

Exploring Changes in Life Science Ontologies with OnEX

Michael Hartung¹, Toralf Kirsten^{1,3}, Anika Groß¹, Erhard Rahm^{1,2}

¹ Interdisciplinary Centre for Bioinformatics, University of Leipzig

² Department of Computer Science, University of Leipzig

³ Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig

<http://www.izbi.de>

<http://dfs.uni-leipzig.de>

<http://imise.uni-leipzig.de>

Motivation

Increasing number of evolving life science ontologies:

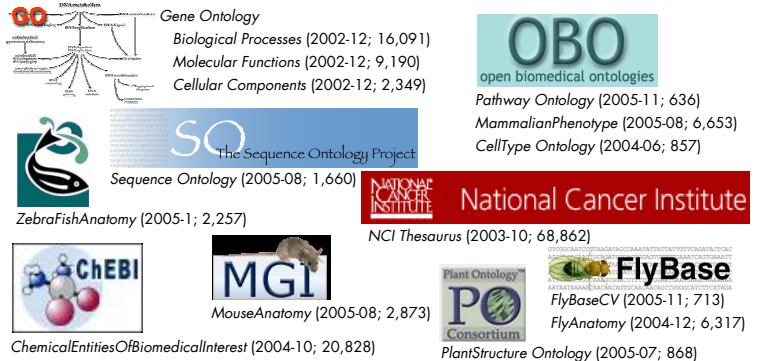
- Causes: new insights and experimental results, revision of existing knowledge
- Result: ontologies need to incorporate changed knowledge → ontology versions that are only valid in specific time periods

Goals:

- Quantitative evolution analysis of life science ontologies**
 - Ontologies in general and their parts, e.g., concepts and relationships
 - Long term evolution analysis (> 2 years)
- Answering of open questions concerning ontology evolution**
 - What are the typical changes in ontologies and how often do they occur?
 - How stable (in)stable are ontologies?
 - Which ontologies are currently highly developed or reside in a final state?
 - How does a single ontology concept evolve over time?
- Online system for ad-hoc evolution analysis**
 - Intuitive and easy-to-use interface for accessing / browsing analysis results

Current ontology developments in the life sciences:

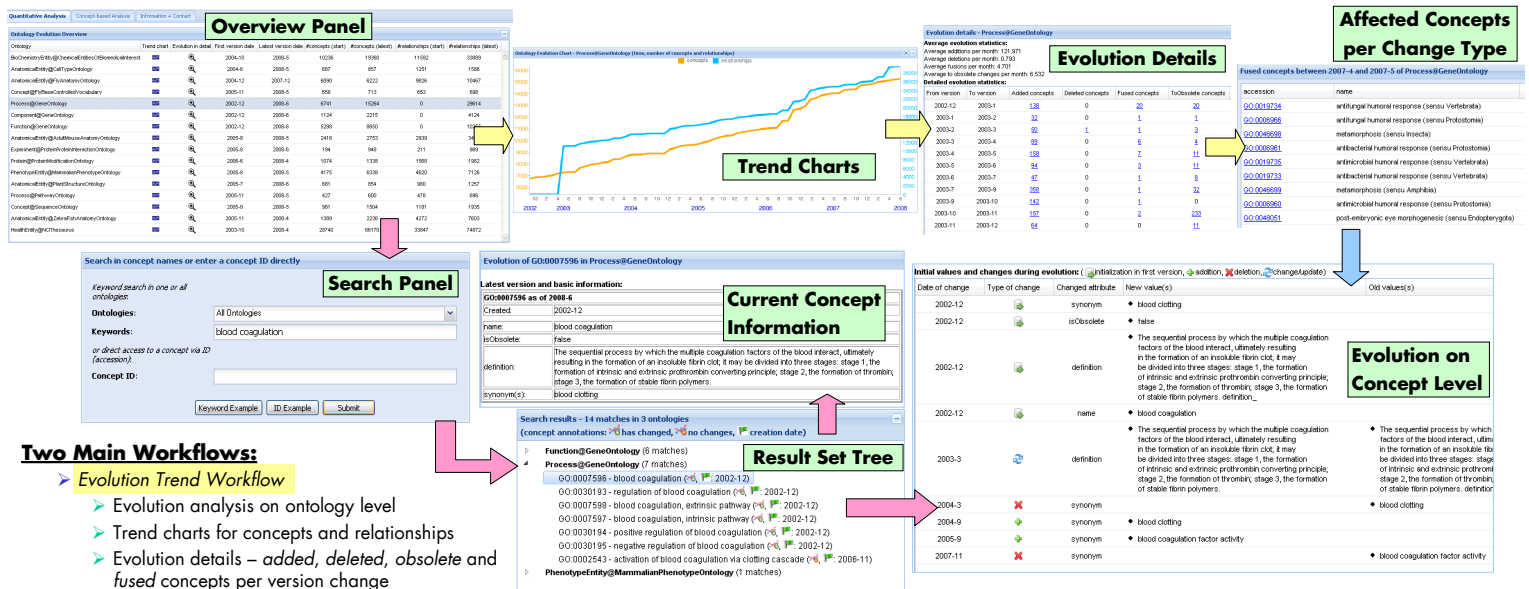
Ontology (available since; number of concepts in latest version)



