

Ontologies in Biomedicine

Mark A. Musen
Stanford University

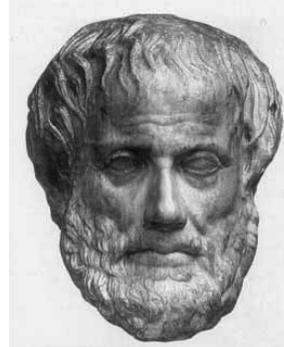


In this workshop ...

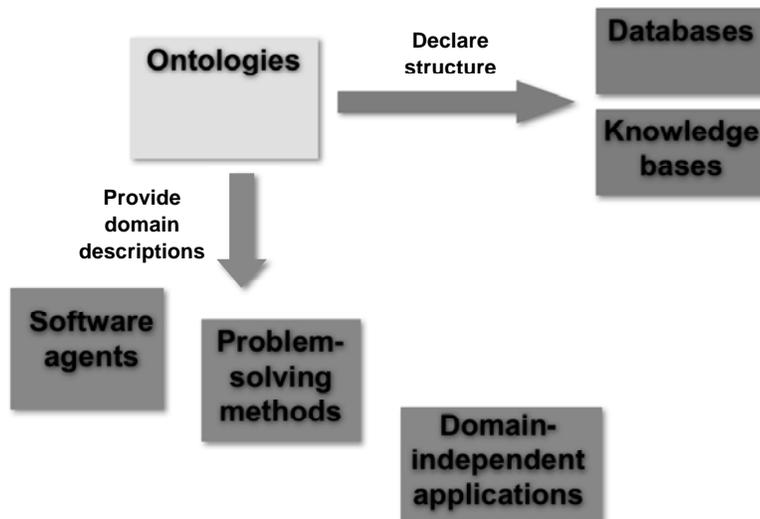
- You will learn about the EON guideline ontology
- You will use the EON guideline ontology to enter knowledge of particular guidelines
- You will wonder how this kind of work relates to other activities in biomedical informatics

What Is An Ontology?

- The study of being (Socrates and Aristotle 400–360 BC)
- Co-opted by Computer Science to connote the explicit description of the conceptualization of a domain:
 - Concepts
 - Properties and attributes of concepts
 - Constraints on properties and attributes
 - Individuals (often, but not always)
- An ontology defines
 - a common vocabulary
 - a shared understanding



Ontologies are just the beginning



Porphyry's depiction of Aristotle's Categories

Supreme genus:

Differentiae:

Subordinate genera:

Differentiae:

Subordinate genera:

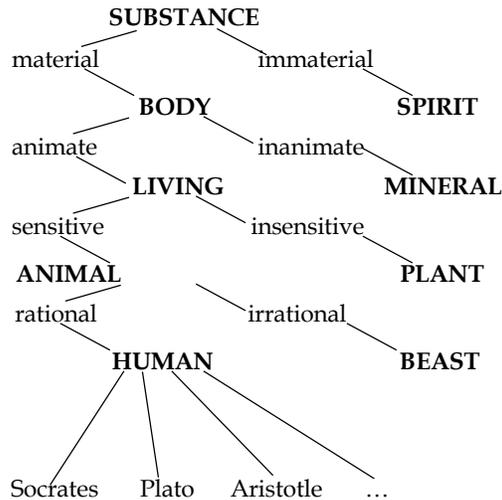
Differentiae:

Proximate genera:

Differentiae:

Species:

Individuals:



Web Images Groups **Directory** News

[Preferences](#) [Directory Help](#)

The web organized by topic into categories.

Arts Movies , Music , Television ,...	Home Consumers , Homeowners , Family ,...	Regional Asia , Europe , North America ,...
Business Industries , Finance , Jobs ,...	Kids and Teens Computers , Entertainment , School ,...	Science Biology , Psychology , Physics ,...
Computers Hardware , Internet , Software ,...	News Media , Newspapers , Current Events ,...	Shopping Autos , Clothing , Gifts ,...
Games Board , Roleplaying , Video ,...	Recreation Food , Outdoors , Travel ,...	Society Issues , People , Religion ,...
Health Alternative , Fitness , Medicine ,...	Reference Education , Libraries , Maps ,...	Sports Basketball , Football , Soccer ,...
World Deutsch , Español , Français , Italiano , Japanese , Korean , Nederlands , Polska , Svenska , ...		

