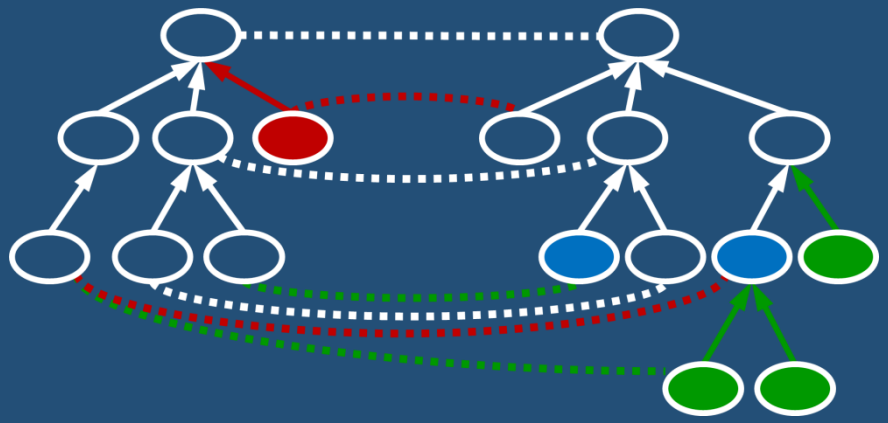


How do Ontology Mappings Change in the Life Sciences?

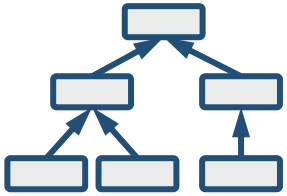
Anika Gross, Michael Hartung,
Andreas Thor, Erhard Rahm

UNIVERSITÄT LEIPZIG

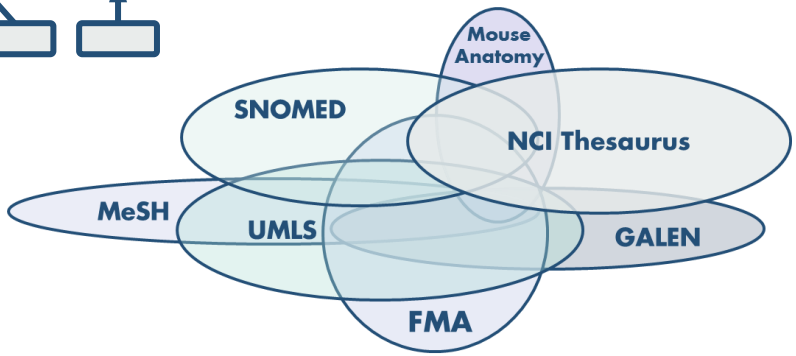
Poster Presentation, 28th June 2012, DILS, College Park, Maryland

Ontologies and Ontology Mappings

Ontologies



- Knowledge representation
- Multiple interrelated ontologies in a domain



Ontology mapping

- Set of semantic correspondences between concepts of different ontologies
- Manual identification or (semi-) automatic matching approaches

