

SMW-LDE

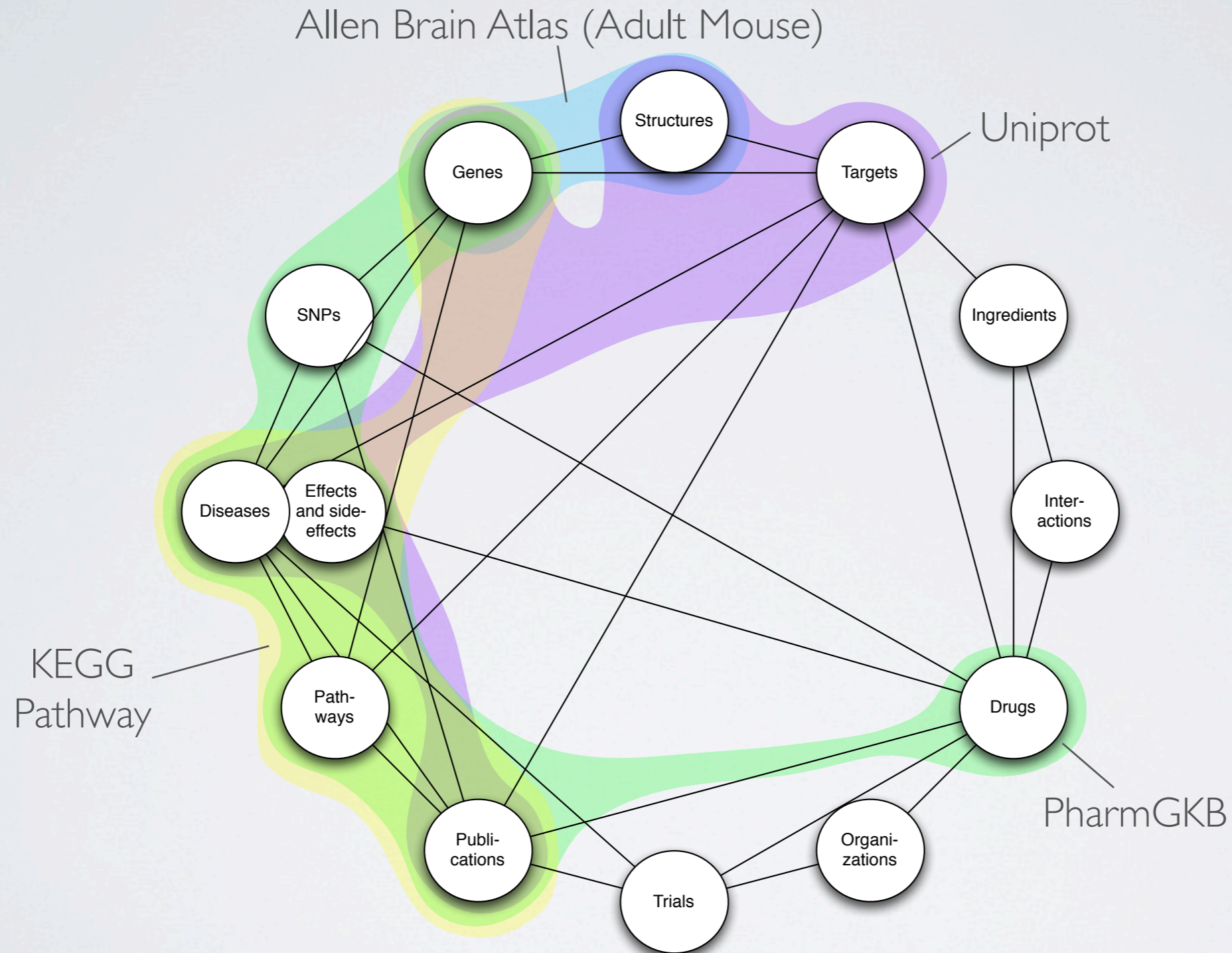
SEMANTIC MEDIAWIKI LINKED DATA EXTENSION

2010-2011

PROJECT OUTLINE

- Connect Allen Brain Atlas Data to public data sets
 - Build upon existing efforts such as LODD and Bio2RDF, where possible
 - Improve usefulness of the Allen Brain Atlas through better integration of other data sources
- Develop an Open Source Semantic Technology Stack
 - Import, Transformation, Identity Resolution, Reconciliation
 - Tightly integrated into SMW+

