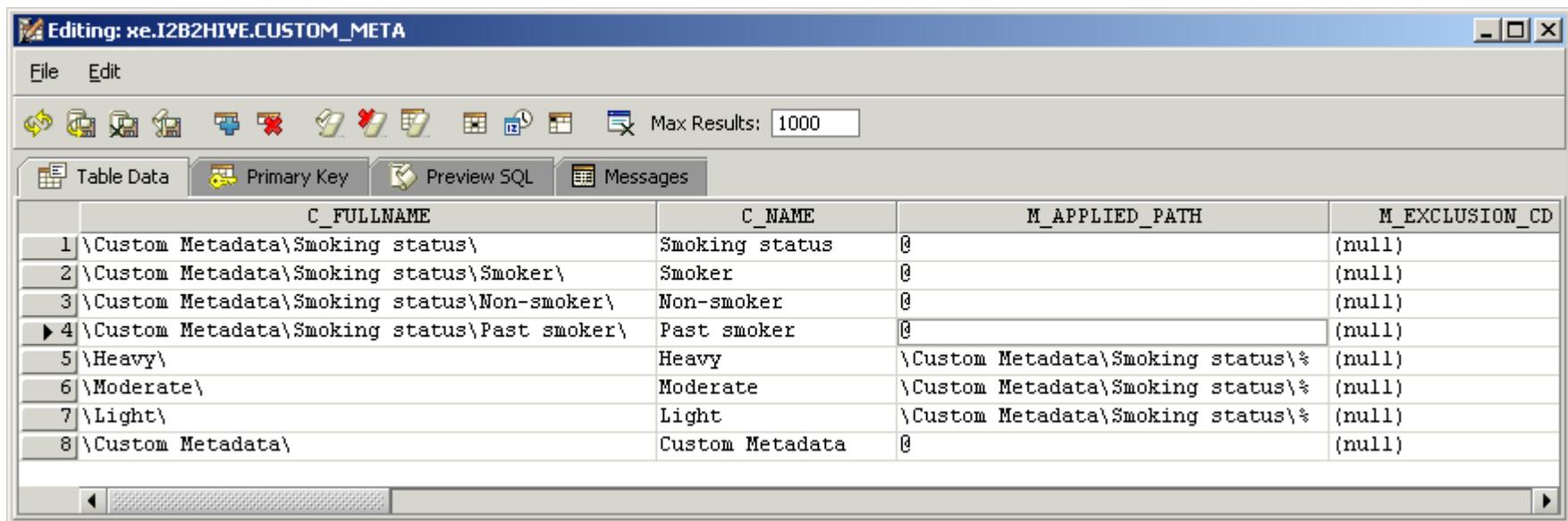


Using Find Terms to Locate Modifier by Code



Metadata db changes to support modifiers

- Modifiers are stored in same metadata table as the terms they apply to.
- Two new columns:
 - m_applied_path
 - m_exclusion_cd

