

Developing Effective Quality Systems in Clinical Trials: An Enlightened Approach

NCI Data Quality Initiatives & Enterprise Vocabulary Services

October 13, 2010

Edward Helton, NCI Imaging Program

Larry Wright, NCI Enterprise Vocabulary Services

Background (1)

Late 1990s cancer research enterprise data quality & tracking

- **Trials diversity in running, reporting and analysis**
 - Treatment, prevention, screening, supportive care, etc.
 - Many sponsors, sites, systems
- **Related research activities even more diverse**
 - Basic research, clinical care, cancer registries, etc.
 - Part of broader communities: federal, health care, etc.
- **Dr. Klausner wanted**
 - A broad, consistent understanding of research activity
 - Sharper, better integrated research effort

Background (2)

EVS born in 1997 as part of this effort, key thrusts:

1. **Analyze and organize** diverse coding
 - LexEVS and other terminology services
2. **Map meanings** between different terminologies
 - NCI Metathesaurus (NCIm)
3. **Reference terminology** with high quality meaning shared between people and computers
 - NCI Thesaurus (NCIt)

EVS provides the **semantic foundation** for sharing and re-use of data and services – **interoperability**.

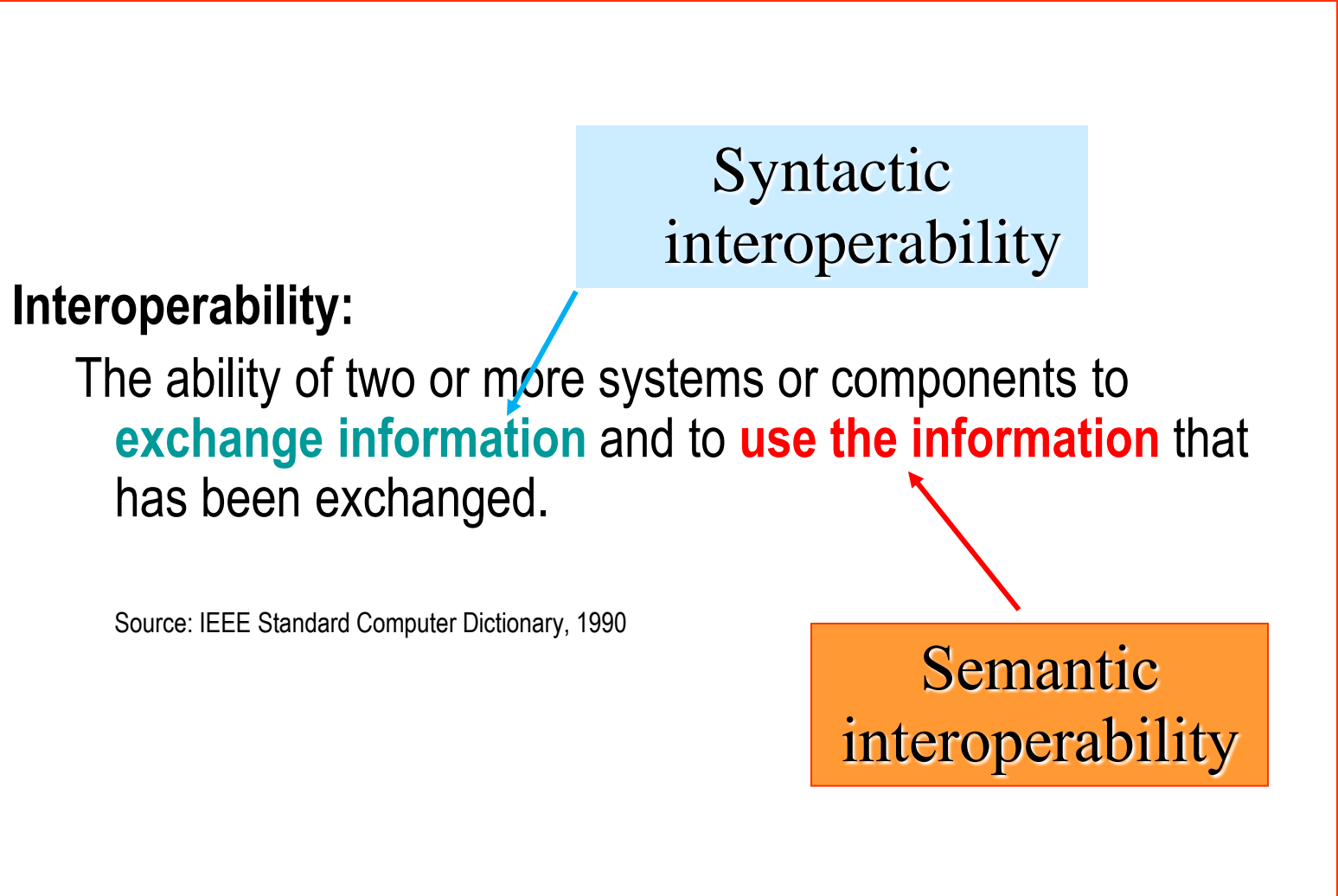
Interoperability

Interoperability:

The ability of two or more systems or components to **exchange information** and to **use the information** that has been exchanged.

Source: IEEE Standard Computer Dictionary, 1990

Syntactic
interoperability



Semantic
interoperability

NCI Design for Interoperability

- **Common API Integration:** Part of the syntactic component of interoperability.
- **Vocabularies/Terminologies/Ontologies:** Provides semantic interoperability, used to record information in and about systems and data.