Open Directory Project

- Started in 1998 as a volunteer effort to develop an open-content directory of Web pages
- In its first year, 4500 editors had indexed 100K Web sites
- By July 2005, 69K editors had indexed 4.6M sites using 580K categories
- On average, between 9K and 10K volunteer editors are working on ODP at any given time

 open directory project

Foundational Model of Anatomy

- Long-term project at University of Washington to create a comprehensive ontology of human anatomy
- 72K concepts, 1.9M relationships
- One of the largest and best developed ontologies in biomedicine



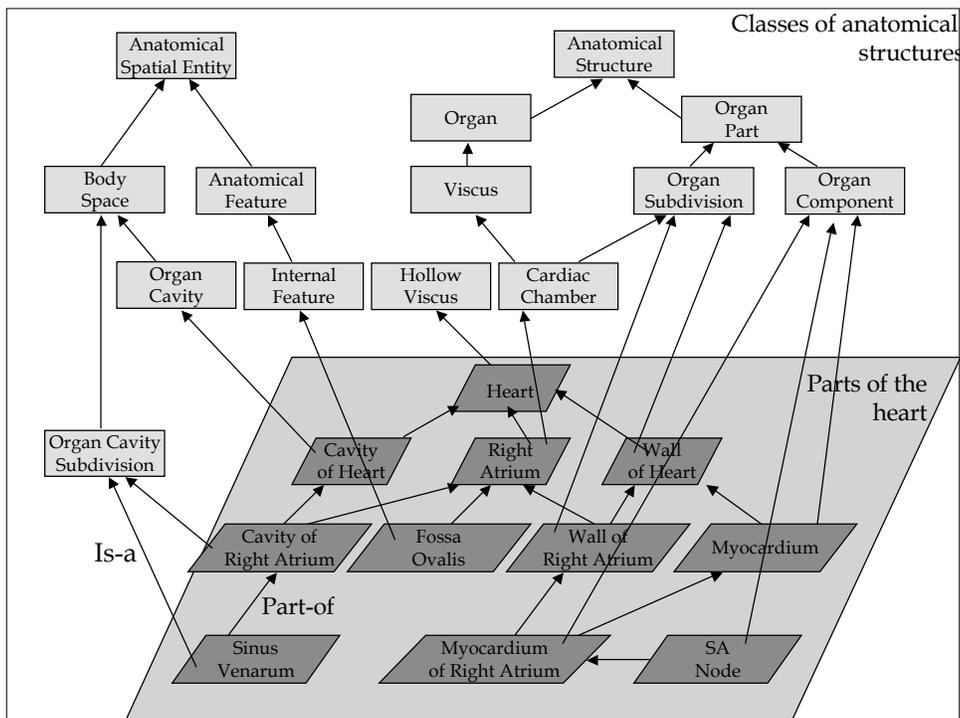
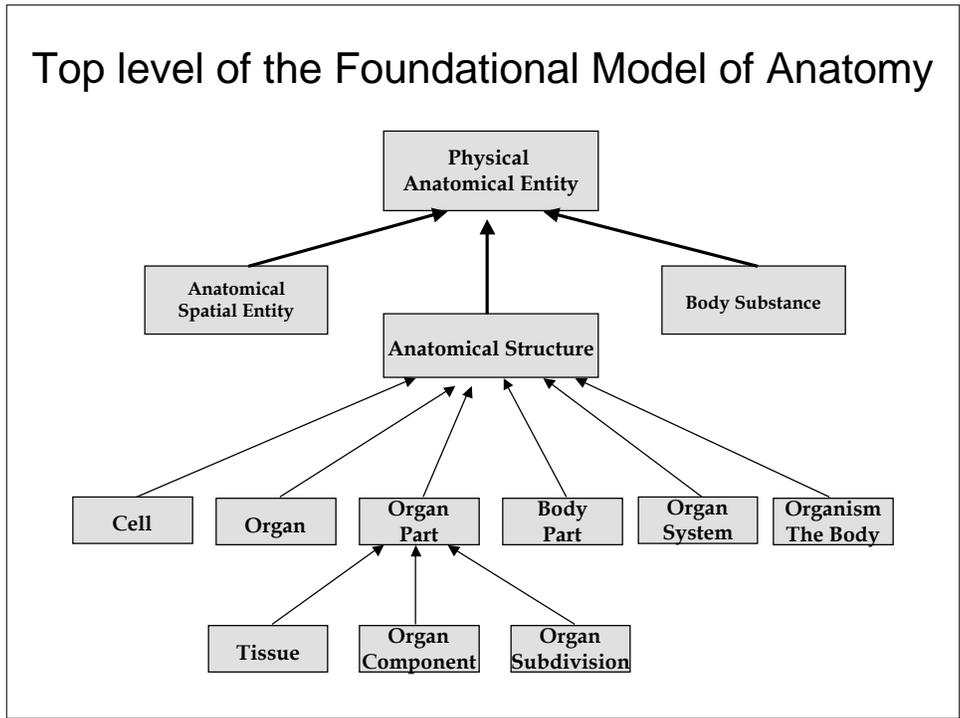
Structural Informatics Group
Structural Informatics Group
Structural Informatics Group