CURRENT DATASETS



not shown: KEGG GENES

CURRENT DATASETS

- Allen Brain Atlas (Mouse Brain)
 - Imported via GRDDL
- KEGG Pathway and GENES
 - Manually extracted using D2R Server
 - http://wiking.vulcan.com/neurobase/kegg_pathway/
 - http://wiking.vulcan.com/neurobase/kegg_genes/
- PharmGKB
 - via Chem2Bio2RDF: <http://cheminfov.informatics.indiana.edu:8080/pharmgkb/>
- Uniprot
 - Provides full RDF dump via FTP

VISUALIZATION

- The prototype includes templates for genes, diseases and pathways
- Each article includes lists detailing the topic's relations to other classes; the list indicates the source of the data
- The user is expected to navigate along these paths to discover new insights

GENE PAGE (I)

Go

Search

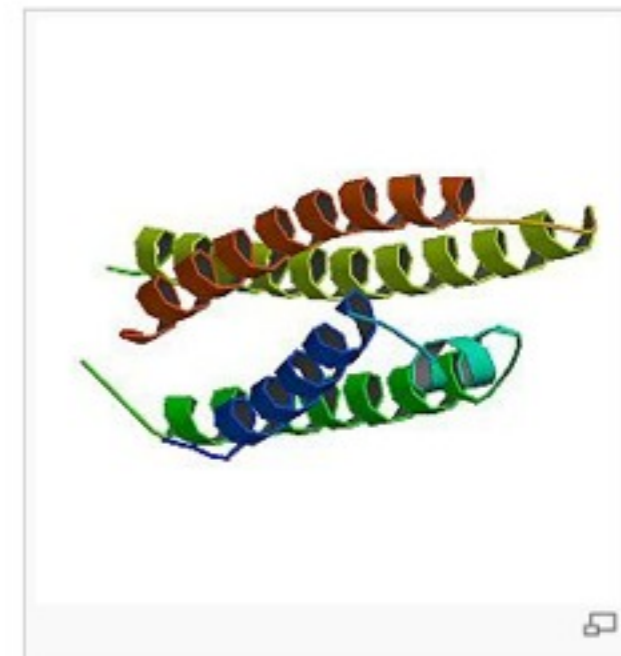
[Home](#) [Project management](#) [Software engineering](#) [Product management](#) [My dashboard](#) [Articles and data](#)[Neurobase Use Case - Gene](#)[Page](#) [Discussion](#)[History](#)[More](#)

Presenilin-1 is a **protein** that in humans is encoded by the *PSEN1* **gene**. Alzheimer's disease (AD) patients with an inherited form of the disease carry mutations in the presenilin proteins (PSEN1; PSEN2) or in the **amyloid precursor protein (APP)**. These disease-linked mutations result in increased production of the longer form of amyloid beta (main component of amyloid deposits found in AD brains). Presenilins are postulated to regulate APP processing through their effects on **gamma secretase**, an **enzyme** that cleaves APP. Also, it is thought that the presenilins are involved in the cleavage of the Notch receptor, such that they either directly regulate gamma secretase activity or themselves are protease enzymes. Multiple alternatively spliced **transcript** variants have been identified for this gene, the full-length natures of only some have been determined.[2]

Contents

[\[hide\]](#)

- [1 Function](#)
- [2 Gene](#)
- [3 Epistatic impact of APOE](#)
- [4 Alzheimer's Disease](#)
- [5 Interactive pathway map](#)
- [6 ***DATA***](#)
 - [6.1 List of Pathways \(and their genes\)](#)
 - [6.2 Related Drugs and Diseases](#)
 - [6.3 Expression Levels of this Gene in different Anatomic Structures](#)
 - [6.4 Publications about this Gene](#)



GENE PAGE (II)