- **Data Elements:** Metadata provides a description of the meaning of information in addition to its value. For example, *Performance Status* would describe both a meaning and what constitutes a valid value.
- **Information Models:** Describe the information design and structure of the data maintained in a system.

No Controlled Terminology? No Interoperability

- **Systems cannot exchange or use information if they use incompatible codes or tokens to signify meaning.**
- **Terminology services provide those tokens, codes and compatibility.**
- **Proper use of terminology assures consistent meaning across and among enterprises.**
- **Requires content and technical quality, with a big dose of harmonization and consensus building.**

Unambiguous Meaning

Semantic Type: Quantitative Concept**Code: C42539**

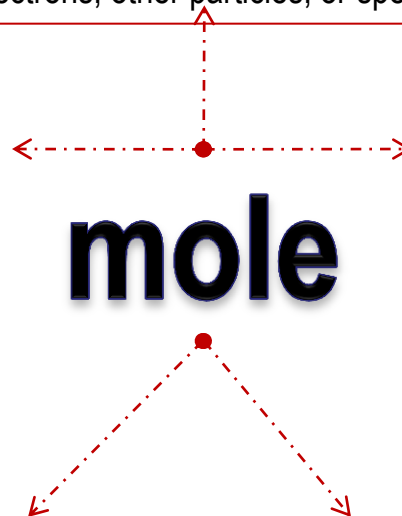
Definition: A unit of amount of substance, one of the seven base units of the International System of Units (Système International d'Unités, SI). It is the amount of substance that contains as many elementary units as there are atoms in 0.012 kg of carbon-12. When the mole is used, the elementary entities must be specified and may be atoms, molecules, ions, electrons, other particles, or specified groups of such particles.

Semantic Type: Mammal**Code: C14876**

Definition: A small, furry creature of the family Talpidae that lives underground and feeds on small invertebrates. The mole has tiny covered eyes that are believed to be able to distinguish night from day, and not much else.

Semantic Type: Occupation or Discipline

Definition: [No use case for this term yet, but welcome CIA inquiries].

**Semantic Type: Neoplastic Process****Code: C7570**

Definition: A neoplasm composed of melanocytes that usually appears as a dark spot on the skin.

Semantic Type: Food or Food Product

Definition: [No use case for this term yet, but welcome inquiries accompanied by samples].

