



**caBIG**<sup>™</sup> cancer Biomedical  
Informatics Grid <sup>™</sup>

# Open Terminology Portal (TOP)

***Frank Hartel, Ph.D.***

***Associate Director, Enterprise Vocabulary Services***

***National Cancer Institute, Center for Biomedical Informatics***

***NCBO Scientific Council***

***October 16, 2007***

# NCI BioPortal and the TOP Project Links



- **NCI BioPortal Project**

<https://gforge.nci.nih.gov/projects/lex-browser/>

- **Open Terminology Portal Project**

<https://gforge.nci.nih.gov/projects/openportal/>

# Why did NCI find BioPortal so interesting?



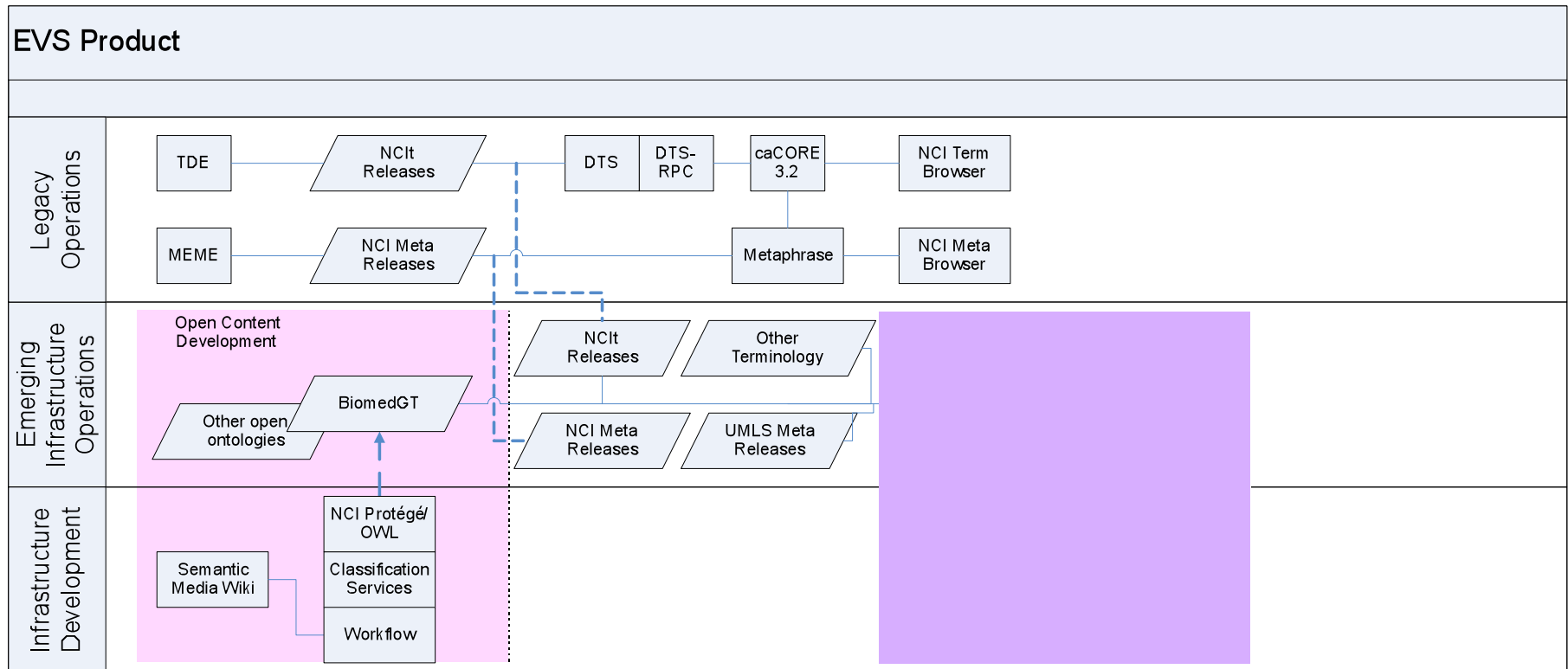
- **Our caCORE and caGrid infrastructure relies on terminology servers to provide access to the base semantics used to construct metadata and data model semantics**
- **During 2007 we have moved away from our legacy terminology server architecture to an open architecture built around LexGrid (LexBIG)**
- **Our legacy terminology browsers are not compatible with the LexBIG infrastructure – but BioPortal largely was and we needed a interactive Web tool for our users**
- **BioPortal had a simple, effective GUI**
- **While NCBO had not built it for distribution, they were willing to let us have the code**