SIG

Overview Projects Personnel Publications Docs Local Info Products

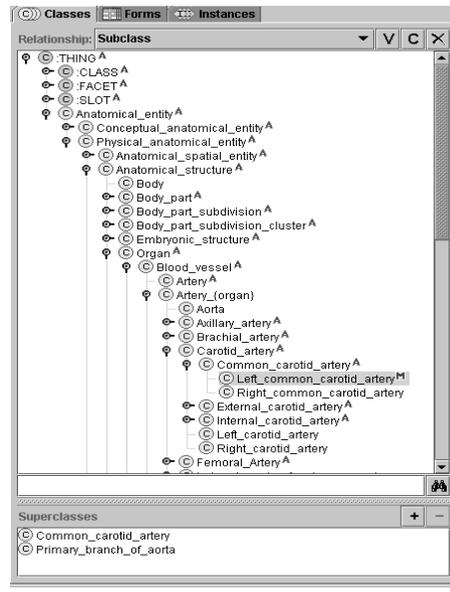
Biological Structure Biomedical Informatics Computer Science and Engineering University of Washington

Top level of the Foundational Model of Anatomy



But we really want ontologies in electronic form

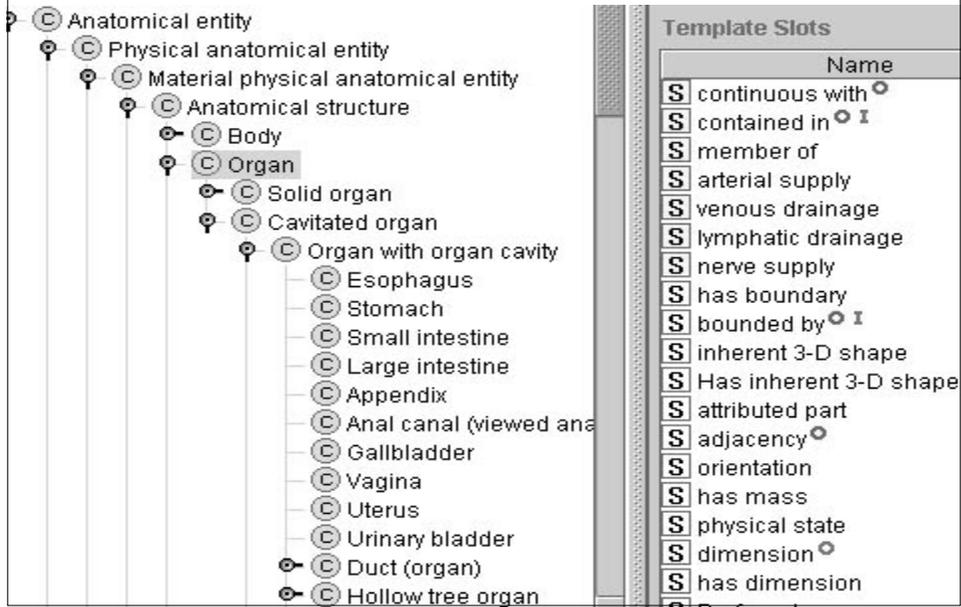
- Ontology contents can be processed and interpreted by computers
- Interactive editors can assist developers in ontology authoring



The components of ontologies

- **Classes:** The primary entities in the world being models (e.g., “organ”)
- **Attributes:** The properties of classes (e.g., “shape”, “location”)
- **Relations:** Statements regarding how one class may relate to others (e.g., “the heart” is-a “organ”)
- **Axioms:** More complex logical statements (e.g., “only paired organs can be left-sided or right-sided”)

Classes and attributes in the FMA



Attributes of a class (e.g., "Esophagus")

CLASS EDITOR

For Class: Esophagus (instance of Organ with organ cavity)

Preferred Name

Esophagus

Has Boundary Has Dimension Dimension: 3-dimension

Has Mass Has Inherent 3-D Sh... Inherent 3-D Shape: Hollow cylinder

Has Physical State Solid Physical State:

Regional Part

- Cervical part of esophagus
- Thoracic part of esophagus
- Abdominal part of esophagus