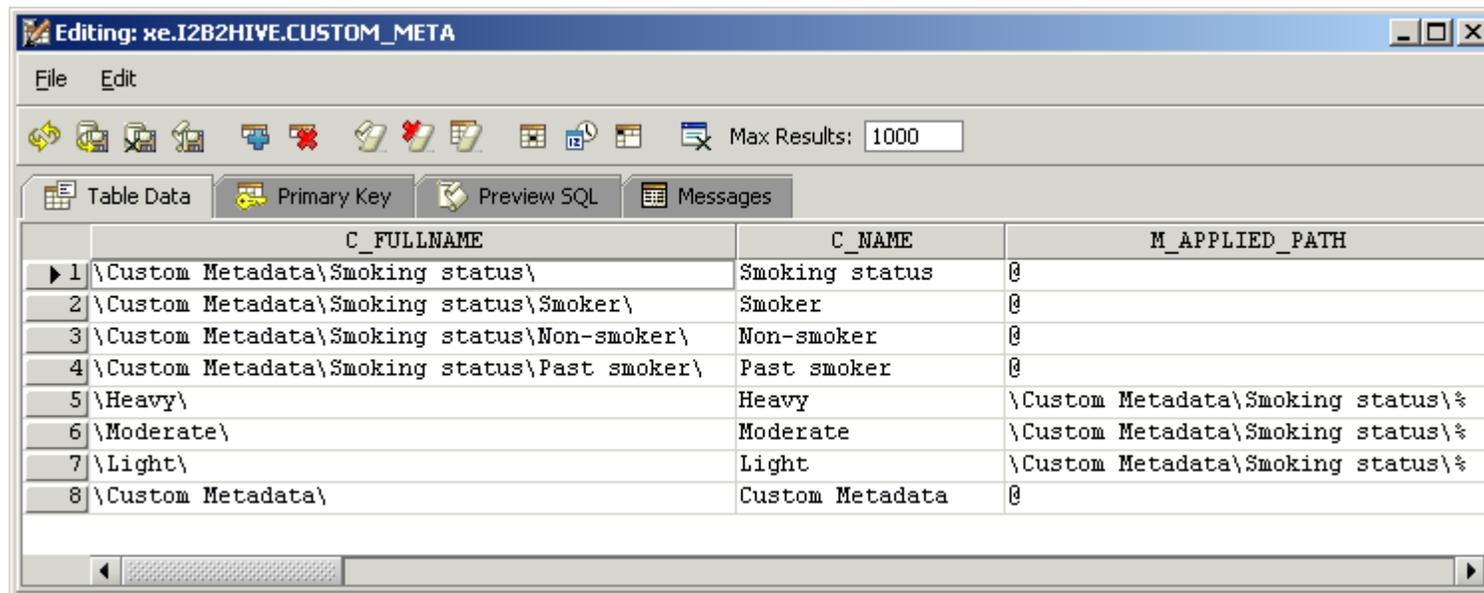


The screenshot shows a window titled "Editing: xe.I2B2HIVE.CUSTOM_META" with a menu bar (File, Edit) and a toolbar. Below the toolbar are tabs for "Table Data", "Primary Key", "Preview SQL", and "Messages". The "Table Data" tab is active, displaying a table with the following data:

	C_FULLNAME	C_NAME	M_APPLIED_PATH	M_EXCLUSION_CD
1	\Custom Metadata\Smoking status\<	Smoking status	@	(null)
2	\Custom Metadata\Smoking status\Smoker\<	Smoker	@	(null)
3	\Custom Metadata\Smoking status\Non-smoker\<	Non-smoker	@	(null)
4	\Custom Metadata\Smoking status\Past smoker\<	Past smoker	@	(null)
5	\Heavy\<	Heavy	\Custom Metadata\Smoking status\%	(null)
6	\Moderate\<	Moderate	\Custom Metadata\Smoking status\%	(null)
7	\Light\<	Light	\Custom Metadata\Smoking status\%	(null)
8	\Custom Metadata\<	Custom Metadata	@	(null)

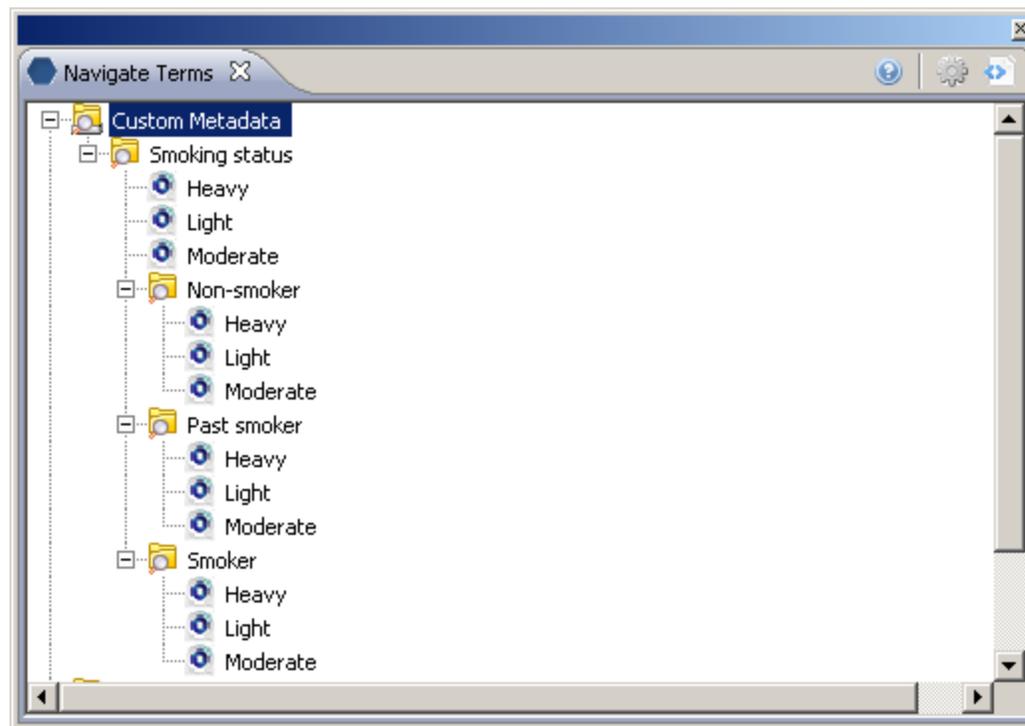
m_applied_path

- Specifies the concept_path of the term the modifier applies to:
 - \Custom Metadata\Smoking status\Smoker\
- Can also be specified to apply to children of the path:
 - \Custom Metadata\Smoking status\%
- Set to '@' for non-modifier (term) metadata entries