# Why did NCI not simply use NCBO BioPortal?

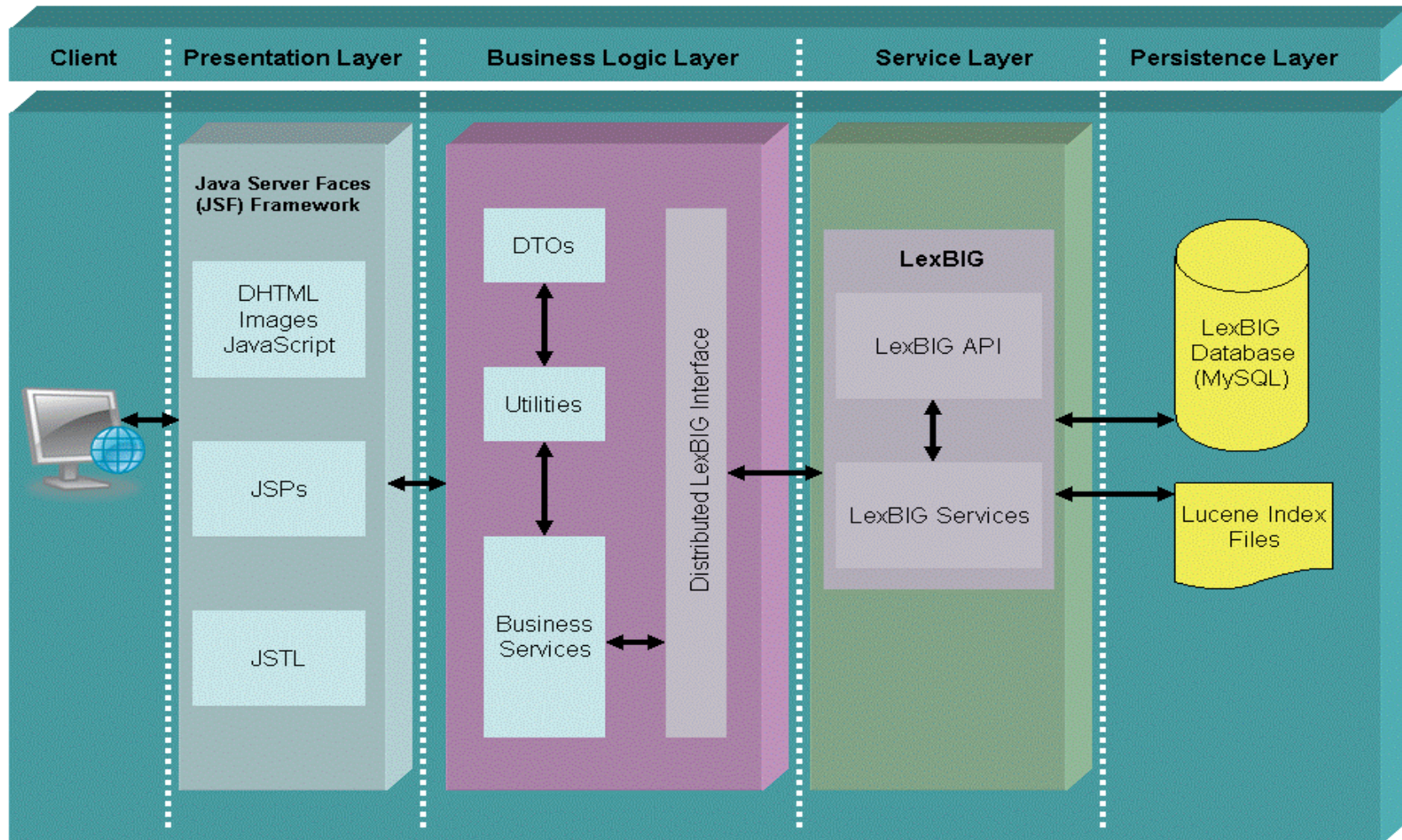


- **We needed to serve a Metathesaurus (NCI Metathesaurus – similar to UMLS Metathesaurus), not only individual ontologies**
- **Our operations model precludes uploading of ontologies by end users**
- **Some of the terminologies we serve have license restrictions prohibiting redistribution or requiring license validation before downloading**
- **Our infrastructure is not the same as the one that NCBO BioPortal assumed**
  - Database is MySQL
  - Index files, database and LexBIG not co-located with BioPortal
  - NCI BioPortal bound to caCORE/LexBIG product and its release cycle
    - Planning, engineering methodology and artifacts, documentation QA, training
    - We use Distributed LexBIG API, not local LexBIO API

# Context of Open Development – NCI Enterprise Vocabulary Services



# NCI BioPortal Architecture Layers



# What did we do with BioPortal to date?

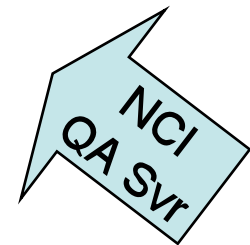


- **We tried to make BioPortal less NCBO specific**
- **Added MySQLSequenceHandler and other code to support MySQL**
- **Removed Oracle specific dependencies**
- **Removed NCBO database and relevant code segments**
- **Removed Apache HTTP server & Tomcat Connector**