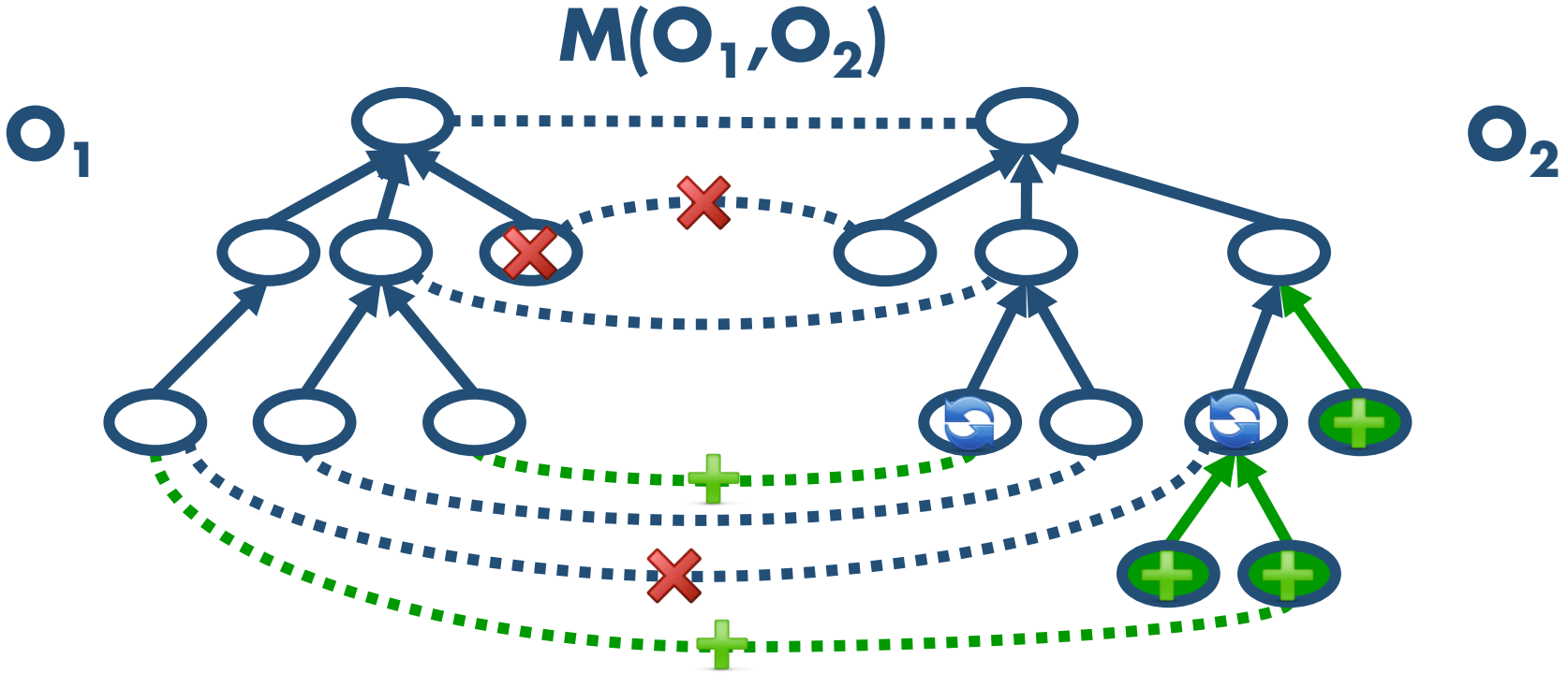
Use of mappings

- Ontology merging - creation of the integrated cross-species anatomy ontology "Uber ontology"
- Knowledge transfer - experiments for different species
- Ontology curation - find missing ontology annotations
- …

Ontology Evolution

- Ongoing research, new findings → continuous modifications
- Periodical release of new ontology versions
- Ontology changes

- **Invalidate previously determined ontology mappings?**



Ontology Evolution

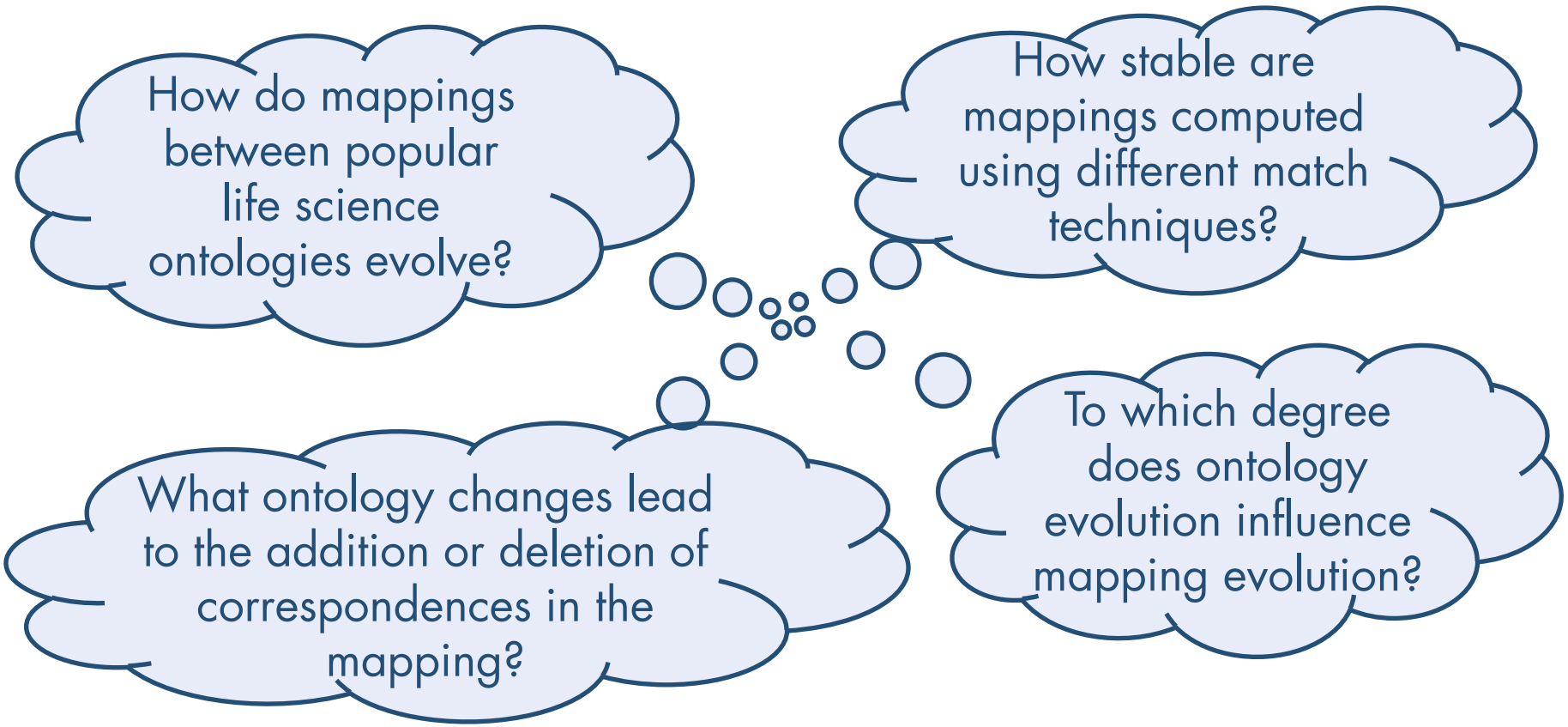
- Ongoing research, new findings → continuous modifications
- Periodical release of new ontology versions
- Ontology changes 