NCI Thesaurus (NCIt)

- **Reference terminology/ontology for clinical, research and other NCI and caBIG systems.**
- **Recognized standard for biomedical coding used both nationally and internationally.**
- **Public domain, open content license.**
- **Built by integrated team of subject matter and technology experts using description logics.**
- **Built in active collaboration with wide range of partners and stakeholders.**

- **NCI Internal Outreach & Integration**
 - PDQ Cancer Clinical Trials Registry & NCI Drug Dictionary
 - CTEP therapeutic trials
 - DCP prevention trials
 - Many others
- **External outreach**
 - Other NIH including NLM, NHLBI, NICHD, NIDCR, BTRIS
 - Interagency: FDA, VHA, CDC, other
 - SDOs: CDISC, NCPDP, HL7, WHO
- **caBIG drew in Cancer Centers & others, public & private**
 - Standards and frameworks for interoperability
 - Building applications and resources

NCIt Drug Information

NCI Drug Dictionary



Page Options

-  [Print This Page](#)
-  [Bookmark & Share](#)

More NCI Dictionaries

- [Dictionary of Cancer Terms](#)
- [Glossary of Statistical Terms](#)
- [NCI Dictionary of Genetics Terms](#)
- [Terminology Resources](#)

Questions about cancer?

- [1-800-4-CANCER](#)
- [LiveHelp online chat](#)



Starts with

Contains

<#> [A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#) [All](#)

imatinib mesylate *R* Patient Information >

The mesylate salt of imatinib, a tyrosine kinase inhibitor with antineoplastic activity. Imatinib binds to an intracellular pocket located within tyrosine kinases (TK), thereby inhibiting ATP binding and preventing phosphorylation and the subsequent activation of growth receptors and their downstream signal transduction pathways. This agent inhibits TK encoded by the bcr-abl oncogene as well as receptor TKs encoded by the c-kit and platelet-derived growth factor receptor (PDGFR) oncogenes. Inhibition of the bcr-abl TK results in decreased proliferation and enhanced apoptosis in malignant cells of Philadelphia-positive (Ph+) hematological malignancies such as CML and ALL; effects on c-kit TK activity inhibit mast-cell and cellular proliferation in those diseases overexpressing c-kit, such as mastocytosis and gastrointestinal stromal tumor (GIST). Check for [active clinical trials](#) or [closed clinical trials](#) using this agent. ([NCI Thesaurus](#))

US brand name: Gleevec

Foreign brand name: Glivec

Code names: CGP 57148
CGP57148B
STI-571

Chemical structure name: 4-[(4-methyl-1-piperazinyl)methyl]-N-[4-methyl-3-[[4-(3-pyridinyl)-2-pyrimidinyl]amino]-phenyl]benzamide methanesulfonate

CTCAE

Version: 4.03

- Exact Match
 Begins With
 Contains
 Name/Code
 Property
 Relationship

[Home](#) | [View Hierarchy](#) | [Help](#)

Anemia (Code E10010)

[Suggest changes to this concept](#)

Terms and Properties

Definition: A disorder characterized by an reduction in the amount of hemoglobin in 100 ml of blood. Signs and symptoms of anemia may include pallor of the skin and mucous membranes, shortness of breath, palpitations of the heart, soft systolic murmurs, lethargy, and fatigability.

Preferred Name: Anemia

Label: Anemia

Common Terminology Criteria for Adverse Events Code: E10010

Synonyms & Abbreviations: [\(see Synonym Details\)](#)

Anemia

External Source Codes:

NCIt Code

C2869 [\(see NCI Thesaurus info\)](#)

MedDRA Code

10002272 [\(see MedDRA info\)](#)

NCI-FDA Terminology Collaboration



- 2001- partnership and agreements in several terminology areas.
 - Structured Product Labeling (SPL)
 - Unique Ingredient Identifier (UNII)
 - Regulated Product Submission (RPS)
 - Individual Case Safety Report (ICSR)
 - Center for Devices and Radiological Health (CDRH)
 - Over 15,000 terms in more than 20 NCI subsets
- FDA PDUFA IV IT Plan:
 - “For terminology standards, the FDA partners with the National Cancer Institute Enterprise Vocabulary Services (EVS). The NCI EVS hosts the FDA terminologies and makes them freely available to the public.”
- FDA terminology resources are available on the NCI portal website:
<http://www.cancer.gov/cancertopics/terminologyresources/FDA>