- **Integrated JBoss with AT&T Graphviz**
  - (All Graphviz output files including .dot, .jpg, and .map) are now generated in a JBoss tmp subdirectory instead.)
- **Modified Windows-specific code; deployable on Linux**
- **Integrated with distributed LexBIG**

# What did we do with BioPortal to date? - ctd



- **Metadata files loaded to MySQL via LexBIG metadata loader**
- **Removed code specific NCBO workflow (user registration, ontology submission, admin functions)**
- **Created Quick and Advanced Search Tabs**
  - Search by source, class property, association, etc.
- **Removed EJB3 related to NCBO database**
- **GUI changes -- Examples**
  - Library statistics read from MYSQL
  - Metathesaurus browse and display
  - On line help, user guide
- **<http://cbvapp-q1003.nci.nih.gov:19480/ncbo/faces/index.xhtml>**



# NCI BioPortal Library Statistics



National Cancer Institute

U.S. National Institutes of Health | www.cancer.gov



Version 1.0

[Home](#) [Help](#)

HELP

[USER GUIDE \(PDF\)](#)  
[ONLINE HELP](#)

QUICK LINKS

- [EVS HOME](#)
- [NCICB HOME](#)
- [NCI HOME](#)
- [NIH HOME](#)
- [HHS HOME](#)
- [FIRSTGOV HOME](#)

TOOLS

[NEW TERM REQUEST](#)

[Browse](#) [Search](#)

## Ontologies

[Download](#) [Visualize](#) [Search](#)