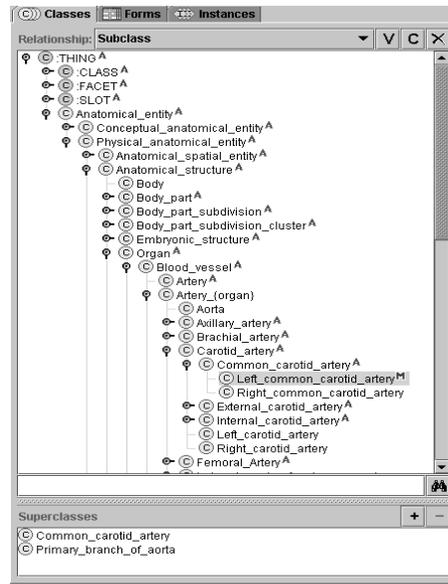
Regional Part Of

- Foregut
- Upper gastrointestinal tract
- Gut
- Gastrointestinal system

“is-a” is a special relation

If a sub-class *is-a* member of a super-class, then

- every instance of the sub-class is also an instance of the super-class (e.g., every member of the set *aorta* is necessarily a member of the set *artery*)
- Values of attributes of the super-class are *inherited* by every instance of the sub-class (e.g., if *arteries* have cylindrical shape, then *aorta* has cylindrical shape)



“Frame-based” knowledge-representation systems

- Allow developers to encode
 - Taxonomic hierarchies of classes
 - Other relations among classes (e.g., “part-of”) in addition to the is-a hierarchy
 - Attributes of classes that take on particular values to define *instances* of the classes
- Support inheritance of attributes and values along taxonomic relations

protégé Search:

HOME | OVERVIEW | DOCUMENTATION | DOWNLOADS | **COMMUNITY** | ABOUT US

welcome to protégé

Protégé is a free, open source ontology editor and knowledge-base framework.

The Protégé platform supports two main ways of modeling ontologies via the Protégé-Frames and Protégé-OWL editors. Protégé ontologies can be exported into a variety of formats including RDF(S), OWL, and XML Schema. (more)

Protégé is based on Java, is extensible, and provides a plug-and-play environment that makes it a flexible base for rapid prototyping and application development. (more)

Protégé is supported by a strong community of developers and academic, government and corporate users, who are using Protégé for knowledge solutions in areas as diverse as biomedicine, intelligence gathering, and corporate modeling.

news

Protégé Short Course
March 28-31, 2006
Stanford, California

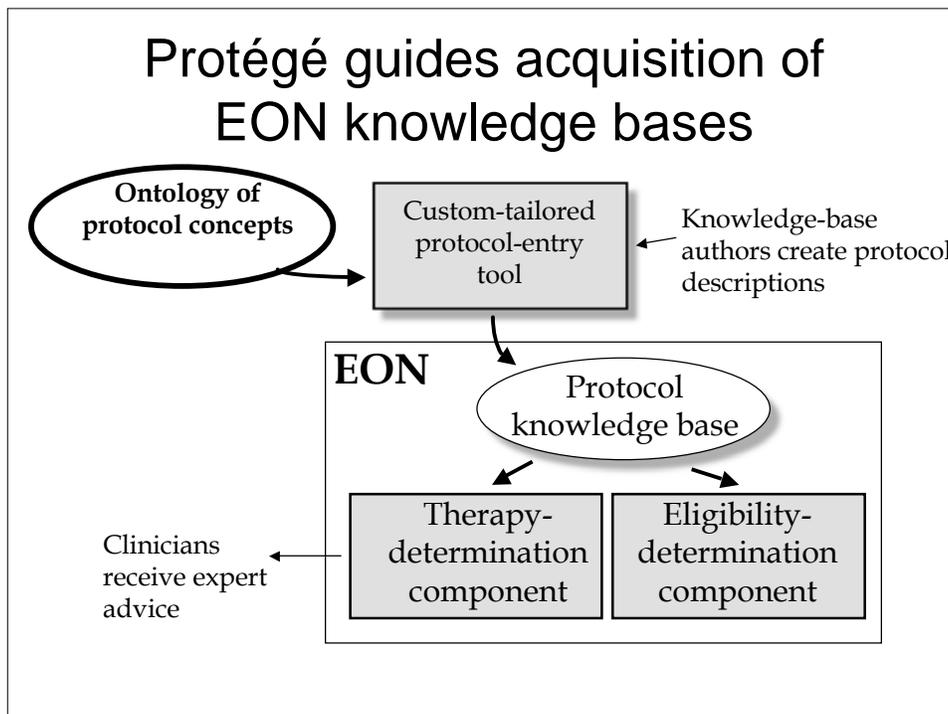
Protégé Conference
July 23-26, 2006
Stanford, California