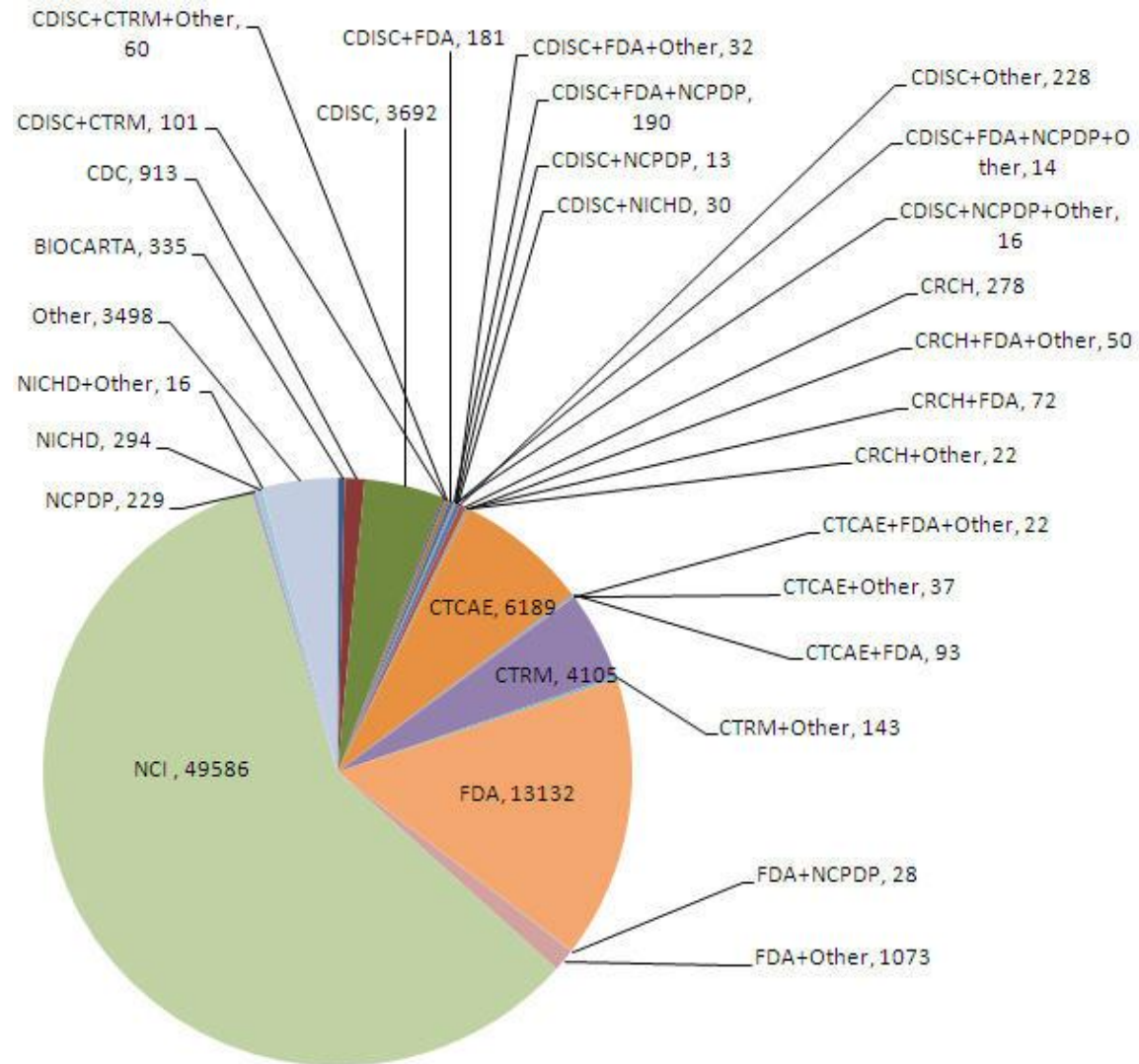
NCI-CDISC Terminology Collaboration

- 2002- partnership and agreements in several terminology areas.