Name	Format	Current Version	Content Location	Action
<a href="#">test2</a>		1.1		<a href="#">Download</a> <a href="#">Visualize</a> <a href="#">Search</a>
<a href="#">GeneOntology.owl</a>		??		<a href="#">Download</a> <a href="#">Visualize</a> <a href="#">Search</a>
<a href="#">HL7.owl</a>		V3		<a href="#">Download</a> <a href="#">Visualize</a> <a href="#">Search</a>
<a href="#">autos</a>		1.0		<a href="#">Download</a> <a href="#">Visualize</a> <a href="#">Search</a>
<a href="#">The MGED Ontology</a>		classified		<a href="#">Download</a> <a href="#">Visualize</a> <a href="#">Search</a>
<a href="#">National Drug File - Reference Terminology, NCI_072606</a>		NCI_072606		<a href="#">Download</a> <a href="#">Visualize</a> <a href="#">Search</a>
<a href="#">NCI MetaThesaurus</a>	RRF	200706	<a href="http://ncicb.nci.nih.gov/core/EVS">http://ncicb.nci.nih.gov/core/EVS</a>	<a href="#">Download</a> <a href="#">Visualize</a> <a href="#">Search</a>
<a href="#">UMLS Semantic Network</a>		3.2		<a href="#">Download</a> <a href="#">Visualize</a> <a href="#">Search</a>
<a href="#">SNOMED Clinical Terms, 2006_07_31</a>		2006_07_31		<a href="#">Download</a> <a href="#">Visualize</a> <a href="#">Search</a>
<a href="#">NCI Thesaurus</a>	OWL-FULL	07.08d	<a href="http://ncicb.nci.nih.gov/core/EVS">http://ncicb.nci.nih.gov/core/EVS</a>	<a href="#">Download</a> <a href="#">Visualize</a> <a href="#">Search</a>
<a href="#">GMP</a>		2.0		<a href="#">Download</a> <a href="#">Visualize</a> <a href="#">Search</a>

[CONTACT US](#)

[PRIVACY NOTICE](#)

[DISCLAIMER](#)

[ACCESSIBILITY](#)

[APPLICATION SUPPORT](#)



FIRSTGOV

# Browse in individual ontologies as on NCBO BioPortal



National Cancer Institute U.S. National Institutes of Health | www.cancer.gov

**EVS** Enterprise Vocabulary Services **BioPortal** Version 1.0

HELP USER GUIDE (PDF) ONLINE HELP Home Help

QUICK LINKS EVS HOME NCICB HOME NCI HOME NIH HOME HHS HOME FIRSTGOV HOME

TOOLS NEW TERM REQUEST

Browse Search

### Medical Dictionary for Regulatory Activities Terminology (MedDRA), 10.0

**Tree View**  
Tree view constructed based on *hasSubclass* hierarchy

- ⊕ Blood and lymphatic system disorders
- ⊖ Cardiac disorders
  - ⊕ Cardiac arrhythmias
  - ⊕ Cardiac disorder signs and symptoms
  - ⊕ Cardiac neoplasms
  - ⊕ Cardiac valve disorders
  - ⊕ Congenital cardiac disorders
  - ⊕ Coronary artery disorders
  - ⊕ Endocardial disorders
  - ⊕ Heart failures
  - ⊕ Myocardial disorders
  - ⊕ Pericardial disorders
- ⊕ Congenital, familial and genetic disorders
- ⊕ Ear and labyrinth disorders
- ⊕ Endocrine disorders

**Class/Type Details**  
No class selected.

**Graph View**  
Graph Type: Local Neighborhood  
No class selected.

# Browse in Meta requires selection of a source



National Cancer Institute

U.S. National Institutes of Health | www.cancer.gov

**EVS** Enterprise Vocabulary Services **BioPortal**

Version 1.0

[Home](#) [Help](#)

HELP

[USER GUIDE \(PDF\)](#)  
[ONLINE HELP](#)

QUICK LINKS

[EVS HOME](#)

[NCICB HOME](#)

[NCI HOME](#)

[NIH HOME](#)

[HHS HOME](#)

[FIRSTGOV HOME](#)

TOOLS

[NEW TERM REQUEST](#)

[Browse](#) [Search](#)

[← Search Results](#)

**Tree View**

Please select a source from the list.