- **Invalidate previously determined ontology mappings?**
- Example: Anatomy reference mapping at *OAEI* *
  - Based on 5 year old versions
  - Quality w.r.t. current ontology versions?
- Re-determination of mappings is an expensive process
  - Manual verification of correspondences
  - Parametrization effort
- Future aim: (semi-)automatic mapping adaptation

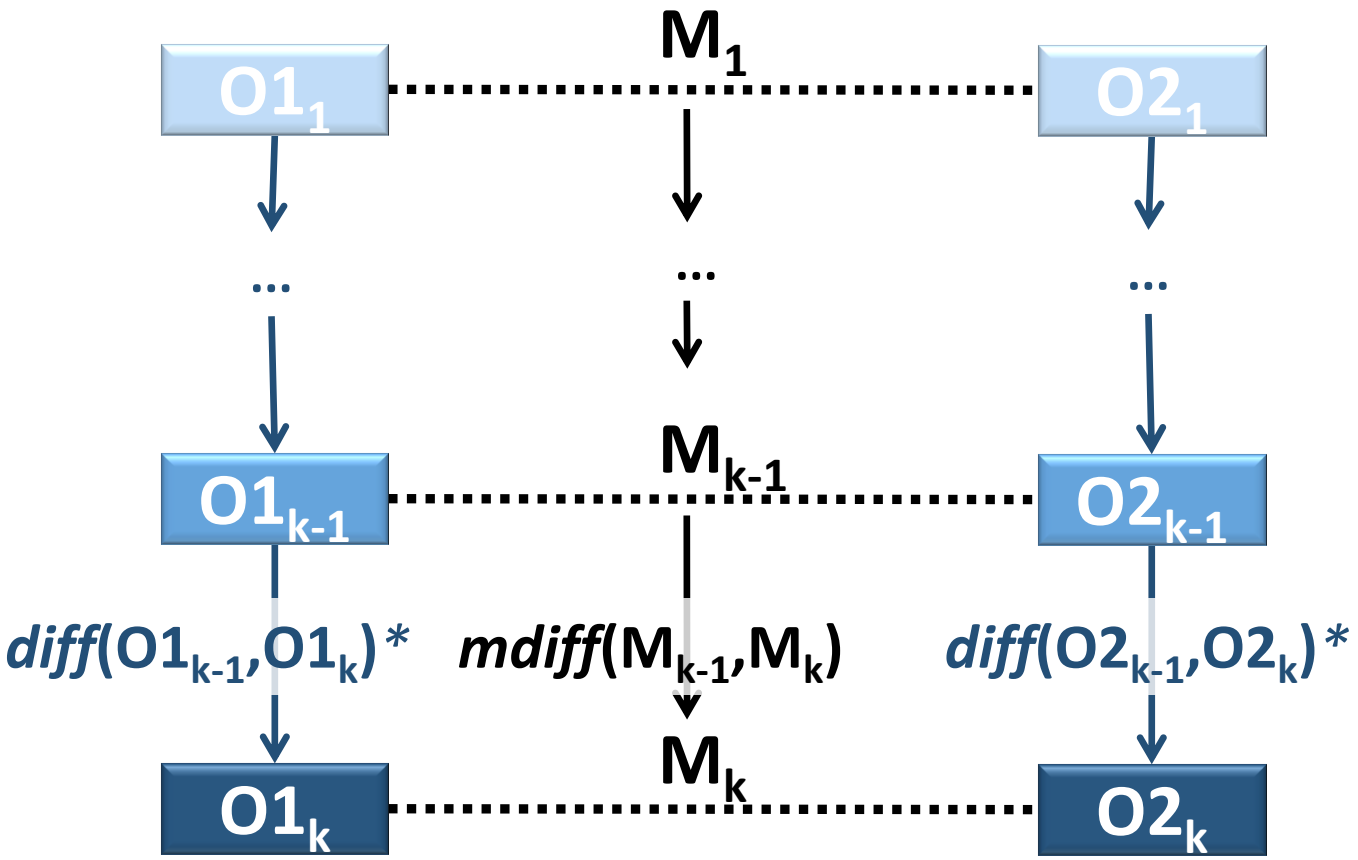
* OAEI = Ontology Alignment Evaluation Initiative