community

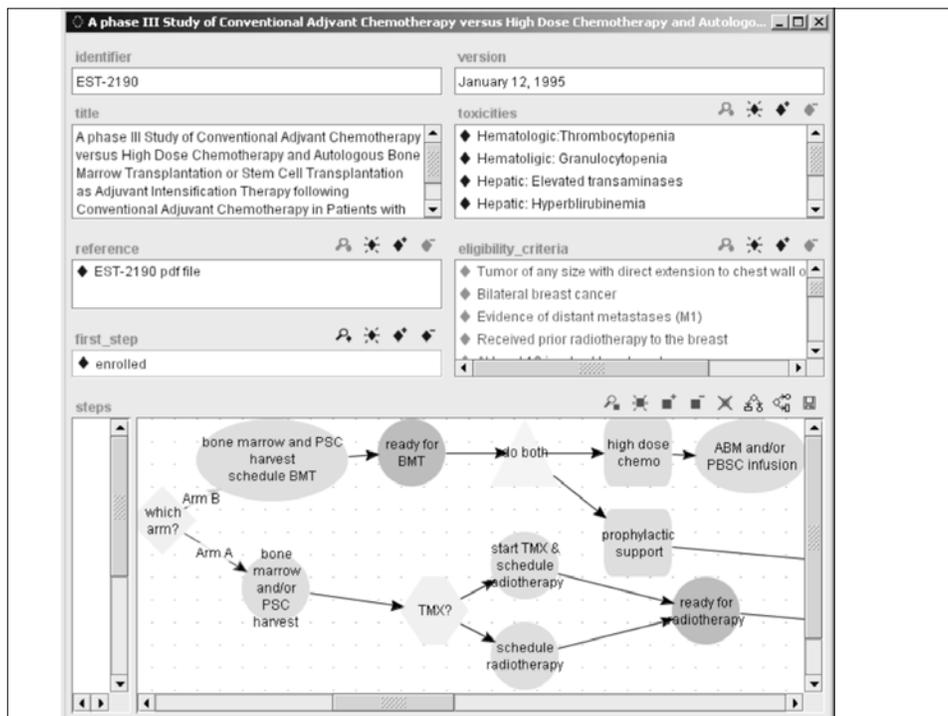
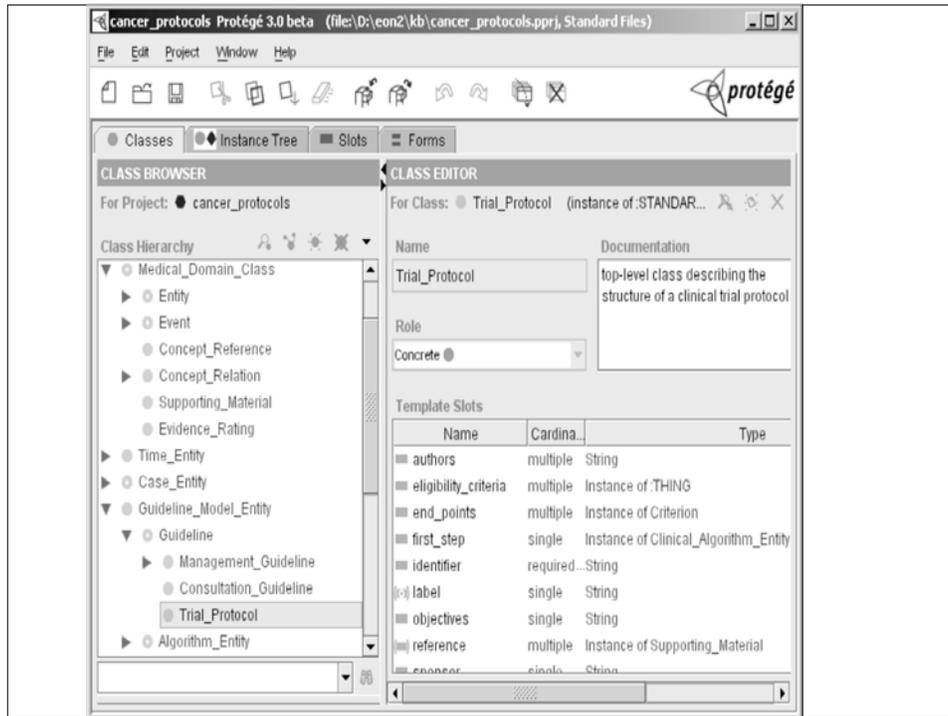
Registered Users	41,557
protége-users list members	16,160
protége-discussion list members	5,369
protége-owl list members	2,066