Sources   
[Continue](#)

- SCTSPA
- SNOMEDCT

**Class/Type Details**

**General**

Class/Type Name **Dental investigation observations**  
Id **C0429136**

**Attributes**

Full Form	<b>Dental investigation observations (finding), observaciones de la evaluación dental (hallazgo)</b>
MTH_PT	<b>observaciones de la evaluación dental</b>
MTH_FN	<b>observaciones de la evaluación dental (hallazgo)</b>
Semantic_Type	<b>Finding</b>

**Graph View**

Graph Type

# In Meta, graph views also require source selection



## NCI MetaThesaurus

[+ Search Results](#)

### Tree View

Tree view constructed based on *hasSubclass* hierarchy

Root set to **Dental investigation observations**

[View Full Ontology](#)

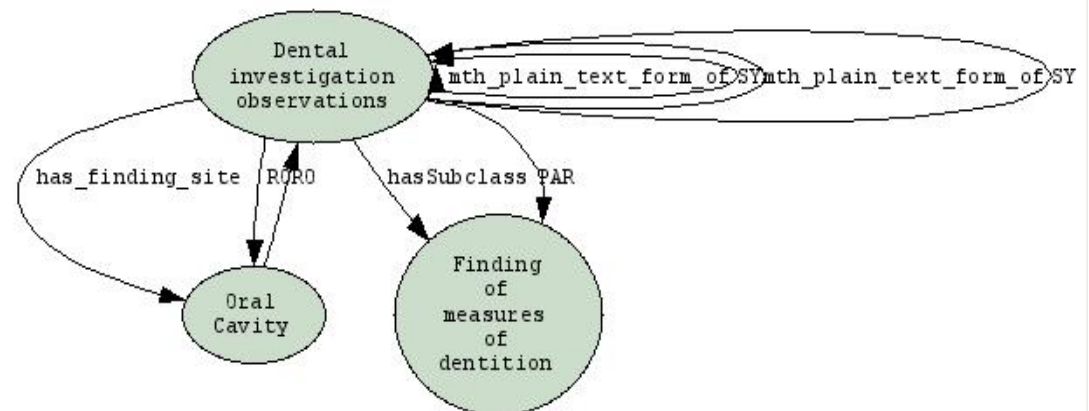
[\[-\] Dental investigation observations](#)

### Class/Type Details

General		Attributes	
Class/Type Name	<b>Dental investigation observations</b>	Full Form	<b>Dental investigation observations (find observaciones de la evaluación dental observaciones de la evaluación dental observaciones de la evaluación dental</b>
Id	C0429136	MTH_PT	<b>observaciones de la evaluación dental</b>
		MTH_FN	<b>observaciones de la evaluación dental</b>
		Semantic_Type	<b>Finding</b>

### Graph View

Graph Type



# Genesis of TOP



- **NCI changes to BioPortal**
  - Made BioPortal code easier to host locally, but...
  - Forked code – need to reintegrate !
  - Demonstrated interest in local instances of BioPortal to meet local needs
- **Others with interest in BioPortal emerged – Mayo Clinic and UK Cancer Grid beside NCI and NCBO**
- **Formed initial partners in Open Terminology Portal Project**
  - Open development, open source, participant driven governance still under development
  - Additional partners are desired

# Current TOP Activities



- **Early days...**
- **Charter adopted**
- **Discussing open development process and project management**
- **Scope definition for near term “de-forking”**
  - Initial goal is to continue to refine the BioPortal code and close the fork
    - Requirements and design goals, architecture under development
    - Close the fork, robust, enterprise scale, site neutral, etc.
- **Requirements gathering for future TOP offerings, initiatives**
  - A bit further in the future, development of services, components and other resources to address patterns of terminology use are being discussed
  - Target architecture
    - Architectural considerations discussions beginning