	Numbers	Status
CDISC Controlled Terminology		
SDTM	3745	Published
CDASH	82	Published
ADaM	29	Published
SEND	1928	Pilot
CDISC (SDTM+CDASH) Metadata		
Domains	45	Published
Observation Classes	200	Harmonization
Variables	654	Harmonization
CDISC Disease Standards		
Cardiovascular	25	Mapping in EVS
Tuberculosis	130+	Mapping in EVS
Polycystic Kidney Disease	200+	Mapping in EVS

- CDISC terminology resources are available on the NCI portal website: <http://www.cancer.gov/cancertopics/terminologyresources/CDISC>

NCIt Terminology Subsets



Concept Code: A unique, permanent identifier



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

EVS Enterprise Vocabulary Services NCI Term Browser

mol Search ?

Exact Match
 Begins With
 Contains
 Name/Code
 Property
 Relationship

Home | View Hierarchy | Subsets | Help Visited Concepts

Quick Links

Mole (Code C42539) ← Concept Code

[Suggest changes to this concept](#)

Terms & Properties Relationships Synonym Details View All View in Hierarchy View History

Synonym Details

Term	Source	Type	Code
mole(s)	ICH	PT	0196
MOLE	FDA	AB	
MOLE	FDA	PT	
mole	HL7	PT	12856
Mole	NCI		
mole	UCUM		
Mole	CDISC		
mol	HL7		12856
mol	ICH		0196
Mol	NCI		
mol	NCI		
mol	UCUM	AB	
mol	CDISC	PT	

Terms

Term Source

6 sources

Additional Source Data

NCImetathesaurus

EVS Enterprise Vocabulary Services

Search: mole [Search] [?]

Exact Begins With Contains

Source: ALL [v]

Home | View NCI Hierarchy | Sources | Help | Term Suggestion

Quick Links [v]

Choose your source

Mole, unit of measurement (CUI C0439189)

[Suggest changes to this concept](#)

Terms & Properties Relationships Synonym Details By Source View All In NCI Hierarchy

Synonym Details

Term	Source [?]	Type [?]	Code
mol	CDISC	PT	MTHU119368
mol	ICH	AB	0196
mol	MGED	PT	MO_998
Mol	NCI	AB	C42539
mol	NCI	AB	C42539
mol	NCI-HL7	AB	12856
mol	SNOMEDCT	PT	258717005
mol	UCUM	AB	MTHU119368
Mole	CDISC	SY	MTHU120819
MOLE	FDA	AB	MTHU105387
MOLE	FDA	PT	MTHU105387
mole	HL7V3.0	PT	MOL
mole	HL7V3.0	PT	mol
Mole	NCI	PT	C42539
mole	NCI-HL7	PT	12856
mole	UCUM	PT	MTHU119367
mole - unit	SNOMEDCT	SY	258717005
mole - unit (qualifier value)	SNOMEDCT	FN	258717005
mole(s)	ICH	PT	0196
Mole, unit of measurement	MTH	PN	NOCODE

11 Sources

NCI Term Browser

<http://nciterns.nci.nih.gov>

The screenshot shows the NCI Term Browser web application. At the top, there is a red header with the National Cancer Institute logo and the text "National Cancer Institute" and "U.S. National Institutes of Health | www.cancer.gov". Below the header, the "EVS Enterprise Vocabulary Services" logo is on the left, and "NCI Term Browser" is on the right. A search bar is located in the top right, with a "Search" button and a help icon. Below the search bar, there are radio buttons for search criteria: "Exact Match", "Begins With", "Contains", "Name/Code", "Property", and "Relationship". A navigation bar contains "Home | Sources | Help" and a "Quick Links" dropdown menu. The main content area has buttons for "Select All", "All but NCI", "Clear", and "Search". Below these buttons, a warning message states: "Select NCI hosted terminologies to search, or click on a source name to go to its browser home page. (WARNING: Select All searches with thousands of hits may be slow; try NCI Metathesaurus separately.)". A list of sources follows, each with a checked checkbox and a link to its browser page. A red box labeled "Sources" points to this list.

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

EVS Enterprise Vocabulary Services NCI Term Browser

Search ?

Exact Match Begins With Contains
 Name/Code Property Relationship

Home | Sources | Help

Quick Links

Select All All but NCI Clear Search

Select NCI hosted terminologies to search, or click on a source name to go to its browser home page.
(WARNING: Select All searches with thousands of hits may be slow; try NCI Metathesaurus separately.)