Contributions

- Investigate evolution of life science ontology mappings
- Generic model for ontology and mapping evolution and their inter-dependencies
- Evaluation for three life science scenarios



General Evolution Scheme



Current version *k*

* *Hartung, M.; Groß, A.; Rahm, E.:*

- ***COnto-Diff**: Generation of Complex Evolution Mappings for Life Science Ontologies, Journal of Biomedical Informatics, 2012.*
- ***CODEX**: Exploration of semantic changes between ontology versions, Bioinformatics 28 (6): 895-896, 2012.*

Change Operations

Ontology changes:

Extension set:

$Ext(O_{v \to v+1})$

Insert new concept, subgraph, relationship, attribute,

…

Reduction set:

$Red(O_{v \to v+1})$

Delete existing concept, subgraph, relationship, set concept to obsolete,

…

Revision set:

$Rev(O_{v \to v+1})$

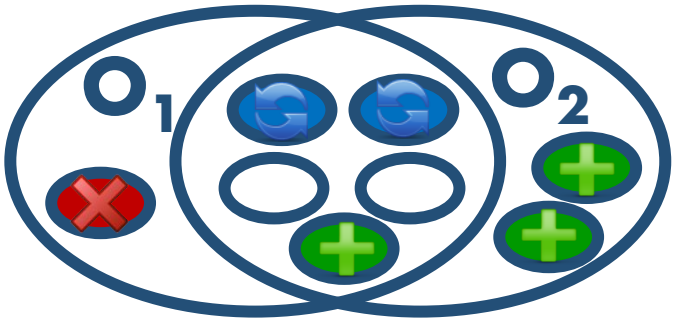
Split, merge, substitute, move concept, change attribute value,

…

Mapping changes

- **Addition set:** $Add(M_{v \to v+1}) = M_{v+1} \backslash M_v$
- **Deletion set:** $Del(M_{v \to v+1}) = M_v \backslash M_{v+1}$

Measures



6/8

Ontology Change Ratio $OCR(O_{v \to v+1})$

- Degree of ontology changes during evolution from O_v to O_{v+1}
- Fraction of concepts in $\boldsymbol{Ext} \cup \boldsymbol{Red} \cup \boldsymbol{Rev}$ versus all concepts $(O_v \cup O_{v+1})$

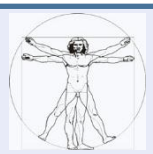
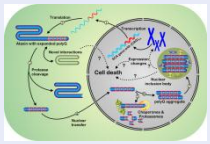
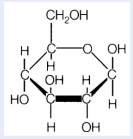
Mapping Change Ratio $MCR(M_{v \to v+1})$

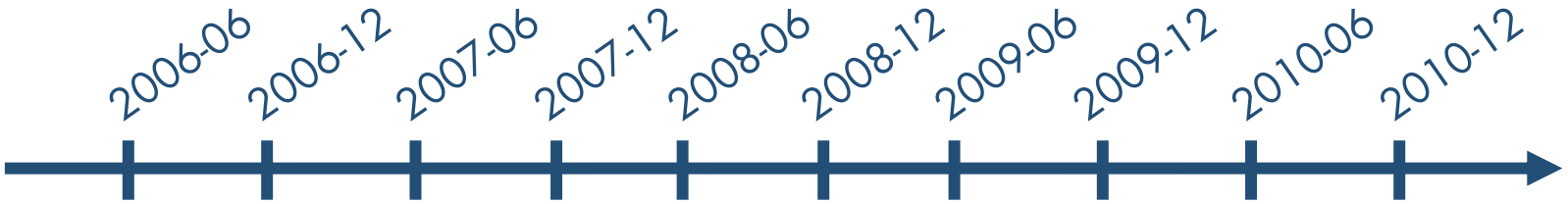
- Degree of mapping changes during evolution from M_v to M_{v+1}
- Fraction of correspondence in $\boldsymbol{Add} \cup \boldsymbol{Del}$ versus all correspondences $(M_v \cup M_{v+1})$