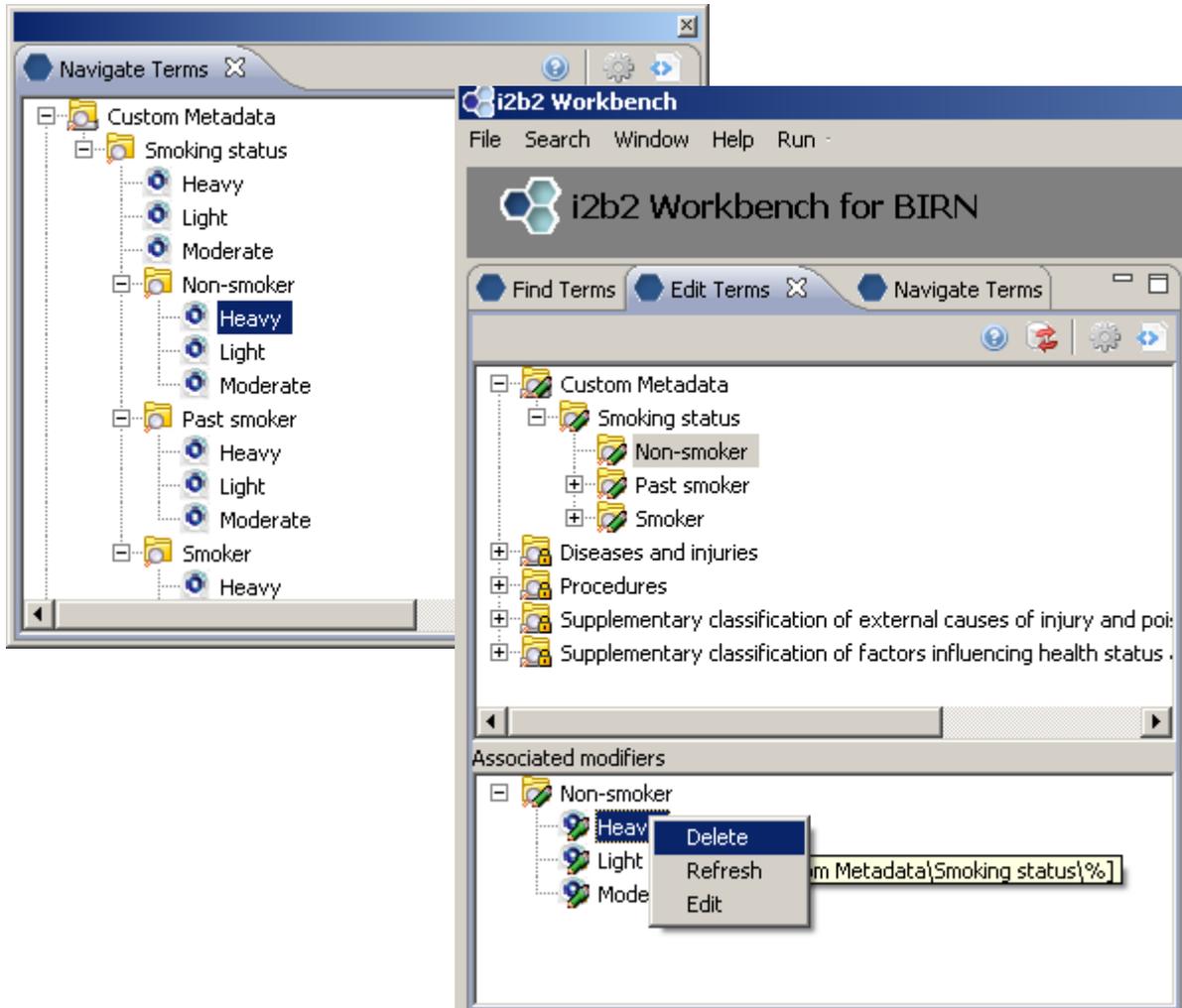


	C_FULLNAME	C_NAME	M_APPLIED_PATH
▶ 1	\Custom Metadata\Smoking status\	Smoking status	@
2	\Custom Metadata\Smoking status\Smoker\	Smoker	@
3	\Custom Metadata\Smoking status\Non-smoker\	Non-smoker	@
4	\Custom Metadata\Smoking status\Past smoker\	Past smoker	@
5	\Heavy\	Heavy	\Custom Metadata\Smoking status\%
6	\Moderate\	Moderate	\Custom Metadata\Smoking status\%
7	\Light\	Light	\Custom Metadata\Smoking status\%
8	\Custom Metadata\	Custom Metadata	@

