When the World Beats a Path to Your Door: Collaboration in the Era of Big Data

Mark A. Musen, M.D., Ph.D.
Stanford Center for Biomedical Informatics Research
Stanford University School of Medicine
musen@Stanford.EDU
It’s all part of the same problem space ...
A Tale of Three Projects

1. Protégé Ontology Editor (1988–present)


3. CEDAR Metadata Editing System (2014–present)
A free, open-source ontology editor and framework for building intelligent systems

Protégé is supported by a strong community of academic, government, and corporate users, who use Protégé to build knowledge-based solutions in areas as diverse as biomedicine, e-commerce, and organizational modeling.

DOWNLOAD NOW    USE WEBPROTÉGÉ

TRUSTED BY OVER 310,610 USERS
Mailing Lists

Get free support for all of your Protégé questions via our mailing lists. The lists are actively monitored by Stanford’s Protégé team, as well as many experienced Protégé users from the community at large.

Click the Subscribe button next to the list you wish to join. Please note that you must be subscribed to a list before you can post messages.

Protégé User Support
User support for all versions of WebProtégé and Protégé Desktop.

Protégé Developer Support
Developer support for all versions of WebProtégé and Protégé Desktop.

Protégé Announcements
Low traffic, announcement-only list for new releases, availability of short courses, and information regarding Protégé Conferences.
Congratulations to Geoff for being the 300,000th registered Protege user! Geoff told us that he downloaded Protege to inspect the RDF representation of FHIR (Fast Healthcare Interoperability Resources), which is a standard for exchanging healthcare information electronically.

We are very happy and proud to have reached this important milestone, and would like to thank the community for their amazing support all throughout these years!
Welcome to the Protege Plugin Library!

The Protege Plugin Library offers a convenient place for the Protege community to find open source and commercial Protege plugins that enhance the Protege application.

The plugins on our Wiki were developed either here at Stanford or by our user community, as extensions to the core Protege system, or to the Protege-Frames editor, or Protege-OWL editor. Follow individual instructions to download and install the plugins that are of interest to you. We always welcome contributions from our user community. If you have developed software that you think would be of interest to others, please register your plugin and your organization (see the green box on the right-hand side of this page).

It may be of interest to our user community that the Protege Wiki uses the Semantic MediaWiki extension and thus is a true semantic wiki. By using semantic technology, we are able to provide a knowledge repository and platform for publishing information about Protege plugins as well as intelligent search capabilities for particular plugins and associated organizations.

Topics

Browse plugins by topics of interest.

- Biomedical Informatics
- Code Examples
- Export
- Import
- Inference
- Natural Language Processing
- Navigation
- Project Management
- Query
- Reasoning
- Search
- Semantic Web
- Software Engineering
- Terminologies
- Validation
- Visualization

Types

Browse plugins by type.

- API
- Application
- Backend
- Import
- Export
- Project
- Reasoner
- Slot Widget
- Tab Widget
- View

News

List of recently updated Protege plug-in pages.

<table>
<thead>
<tr>
<th>Plugin</th>
<th>Last update</th>
</tr>
</thead>
<tbody>
<tr>
<td>VOWL</td>
<td>12 November 2016</td>
</tr>
<tr>
<td>DL-Learner</td>
<td>1 November 2016</td>
</tr>
<tr>
<td>ELK</td>
<td>11 January 2016</td>
</tr>
</tbody>
</table>

Register your Plugin

Add a new plugin in 2 steps:

1. Enter the name of your plugin, which will then be the title of a new article in this wiki.
2. Enter as much information as possible in the provided form (don't forget to save your changes).

Register your Organization

Add a new organization in 2 steps:

1. Enter the name of your organization, which will then be the title of a new article in this wiki.
2. Enter as much information as possible in the provided form (don't forget to save your changes).
Collaboration in Protégé

• Happens nearly spontaneously
• Is encouraged by very active e-mail lists and presence on social media
• Is enabled by having standard mechanisms for developing plug-ins
• Is self-reinforcing when there are thousands of engaged users
A Tale of Three Projects

1. Protégé Ontology Editor (1988–present)


3. CEDAR Metadata Editing System (2014–present)
Page Requests
Unique IP Addresses
Who is using NCBO technology?
Collaboration with BioPortal

• Is modest in comparison to the size of the user community
• Is impeded by lack of a plug-in architecture for the main server
• Typically takes place silently when users develop new technology that simply calls our APIs without the need to interact with NCBO personnel
A Tale of Three Projects

1. Protégé Ontology Editor (1988–present)


3. CEDAR Metadata Editing System (2014–present)
DISEASE, CENTER, PROVIDER

- **DISEASE**
  - CL_Disease
  - CL_Disease_Detail
  - CL_Disease_Site_Onset
  - CL_Disease_Age_Onset

- **CENTER**
  - CL_Center_Name
  - CL_Center_Specific_ID

- **PROVIDER**
  - CL_Provider_Name
  - CL_Provider_Catalog_ID
DISEASE, CENTER, PROVIDER

**DISEASE**

- **CL_Disease**
  - melanoma
  - Malignant melanoma of other specified sites of skin
    - Malignant melanoma of skin
    - Malignant melanoma of skin of ear and external auditory canal
    - Malignant melanoma of skin of eyelid, including canthus
    - Malignant melanoma of skin of lip
    - Malignant melanoma of skin of lower limb, including hip
    - Malignant melanoma of skin of other and unspecified parts of face
    - Malignant melanoma of skin of unspecified part of body

**CENTER**

- **CL_Center_Name**
- **CL_Center_Specific_ID**

**PROVIDER**

- **CL_Provider_Name**
- **CL_Provider_Catalog_ID**
LINCS aims to create a network-based understanding of biology by cataloging changes in gene expression and other cellular processes that occur when cells are exposed to a variety of perturbing agents.
ANNOUNCEMENTS

The BD2K Guide To The Fundamentals Of Data Science

BioCADDIE 2016 Dataset Retrieval Challenge

DataMed Prototype - Call For Feedback

DataMed V1.0 Is Live

ACCESS DATAMED

METADATA STANDARDS
Collaboration with CEDAR

• Is generating “buzz,” as lots of groups want to use our technology
• Is pragmatically difficult, since adoption of CEDAR necessarily leads to a change in user workflow
• Initially incurs cost in staff time—both at our site and at those of our collaborators
• Requires an explicit strategy, explicit coordination, and explicit funding
A Tale of Three Projects

1. Protégé Ontology Editor (1988–present)
3. CEDAR Metadata Editing System (2014–present)
Collaboration: “It’s complicated”

- Collaboration may be an emergent phenomenon, given the right ecosystem
- Collaboration may take time to take root and become part of the culture
- With complex systems such as CEDAR, collaboration must be
  - Planned
  - Managed
  - Funded