Impact Ratio $IR(O_{Ch}, M_{Ch})$

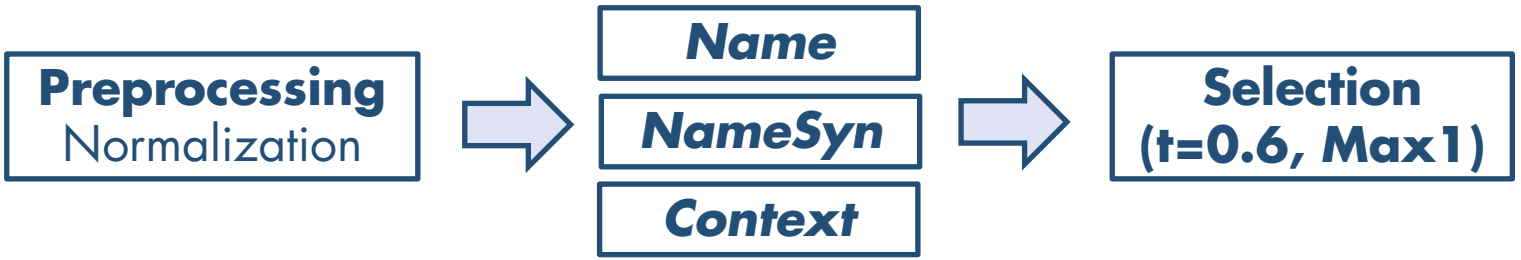
- Share of changed concepts that actually had an impact on changed correspondences
- Fraction of additive ontology changes that led to new correspondences: $IR(\boldsymbol{Ext}, \boldsymbol{Add})$

Evaluation Setup

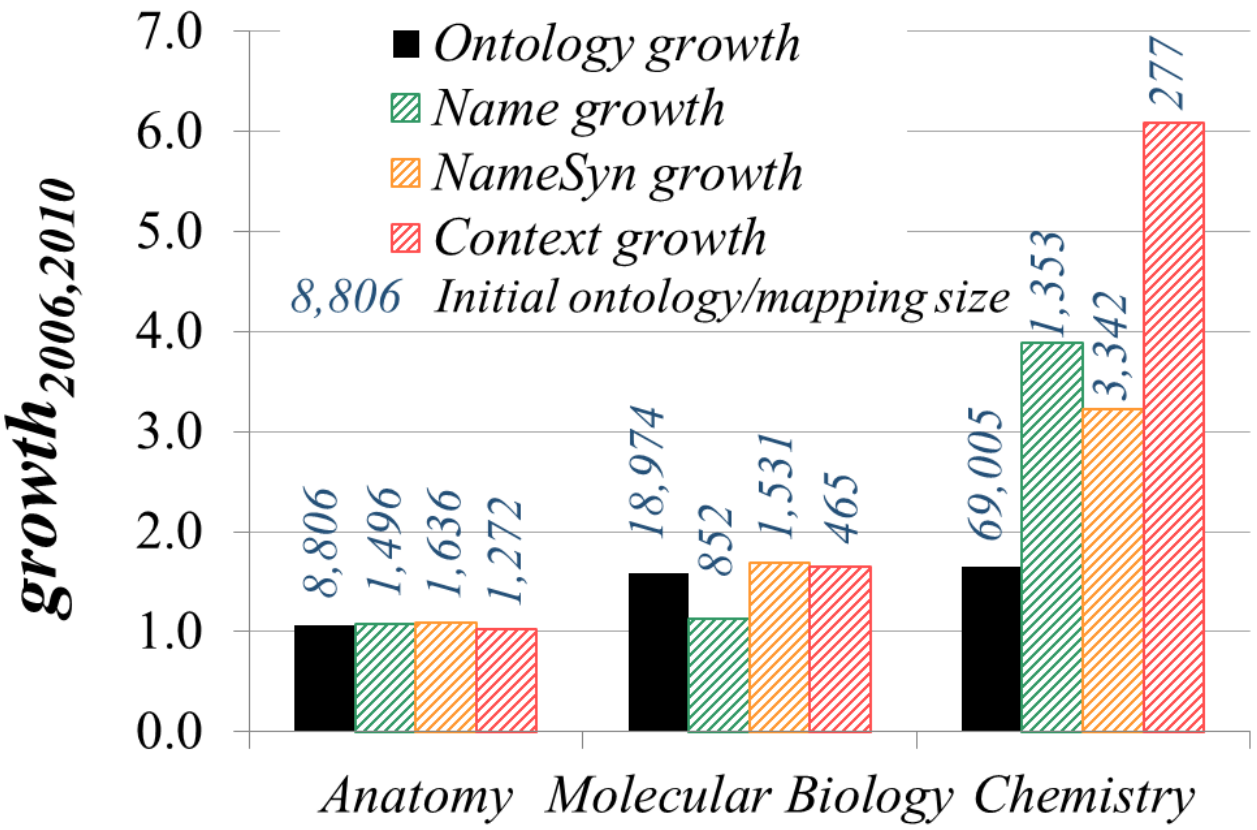
	***Anatomy***	Adult Mouse Anatomical Dictionary (MA)	NCI Thesaurus Anatomy part (NCITa)
	***Molecular Biology***	Molecular Functions (MF)	Biological Processes (BP)
	***Chemistry***	Chemical Entities of Biological Interest (ChEBI)	NCI Thesaurus (NCIT)