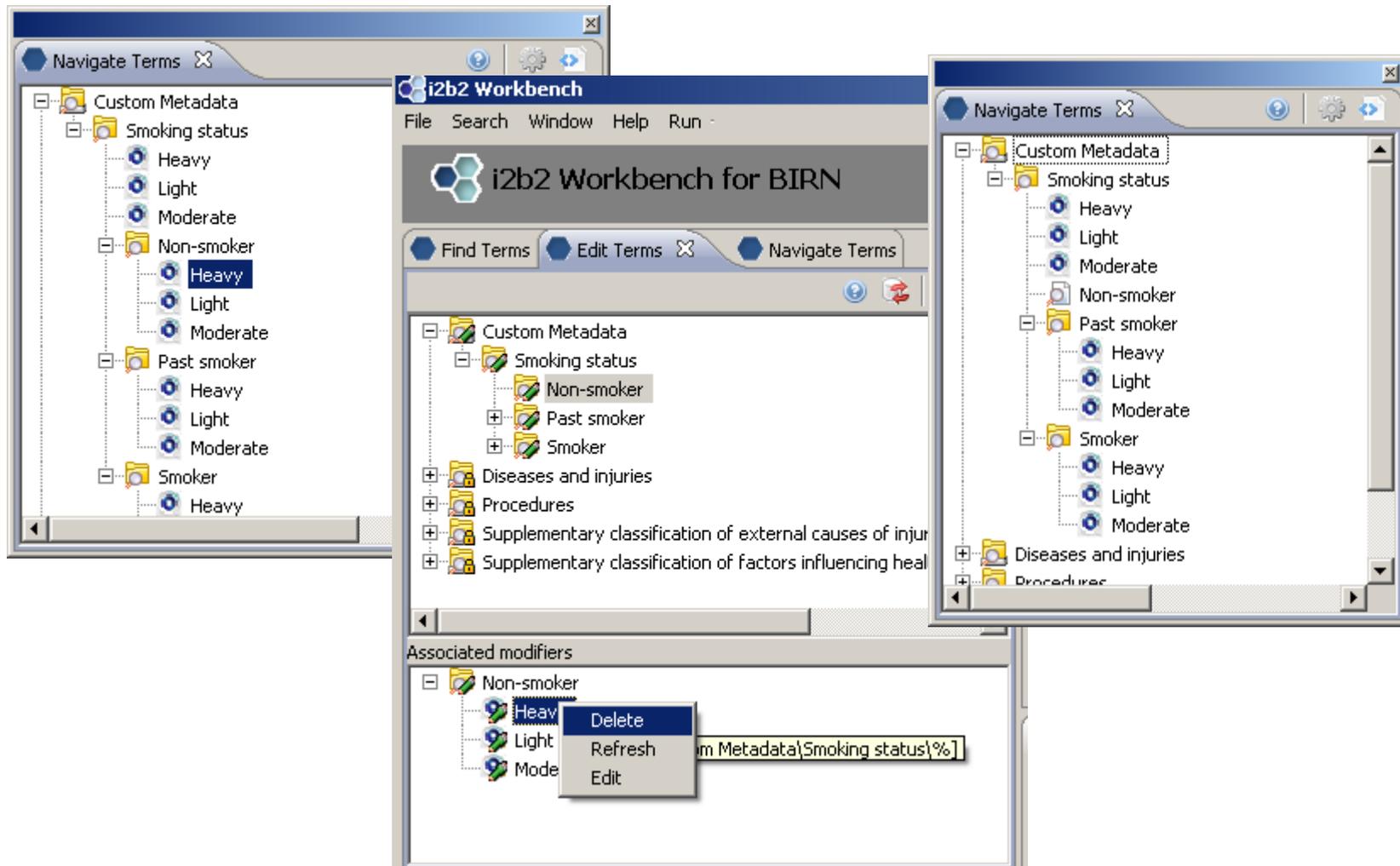
Excluding a modifier



Excluding a modifier

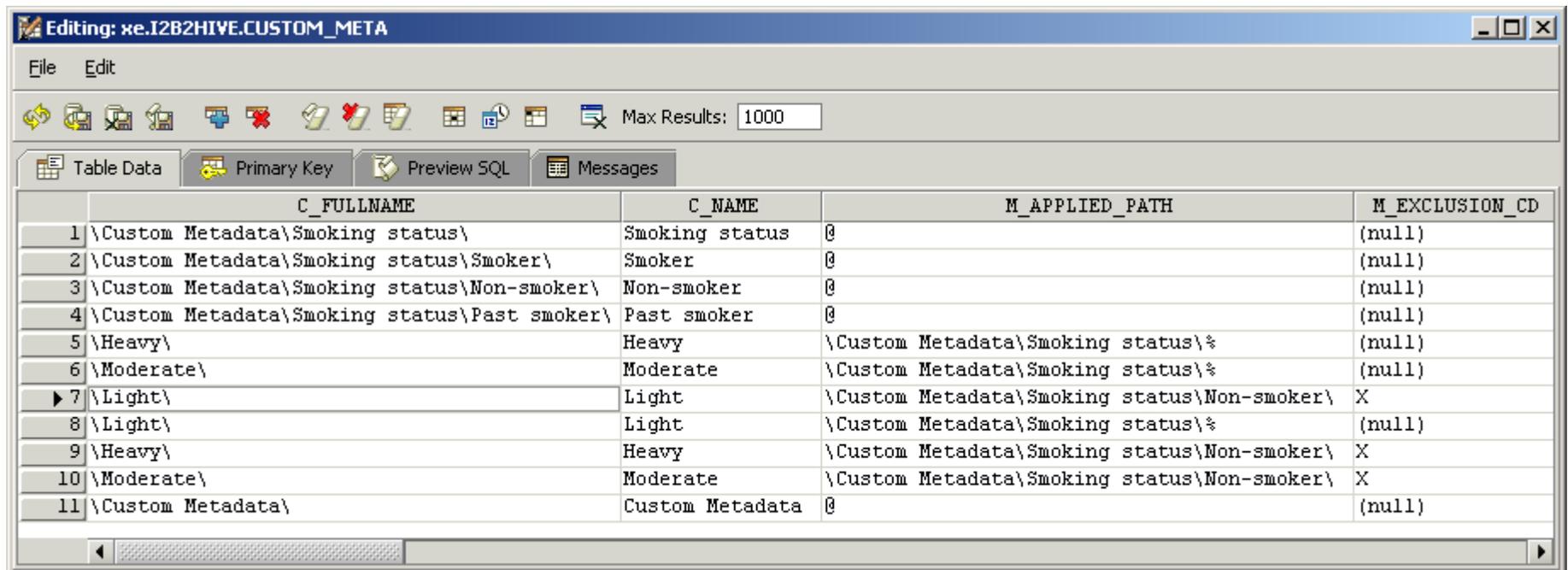


Excluding a modifier



m_exclusion_cd

- If non-null ('X') the modifier is excluded from the concept_path specified in m_applied_path