- [NCI Metathesaurus \(201004\)](#)
- [NCI Thesaurus \(10.08e\)*](#)
- [BiomedGT: Biomedical Grid Terminology \(09.11b\)](#)
- [CTCAE: Common Terminology Criteria for Adverse Events \(4.03\)](#)
- [gene_ontology: Gene Ontology \(September 2010\)](#)
- [HL7 RIM: HL7 Reference Information Model \(V 02-31\)](#)
- [HUGO: Human Genome Organization \(July 2010\)](#)
- [ICD-10: International Classification of Diseases, Tenth Revision \(2009\)*](#)
- [ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification \(2010\)*](#)
- [LOINC: Logical Observation Identifiers Names and Codes \(224\)*](#)
- [MedDRA: Medical Dictionary for Regulatory Activities Terminology \(12\)*](#)
- [MGED Ontology: Microarray Gene Expression Data Ontology \(1.3.1\)](#)
- [NanoParticle Ontology: NanoParticle Ontology \(1.0_Jan_29_2010\)](#)
- [NCI Thesaurus Draft: National Cancer Institute Thesaurus Draft \(10.08e\)](#)
- [NDF-RT: National Drug File Reference Terminology Public Inferred Edition \(NDFRT_1009\)](#)
- [OBI: Ontology for Biomedical Investigations \(rc1\)](#)
- [PDQ: Physician Data Query \(Apr. 2010\)](#)
- [RadLex: Radiology Lexicon \(2.01\)*](#)
- [SNOMED CT: Systematized Nomenclature of Medicine-Clinical Terms \(2010_01_31\)*](#)
- [UMLS Semantic Network: UMLS Semantic Network \(3.2\)](#)
- [Zebrafish: Zebrafish Model Organism Database \(1.2_June_14_2010\)](#)

* caBIG approved.

Sources

Progress on Quality and Interoperability



- **EVS offers quality terminology infrastructure**
 - Consistent access to controlled terminology
 - Mapping between meanings
 - Reference terminology for both people and computers
 - Extensive internal and external quality measures
 - Being deployed throughout NCI trial & informatics systems
- **Community participation builds quality, adoption & standards**
 - Partnerships crucial to high-quality consensus on meaning
 - Interoperability a core component of quality
 - Consensus and interoperability build durable standards

EVS Products & Services Are Open



- **NCI Thesaurus is Open Content**
<http://evs.nci.nih.gov/terminologies>
- **NCI Metathesaurus is Mostly Open Source**
(See Each Source's License)
http://ncim.nci.nih.gov/ncimbrowser/pages/source_help_info.jsf
- **NCI EVS Servers Are Freely Accessible**
 - **On the Web:**
<http://nciterms.nci.nih.gov>
<http://ncimeta.nci.nih.gov>
 - **Via API:**
https://cabig.nci.nih.gov/tools/LexEVS_API
 - **On caGrid:**
<https://cabig.nci.nih.gov/workspaces/Architecture/caGrid>
- **All Software Developed by NCI EVS is Public Open Source and Free for the Asking:**
<http://ncicb.nci.nih.gov/download/#ETools>

Methods of Data Retrieval

- **NCI ftp site:**
<http://evs.nci.nih.gov/ftp1/FDA>
- **NCI partner web sites (CDISC, FDA, etc.)**
- **Request a report from NCI staff:**
http://ncit.nci.nih.gov/ncitbrowser/pages/contact_us
- **NCIt Browser by subset :**
<http://ncit.nci.nih.gov/pages/subset.jsf>
- **Cancer.gov:**
<http://www.cancer.gov/cancertopics/terminologyresources>