- Meta-data based matchers



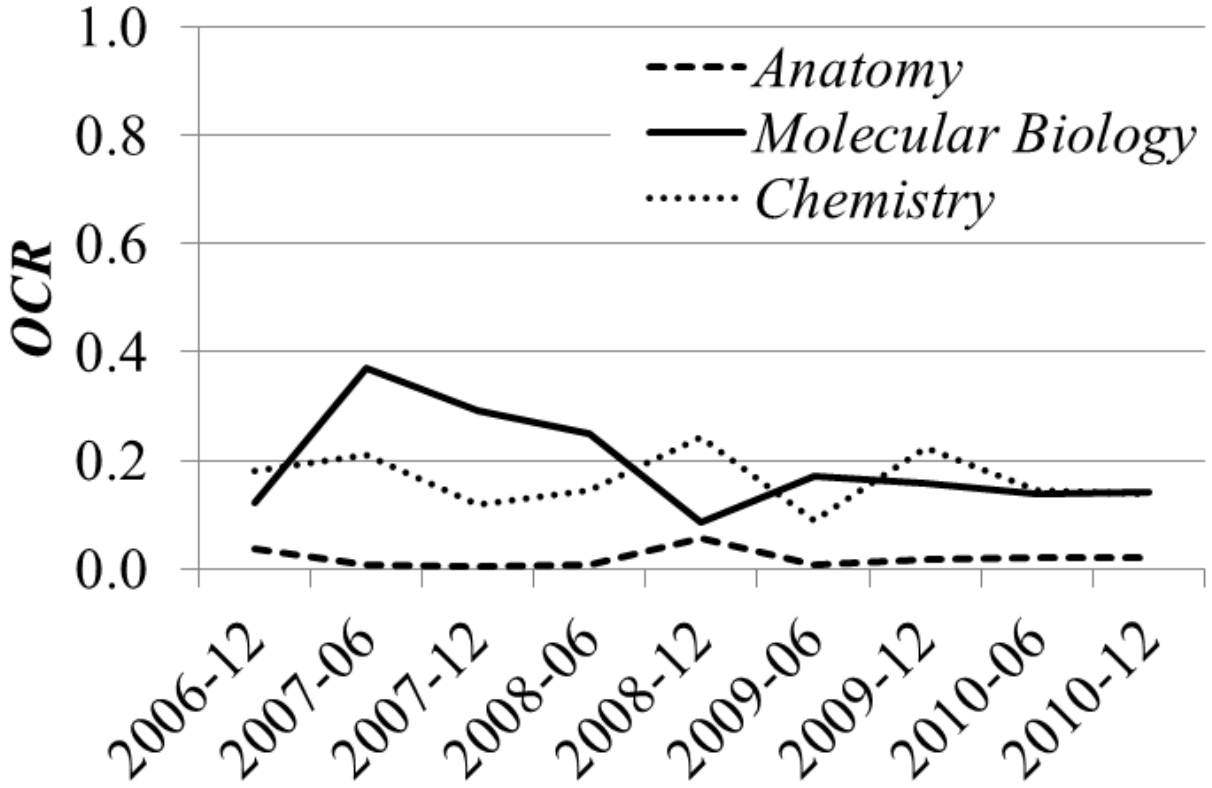
Ontology and Mapping Growth



2006 → 2010:

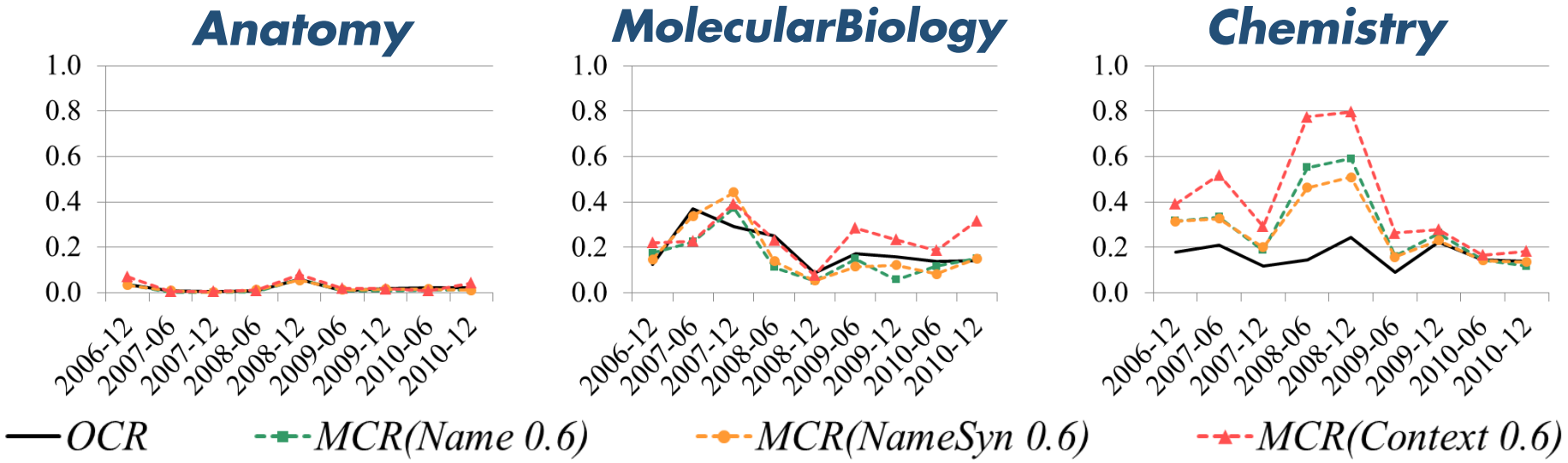
- Slight ontology growth for *Anatomy* (10%)
- 60-70% for *MolecularBiology* and *Chemistry*
- Mapping growth similar to ontology growth, except for chemistry

Ontology Change Ratio



- *Anatomy*: only few changes
- *Molecular Biology*: high change rates until 2007 (nearly 40%)
- *Chemistry*: change rates around 20%

Comparison of Change Ratios



Correlation between the ontology and mapping change factors

Domains *Anatomy*: few mapping changes, relatively stable

MolecularBiology, Chemistry: high degree of mapping changes , between 10 and 80 %

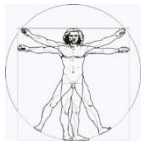
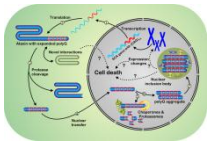
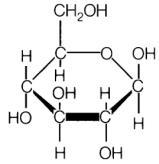
Matcher *Name*: relatively stable

Context: most heavily influenced by ontology evolution, 2008, *Chemistry* 80% mapping changes

Conclusions & Future Work

- Study the evolution of ontology mappings
  - General evolution scheme and measures (change factors)

- Evaluation for ontology mappings in three life science domains + comparison of three match strategies