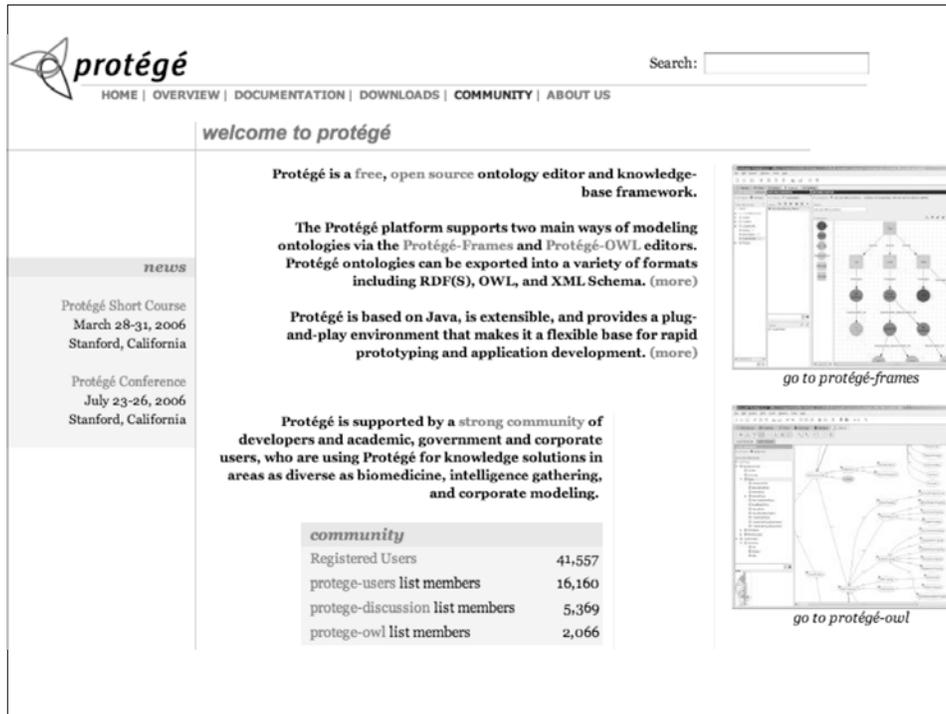
go to protégé-frames

go to protégé-owl



B. Ontologies in Biomedicine





The screenshot shows the Protégé website homepage. At the top left is the Protégé logo, followed by navigation links: HOME | OVERVIEW | DOCUMENTATION | DOWNLOADS | **COMMUNITY** | ABOUT US. A search box is located at the top right. Below the navigation is a "welcome to protégé" message. The main content area features three paragraphs of text describing Protégé as a free, open source ontology editor and knowledge-base framework, supported by two main modeling editors (Protégé-Frames and Protégé-OWL), based on Java, and supported by a strong community. To the left is a "news" sidebar with two entries: "Protégé Short Course March 28-31, 2006 Stanford, California" and "Protégé Conference July 23-26, 2006 Stanford, California". To the right are two screenshots of the Protégé-Frames and Protégé-OWL editors, each with a "go to" link below it. At the bottom center is a "community" table.

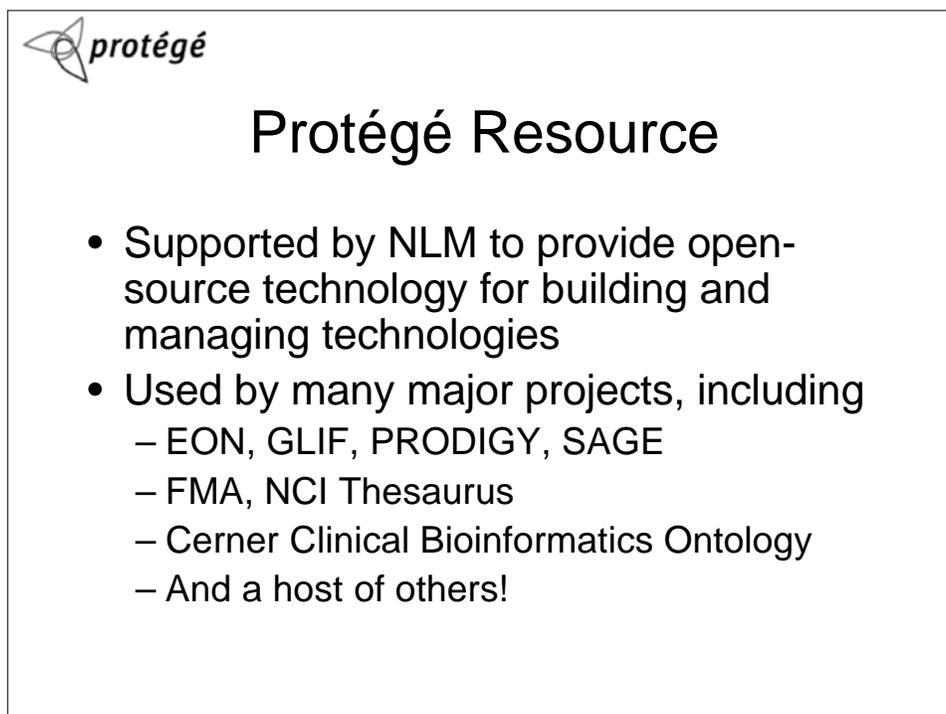
Protégé is a free, open source ontology editor and knowledge-base framework.