Shared Content Standards



NICHD
NHLBI
NINDS
NLM
NIH
"Roadmap"
caBIG



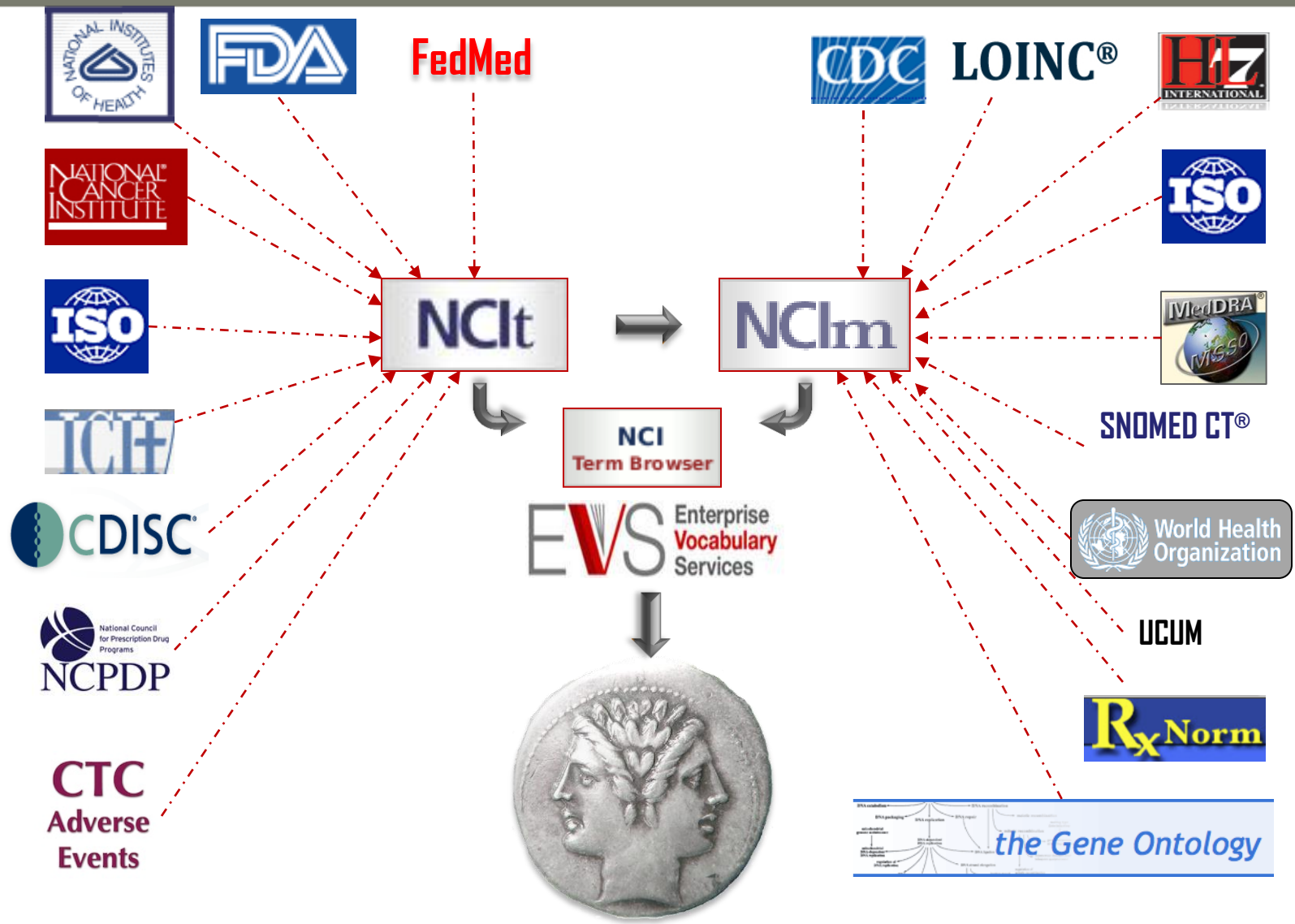
UNIs
ICSR
SPL
RPS
CDRH
Admin
Procedures
Other



SDTM
CDASH SEND
ADaM
Glossary
SHARE

Therapeutic Area
Standards

Consolidated Content Services



Contact Information



Lawrence W Wright
Associate Director
Enterprise Vocabulary Services
National Cancer Institute
lwright@mail.nih.gov

Margaret Haber
Associate Director
Enterprise Vocabulary Services
National Cancer Institute
mhaber@mail.nih.gov