Editing: xe.I2B2HIVE.CUSTOM_META

File Edit

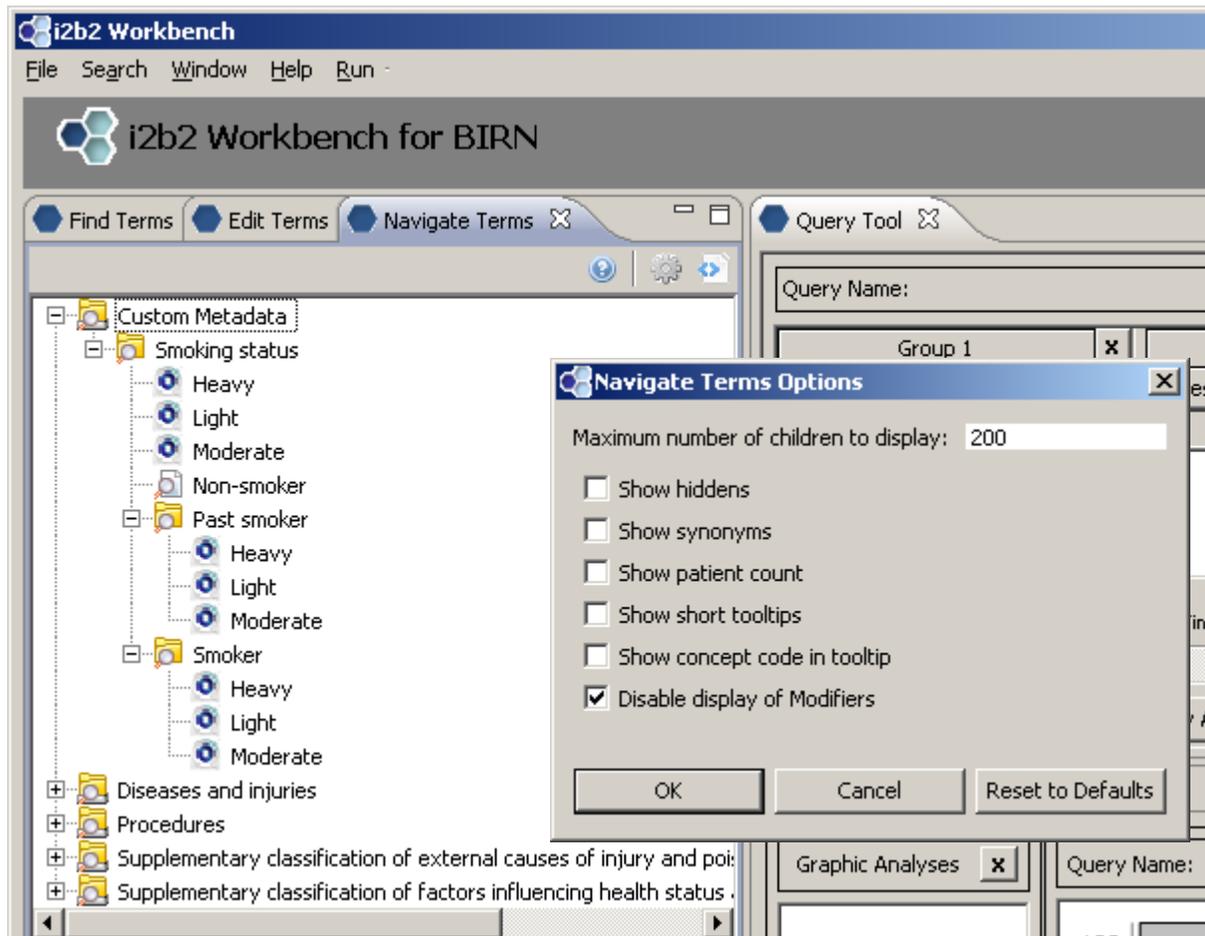
Max Results: 1000

Table Data Primary Key Preview SQL Messages

	C_FULLNAME	C_NAME	M_APPLIED_PATH	M_EXCLUSION_CD
1	\Custom Metadata\Smoking status\	Smoking status	@	(null)
2	\Custom Metadata\Smoking status\Smoker\	Smoker	@	(null)
3	\Custom Metadata\Smoking status\Non-smoker\	Non-smoker	@	(null)
4	\Custom Metadata\Smoking status\Past smoker\	Past smoker	@	(null)
5	\Heavy\	Heavy	\Custom Metadata\Smoking status\%	(null)
6	\Moderate\	Moderate	\Custom Metadata\Smoking status\%	(null)
7	\Light\	Light	\Custom Metadata\Smoking status\Non-smoker\	X
8	\Light\	Light	\Custom Metadata\Smoking status\%	(null)
9	\Heavy\	Heavy	\Custom Metadata\Smoking status\Non-smoker\	X
10	\Moderate\	Moderate	\Custom Metadata\Smoking status\Non-smoker\	X
11	\Custom Metadata\	Custom Metadata	@	(null)