The Protégé platform supports two main ways of modeling ontologies via the Protégé-Frames and Protégé-OWL editors. Protégé ontologies can be exported into a variety of formats including RDF(S), OWL, and XML Schema. (more)

Protégé is based on Java, is extensible, and provides a plug-and-play environment that makes it a flexible base for rapid prototyping and application development. (more)

Protégé is supported by a strong community of developers and academic, government and corporate users, who are using Protégé for knowledge solutions in areas as diverse as biomedicine, intelligence gathering, and corporate modeling.

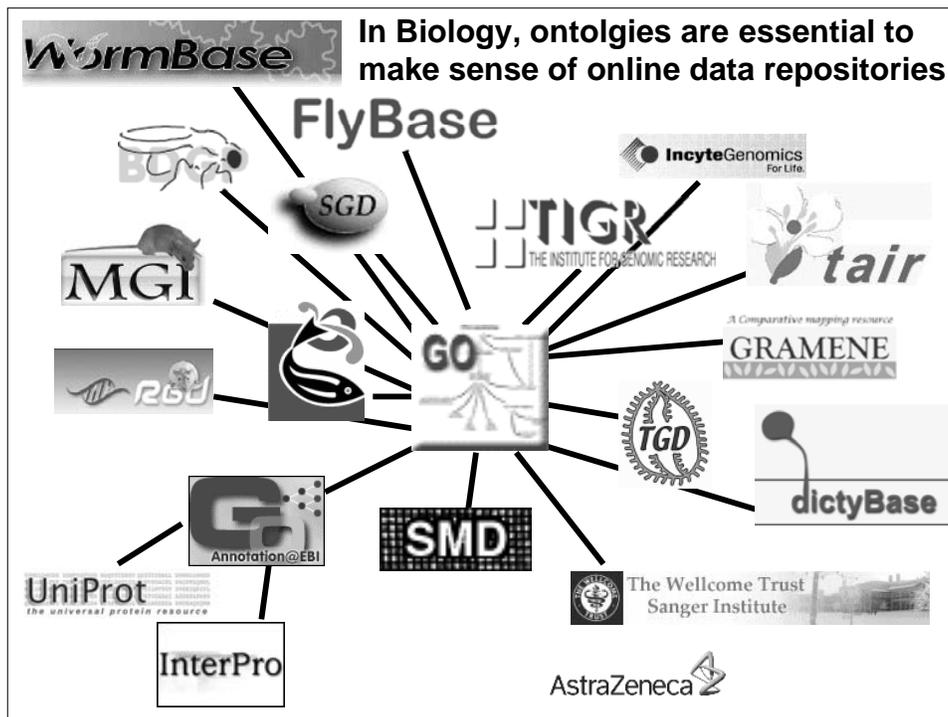
community	
Registered Users	41,557
protege-users list members	16,160
protege-discussion list members	5,369
protege-owl list members	2,066



The slide features the Protégé logo at the top left and the title "Protégé Resource" in the center. Below the title is a bulleted list of key points about the project's support and usage.

Protégé Resource

- Supported by NLM to provide open-source technology for building and managing technologies
- Used by many major projects, including
 - EON, GLIF, PRODIGY, SAGE
 - FMA, NCI Thesaurus
 - Cerner Clinical Bioinformatics Ontology
 - And a host of others!



Biologist have adopted ontologies

- To provide canonical representation of scientific knowledge
- To annotate experimental data to enable interpretation and comparison across databases
- To facilitate knowledge-based applications for
 - Decision support
 - Natural language-processing
 - Data integration

