How do Ontology Mappings Change in the Life Sciences?

Anika Groß, Michael Hartung, Andreas Thor, Erhard Rahm

1 Department of Computer Science, University of Leipzig
2 Interdisciplinary Centre for Bioinformatics, University of Leipzig

http://dbs.uni-leipzig.de
http://www.izbi.de

Motivation

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

Ontology Mappings
- Set of semantic correspondences between concepts of different ontologies
- Possible invalidation of previously determined ontology mappings due to ontology evolution

Aims
- Investigate evolution of life science ontology mappings
- Study impact of ontology evolution on mapping evolution
- Future Work: Use known ontology changes to semi-automatically adapt ontology mappings

Change Operations

Ontology Changes
- Extension set Ext(O\[\text{new}\])
  - Add new concept, subgraph, relationship, attribute, ...
- Reduction set Red(O\[\text{old}\])
  - Delete existing concept, subgraph, relationship, ...
- Revision set Rev(O\[\text{old}\])
  - Split, merge, substitute, move concept, ...

Ontology Mapping Changes
- Addition set Add(M\[\text{new}\]) = M\[\text{new}\] \setminus M\[\text{old}\]
- Deletion set Del(M\[\text{old}\]) = M\[\text{old}\] \setminus M\[\text{new}\]

General evolution scheme

Ontology and Mapping Growth

<table>
<thead>
<tr>
<th>Anatomy</th>
<th>Molecular Biology</th>
<th>Chemistry</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.0</td>
<td>6.0</td>
<td>7.0</td>
</tr>
<tr>
<td>6.7</td>
<td>6.6</td>
<td>6.7</td>
</tr>
<tr>
<td>6.5</td>
<td>6.4</td>
<td>6.5</td>
</tr>
<tr>
<td>6.3</td>
<td>6.2</td>
<td>6.3</td>
</tr>
<tr>
<td>6.1</td>
<td>6.0</td>
<td>6.1</td>
</tr>
</tbody>
</table>

Change Ratios

Ontology Change Ratio (OCR)

OCR(\text{O}_{\text{new}}) = \frac{|\text{Ext}(|O_{\text{new}}|) \cup |\text{Red}(|O_{\text{old}}|) \cup |\text{Rev}(|O_{\text{old}}|)|}{|O_{\text{old}}|}

Mapping Change Ratio (MCR)

MCR(\text{M}_{\text{new}}, \text{M}_{\text{old}}) = \frac{|\text{Add}(\text{M}_{\text{new}}) \cup |\text{Del}(\text{M}_{\text{old}})|}{|\text{M}_{\text{old}}|}

Impact Ratio (IR)

IR(\text{M}_{\text{new}}, \text{M}_{\text{old}}) = \frac{|\text{Add}(\text{M}_{\text{new}}) \cup |\text{Del}(\text{M}_{\text{old}})|}{|\text{M}_{\text{old}}|}

Evaluation

Three Life Science Match Problems
- Analyze versions between 2006 and 2010

3 Meta-data based Matchers

Preprocessing
- Normalization

Selection
- [0.6, Max 1]

Ontology and Mapping Change Ratios

Correlation between ontology and mapping change factors
- Different stability for different match techniques
- Context → most unstable

Impact of Ontology Changes on Mapping Changes

<table>
<thead>
<tr>
<th>Anatomical Change</th>
<th>OCR</th>
<th>MCR(Name + 0.6)</th>
<th>MCR(NameSync + 0.6)</th>
<th>MCR(Context + 0.6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anatomy</td>
<td>95</td>
<td>18.7%</td>
<td>0.1%</td>
<td>7.8%</td>
</tr>
<tr>
<td>Molecular Biology</td>
<td>259</td>
<td>8.8%</td>
<td>2.4%</td>
<td>3.8%</td>
</tr>
<tr>
<td>Chemistry</td>
<td>375</td>
<td>11.2%</td>
<td>1.2%</td>
<td>5.3%</td>
</tr>
</tbody>
</table>

Mapping Changes

More correspondence additions + High degree of deletions

Ontology and Mapping Change Ratios

Correlation between ontology and mapping change factors
- Different stability for different match techniques
- Context → most unstable

Impact of Ontology Changes on Mapping Changes

<table>
<thead>
<tr>
<th>Anatomical Change</th>
<th>OCR</th>
<th>MCR(Name + 0.6)</th>
<th>MCR(NameSync + 0.6)</th>
<th>MCR(Context + 0.6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anatomy</td>
<td>95</td>
<td>18.7%</td>
<td>0.1%</td>
<td>7.8%</td>
</tr>
<tr>
<td>Molecular Biology</td>
<td>259</td>
<td>8.8%</td>
<td>2.4%</td>
<td>3.8%</td>
</tr>
<tr>
<td>Chemistry</td>
<td>375</td>
<td>11.2%</td>
<td>1.2%</td>
<td>5.3%</td>
</tr>
</tbody>
</table>

Ontology and Mapping Growth

<table>
<thead>
<tr>
<th>Anatomical Change</th>
<th>OCR</th>
<th>MCR(Name + 0.6)</th>
<th>MCR(NameSync + 0.6)</th>
<th>MCR(Context + 0.6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anatomy</td>
<td>95</td>
<td>18.7%</td>
<td>0.1%</td>
<td>7.8%</td>
</tr>
<tr>
<td>Molecular Biology</td>
<td>259</td>
<td>8.8%</td>
<td>2.4%</td>
<td>3.8%</td>
</tr>
<tr>
<td>Chemistry</td>
<td>375</td>
<td>11.2%</td>
<td>1.2%</td>
<td>5.3%</td>
</tr>
</tbody>
</table>

Anatomy relatively stable

Heavy changes for Molecular Biology and Chemistry

Most correspondence additions are caused by ontology extensions

Most correspondence deletions are caused by ontology reductions

Surprisingly high degree of mapping changes caused by ontology revisions