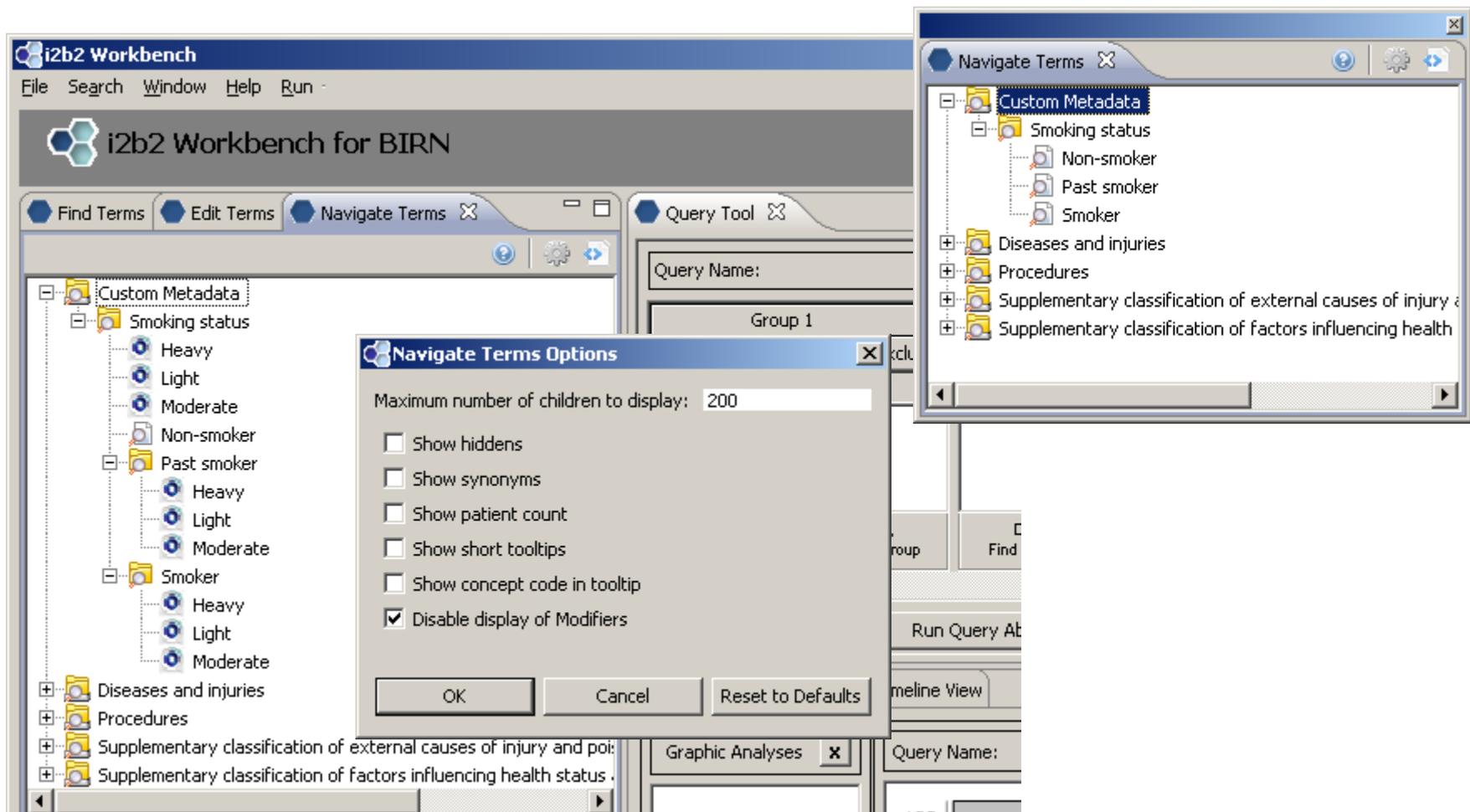
Disabling modifiers

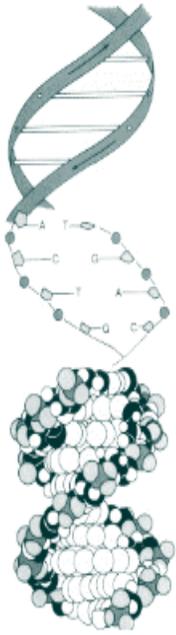
- Mechanism to disable the display and acquisition of modifiers.



Disabling modifiers

- Refresh navigate tree





Ontology Query Building

Using modifiers in a query

The screenshot displays the i2b2 Workbench for BIRN interface. On the left, a tree view under 'Custom Metadata' shows 'Smoking status' expanded to 'Smoker' > 'Light', which is highlighted. The main 'Query Tool' window is divided into three groups:

Group 1			Group 2			Group 3		
Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude	Dates	Occurs > 0x	Exclude
Treat Independently			Treat Independently			Treat Independently		
Smoker [Light]								
The terms of this group are joined then intersected with other groups			Drag terms from Navigate, Find and Workplace into this group			Drag terms from N Find and Workplace in		

At the bottom of the Query Tool, there is a checkbox for 'Get Everyone' and a 'Run Query Above' button.

Dragging a modifier to Query tool

- When you drag a modifier, you are really dragging a concept term / modifier pair.

```
<concept>
  <level>2</level>
  <key>\\CUST\Custom Metadata\Smoking status\Smoker</key>
  <name>Smoker</name>
  <synonym_cd>N</synonym_cd>
  <visualattributes>FAE</visualattributes>
  <basecode>LCS-I2B2:smoker</basecode>
  <tooltip>Custom Metadata \ Smoking status \ Smoker</tooltip>
  <modifier>
    <level>1</level>
    <applied_path>\Custom Metadata\Smoking status\%</applied_path>
    <key>\\CUST\Light</key>
    <fullname>\Light</fullname>
    <name>Light</name>
    <visualattributes>RAE</visualattributes>
    <synonym_cd>N</synonym_cd>
    <basecode>lgt</basecode>
    <tooltip>Light [\Custom Metadata\Smoking status\%]</tooltip>
  </modifier>
</concept>
```

Putting it all together to form a query

```
select patient_num from observation_fact
  where concept_cd IN (select concept_cd
    from concept_dimension where concept_path
    like '\Custom Metadata\Smoking status\Smoker\%')
```

```
and modifier_cd IN (select modifier_cd from
  modifier_dimension where modifier_path
  like '\Light\%')
```

Edit Term Wizard

Query Dimension Settings

These settings are inherited from the parent node.

*Table Name:

*Column Name:

*Fact Table Column Name:

*Operator:

*Column Data Type:

* denotes required field.

Edit Modifier Wizard

Query Dimension Settings

These settings are inherited from the parent node.

*Table Name:

*Column Name:

*Fact Table Column Name:

*Operator:

*Column Data Type:

* denotes required field.