A Portion of the OBO Library

Domain	Prefix	Ontology	Defs file
Arabidopsis gross anatomy	TAIR	arabidopsis_anatomy.ontology	arabidopsis_anatomy.definitions
Arabidopsis development	TAIR	arabidopsis_development.ontology	arabidopsis_development.definitions
Cell type	CL	cell.obo	Included in cell.obo
Cereal plant gross anatomy	GRO	anatomy_gr_ont	anatomy_gr_def
Cereal plant development	GRO	temporal_gr_ont	temporal_gr_def
Cereal plant trait ontology	TO	trait_ontology	trait_definitions
Chemical entities of biological interest	CHEBI	ontology.obo	Included in ontology.obo
Protein covalent bond	CV	[none]	[none]
Protein-protein interaction	MI	psi-mi.dag	psi-mi.def
Maize gross anatomy	ZEA	Zea_mays_anatomy_ontology.txt	Zea_mays_anatomy_ontology_definitions.txt
Dictyostellium anatomy	DDANAT	anatomy.ontology	anatomy.definitions
Drosophila gross anatomy	FBbt	fly_anatomy.ontology	fly_anatomy.definitions
Habronattus courtship		protege_source	Included in protege_source
Loggerhead nesting		protege_source	Included in protege_source
Human anatomy and development	EV	ontologies	[none]
Microarray experimental conditions		MGEDOntology.daml	Included in MGEDOntology.daml
Physical-chemical methods and properties	FIX	fix.ontology	[none]
Fungal gross anatomy	FAO	fungus_anatomy.ontology	fungus_anatomy.definitions
Molecular function	GO	gene_ontology.obo	Included in gene_ontology.obo
Biological process	GO	gene_ontology.obo	Included in gene_ontology.obo
Cellular component	GO	gene_ontology.obo	Included in gene_ontology.obo


NIH Roadmap
ACCELERATING MEDICAL DISCOVERY TO IMPROVE HEALTH



- ▶ [Overview](#)
- ▶ [NIH Roadmap Initiatives](#)
- ▶ [Funding Opportunities](#)
- ▶ [Funded Research](#)
- ▶ [Roadmap Related Activities](#)
- ▶ [Public Meetings and Workshops](#)
- ▶ [Frequently Asked Questions](#)
- ▶ [News and Information](#)
- ▶ [NIH Roadmap Institute and Center Liaisons](#)
- ▶ [Subscribe to the NIH Roadmap E-mail list](#)

New Pathways to Discovery

- ▶ [Building Blocks, Biological Pathways, and Networks](#)
- ▶ [Molecular Libraries and Imaging](#)
- ▶ [Structural Biology](#)
- ▶ [Bioinformatics and Computational Biology](#)
- ▶ [Nanomedicine](#)

Research Teams of the Future

- ▶ [High-Risk Research](#)
 - [NIH Director's Pioneer Award](#)
- ▶ [Interdisciplinary Research](#)
- ▶ [Public-Private Partnerships](#)

Re-engineering the Clinical Research Enterprise

- ▶ [Re-engineering the Clinical Research Enterprise Initiatives](#)
 - [Clinical Research Networks and NECTAR](#)
 - [Clinical Outcomes Assessment](#)
 - [Clinical Research Training](#)
 - [Clinical Research Policy Analysis and Coordination](#)
 - [Translational Research](#)

What's New

- ▶ [Press Release: NIH Launches Major Program to Transform Clinical and Translational Science](#)
- ▶ [RFA: Planning Grants for Institutional Clinical and Translational Science Awards](#)
- ▶ [RFA: Institutional Clinical and Translational Science Award](#)
- ▶ [Program: Institutional Clinical and Translational Science Award Program Information](#)
- ▶ [Meeting: Interdisciplinary Research Centers Workshop](#)
- ▶ [Press Release: 2005 NIH Director's Pioneer Award Recipients Announced](#)
- ▶ [Press Release: NIH Roadmap Continues to Move Forward on All Fronts](#)
- ▶ [Meeting Summary: BAA Roadmap Steering Committee, May 2005](#)
- ▶ [What's New – Archives](#)



- One of 7 National Centers for Biomedical Computing
- Goal is to create new technology for accessing, using, and disseminating ontologies in biomedicine
- Led by Mark Musen (Stanford) and Suzanna Lewis (Berkeley)



- Stanford: Tools for ontology alignment, indexing, and management (Mark Musen)
- Lawrence–Berkeley Labs: Tools to use ontologies for data annotation (Suzanna Lewis)
- Mayo Clinic: Tools for access to large controlled terminologies (Christopher Chute)
- Victoria: Tools for ontology visualization (Margaret-Anne Story)
- University at Buffalo: Dissemination of best practices for ontology engineering (Barry Smith)

cBio plans to offer technologies



- To help build and extend ontologies
- To locate ontologies and to relate them to one another
- To visualize relationships and to aid understanding
- To facilitate evaluation and annotation of ontologies

cBio *BioPortal* will offer Web-based tools to



- Browse ontology repositories
- Edit individual ontologies
- Index and align different ontologies
- Visualize ontologies and alignments among ontologies
- Enable peer review of ontologies

Ontologies are everywhere in biomedicine!

- From controlled terminologies to complex guideline models
- Can be edited with widely available tools such as Protégé
- Are being curated and disseminated via the National Center for Biomedical Ontology
- Form the basis of modern knowledge-based systems to deliver decision support