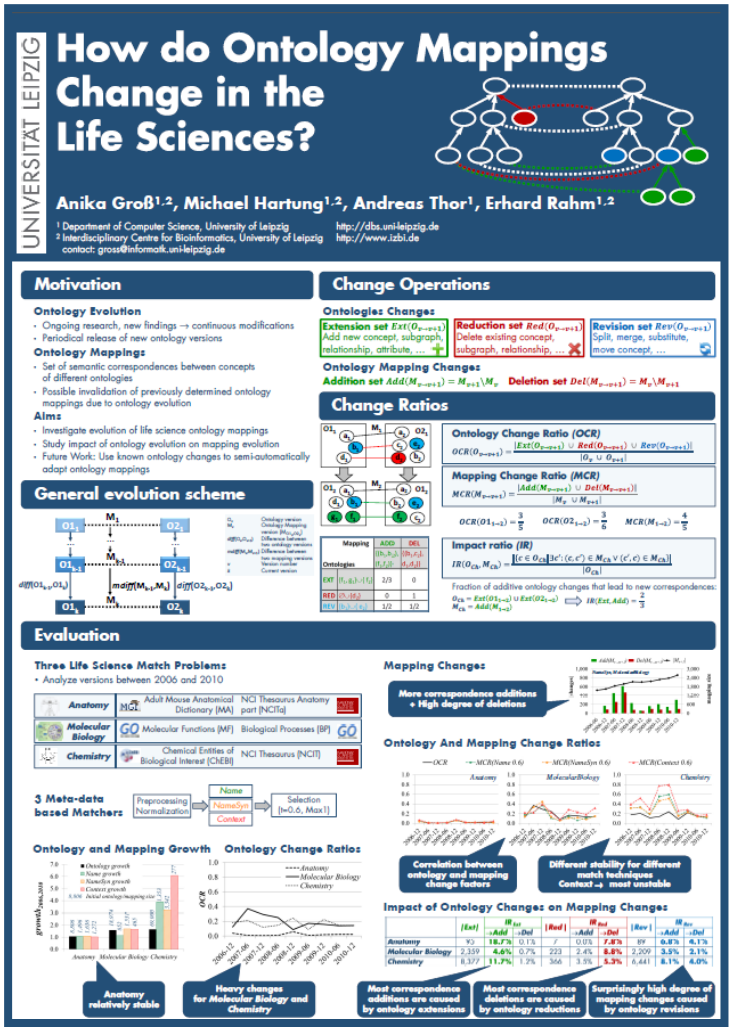
  - Correlation between ontology and mapping change factors
  - Different stability for different match techniques and domains
  - Impact of ontology on mapping changes
    - Most correspondence *Add* / *Del* are caused by ontology *Ext* / *Red*
    - Surprisingly high degree of mapping changes caused by ontology *Rev*

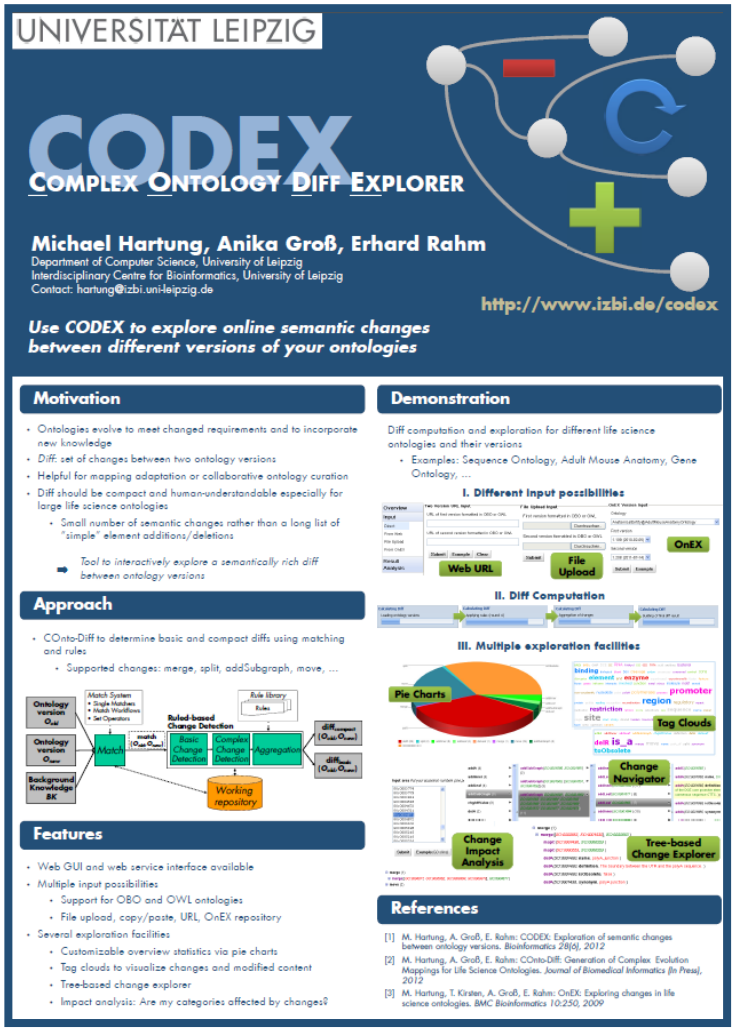
Future Work

- Use known ontology changes to semi-automatically adapt ontology mappings (without completely new mapping determination)

How do Ontology Mappings Change in the Life Sciences?

Poster Session This Afternoon





FUNDING - German Research Foundation
Grant RA497/18-1 - "EVOLUTION OF ONTOLOGIES AND MAPPINGS"