- Gene Ontology**
 - Biological Processes (2002-12; 16,091)
 - Molecular Functions (2002-12; 9,190)
 - Cellular Components (2002-12; 2,349)
- OBO** (open biomedical ontologies)
 - Pathway Ontology (2005-11; 636)
 - MammalianPhenotype (2005-08; 6,653)
 - CellType Ontology (2004-06; 857)
- Sequence Ontology** (2005-08; 1,660)
- ZebrafishAnatomy** (2005-1; 2,257)
- National Cancer Institute**
 - NCI Thesaurus (2003-10; 68,862)
- Chemical Entities of Biomedical Interest** (2004-10; 20,828)
- MouseAnatomy** (2005-08; 2,873)
- Plant Structure Consortium**
 - Plant Structure Ontology (2005-07; 868)
- FlyBase**
 - FlyBaseCV (2005-11; 713)
 - FlyAnatomy (2004-12; 6,317)

*Hartung, M.; Kirsten, T.; Rahm, E.: Analyzing the Evolution of Life Science Ontologies and Mappings. Proc. 5th International Workshop on Data Integration in the Life Sciences (DILS), Paris, 2008

OnEX Application – <http://www.izbi.de/onex>



The screenshot displays the OnEX application interface with several key components:

- Overview Panel:** A table showing evolution statistics for various ontologies, including columns for ontology name, first version, latest version, and number of concepts.
- Trend Charts:** A line graph showing the growth of concepts over time for selected ontologies.
- Evolution Details:** A table showing changes in concepts over time, including the type of change (e.g., synonym, deletion) and the affected concepts.
- Affected Concepts per Change Type:** A table listing concepts affected by specific changes, such as 'blood clotting'.
- Search Panel:** A search interface for finding concepts by name or ID.
- Current Concept Information:** A detailed view of a specific concept, including its definition and synonyms.
- Result Set Tree:** A hierarchical tree view showing the evolution of a concept across different ontologies and versions.

Two Main Workflows:

- Evolution Trend Workflow**
 - Evolution analysis on ontology level
 - Trend charts for concepts and relationships
 - Evolution details – added, deleted, obsolete and fused concepts per version change
- Concept Evolution Workflow**
 - Search in specific or across all ontologies
 - Evolution analysis on concept level – attributes and relationships of a concept

Current Content:

- Ontologies of different life science fields, e.g., proteomics, anatomy, phenotype, biomedical chemistry and cancer research
- Approx. 520 versions of 16 life science ontologies accessible

Future Work

- Measures for assessing ontology stability by utilization of evolution information
- Discovery of "Hot / Cold Topics" in an ontology by
 - Exploring changes in sub graphs of an ontology
 - Studying changes in annotations to ontology concepts

- Inclusion of structural changes
 - Changes in the semantic neighborhood of a concept, e.g., parents and children
 - Changes in paths of a concept

Example:
GO:0007596
"blood coagulation" path changes
[2004-05 - 2008-12]