Extending queries beyond concept_dimension

■ Provider_dimension

- select patient_num from observation_fact where provider_id in
(select provider_id from provider_dimension where provider_path
like '\i2b2\Providers\Emergency\%')

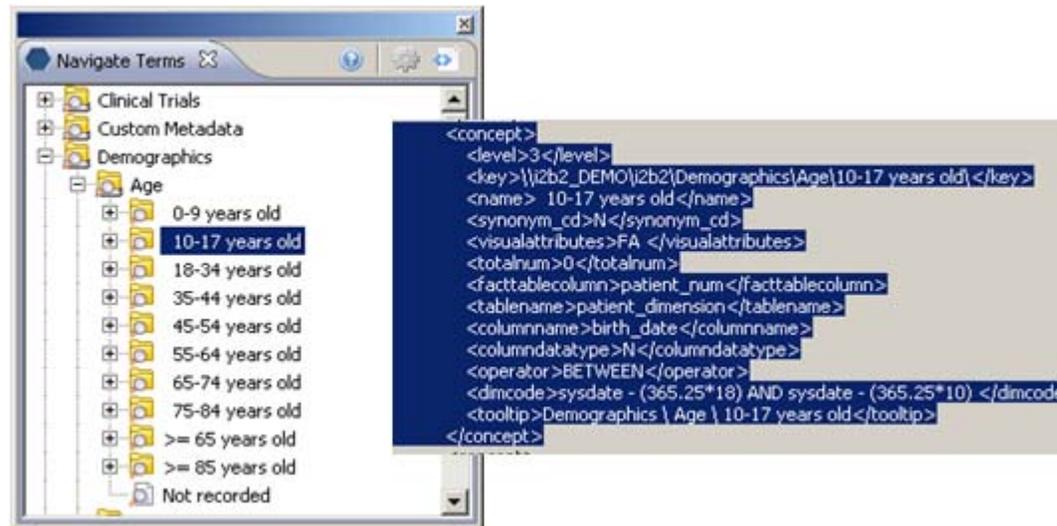
■ Patient_dimension

- select patient_num from observation_fact where patient_num in
(select patient_num from patient_dimension where race_cd = 'white')

■ Visit_dimension

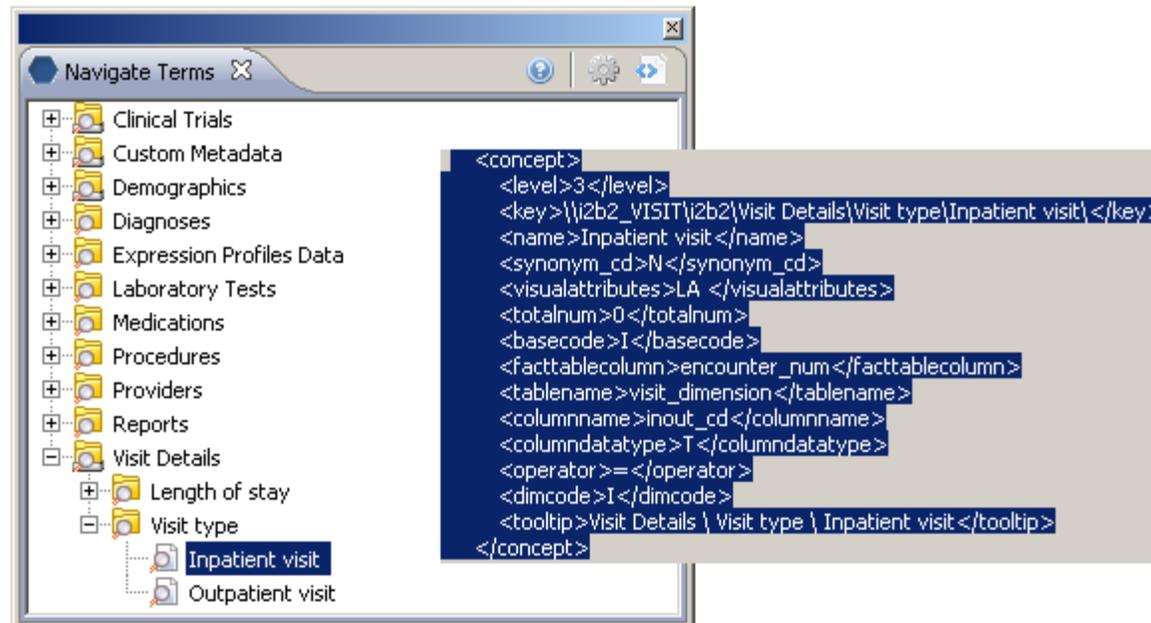
- select patient_num from observation_fact where encounter_num in
(select encounter_num from visit_dimension where length_of_stay = 2)

Patient_dimension based queries



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

Visit_dimension based queries



```
<concept>  
<level>3</level>  
<key>\\i2b2_VISIT\i2b2\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>
```

select patient_num from observation_fact where encounter_num in
(select encounter_num from visit_dimension where inout_cd = 'I')

Where are we going from here?

- Stand alone ontology creation tool that uses NCBO services
 - Likely to be Protégé plug-in
- Advocate for uniform representation of SNP's
- Develop Ontology Mapping Cell
- Community guideline development on Wiki for
 - Most useful ontologies and versions of ontologies
 - Use of modifiers
 - Value representations