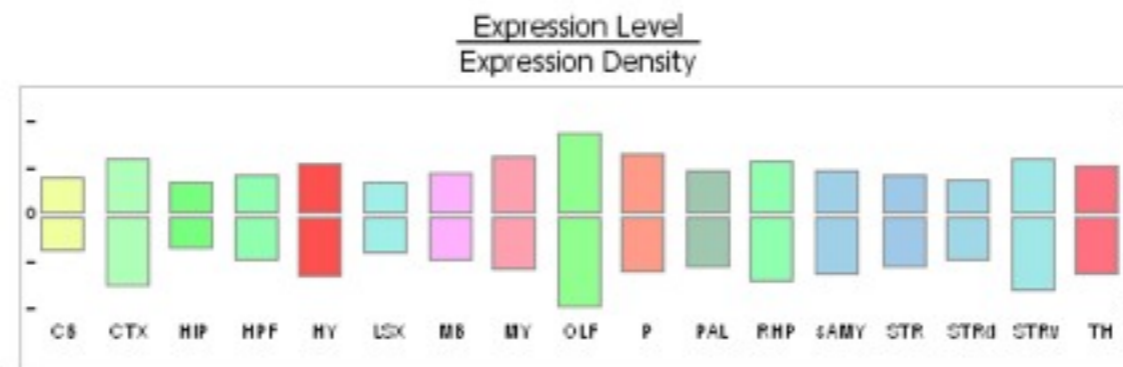
List of Pathways (and their genes)

| | PSEN1 | PSENEN | APH1A | APH2 |
|--------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Wnt signaling pathway | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| Notch signaling pathway | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> |
| Neurotrophin signaling pathway | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input checked="" type="checkbox"/> |
| Alzheimer's disease | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input checked="" type="checkbox"/> |

Related Drugs and Diseases

| | Donepezil hydrochloride | Galantamine | Rivastigmine | Memantine |
|-----------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Alzheimer Disease | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> |
| Cardiomyopathy, Dilated | <input type="checkbox"/> | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| Spondylothoracic Dysostosis | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Expression Levels of this Gene in different Anatomic Structures

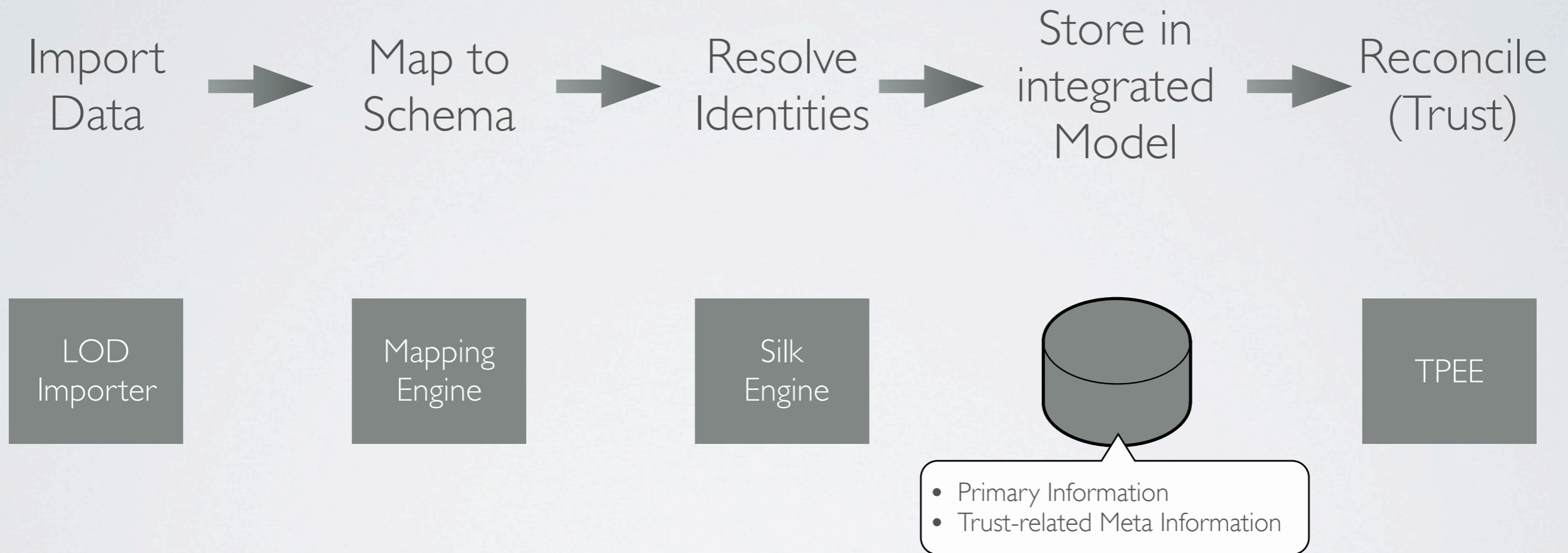


from AIBS data of the mouse brain:

Publications about this Gene

1. Cummings JL, Frank JC, Cherry D, Kohatsu ND, Kemp B, Hewett L, Mittman B (2002). "Guidelines for managing Alzheimer's disease: Part I. Assessment". *American Family Physician* 65 (11): 2263–2272. PMID 12074525. <http://www.aafp.org/afp/20020601/2263.html>.
2. Cummings JL, Frank JC, Cherry D, Kohatsu ND, Kemp B, Hewett L, Mittman B (2002). "Guidelines for managing Alzheimer's disease: Part II.

PROCESS



LOD IMPORTER: EXAMPLE

- An HTTP Crawl of brain-map.org results in the following RDF obtained from <http://mouse.brain-map.org/brain/Ace.xml> via GRDDL:

```

smw:TemporaryGraph_brain-map_05/20/2010-12:30:45 {
  <http://brain-map.org/mouse/brain/Ace.xml>    rdf:type          brain:gene .
  <http://brain-map.org/mouse/brain/Ace.xml>    brain:chromosome   "11" .
  <http://brain-map.org/mouse/brain/Ace.xml>    brain:entrezgeneid "11421" .
  <http://brain-map.org/mouse/brain/Ace.xml>    brain:gene-expressions _:exprs .
  _:exprs                                       brain:gene-expression _:e .
  _:e                                           rdf:type           brain:gene-expression .
  _:e                                           brain:avgdensity   "3.5120085477829" .
  _:e                                           brain:avglevel     "4.00683224201202" .
  _:e                                           brain:geneid       "11210" .
  _:e                                           brain:projectcode  "0310" .
  _:e                                           brain:rgb          "#b0ffb8" .
  _:e                                           brain:structureid  "85" .
  _:e                                           brain:structurelabel "CTX" .
  _:e                                           brain:structurename "Cerebral cortex" .
}

```

MAPPING ENGINE: EXAMPLE

```

smw:ImportGraph_brain-map_05/20/2010-12:30:45 {
  <http://brain-map.org/mouse/brain/Ace.xml> rdf:type          brain:gene .
  <http://brain-map.org/mouse/brain/Ace.xml> brain:entrezgeneid  "11421" .
  <http://brain-map.org/mouse/brain/Ace.xml> brain:gene-expressions  _:exprs .
  _:exprs brain:gene-expression  _:e .
  _:e brain:avgdensity "3.5120085477829" .
  _:e brain:avglevel "4.00683224201202" .
  _:e brain:structureid "85" .
  _:e brain:structurelabel "CTX" .
  _:e brain:structurename "Cerebral cortex" .
}
    
```



```

smw:ImportGraph_brain-map_05/20/2010-12:30:45 {
  <http://brain-map.org/mouse/brain/Ace.xml> rdf:type          smw:Gene .
  <http://brain-map.org/mouse/brain/Ace.xml> smw:EntrezId      "11421" .
  <http://brain-map.org/mouse/brain/Ace.xml> smw:hasGeneExpression  _:e .
  _:e rdf:type          smw:GeneExpression .
  _:e smw:ExprAvgDensity "3.5120085477829" .
  _:e smw:ExprAvgLevel  "4.00683224201202" .
  _:e smw:ABARegion     _:r .
  _:r rdf:type          smw:Region .
  _:r smw:ABAStructureID "85" .
  _:r smw:ABAStructureLabel "CTX" .
  _:r smw:ABAStructureName "Cerebral cortex" .
}
    
```

MAPPING ENGINE: EXAMPLE

MAIN ELEMENTS OF THE R2R MAPPING LANGUAGE:

- **Class Mappings** define the correspondences of a class from the target vocabulary with the source vocabulary.
- **Property Mappings** define correspondences of a property from the target vocabulary with the source vocabulary.
- **Transformations** define how target values are transformed, for example using string or numerical operators.
- **Target Patterns and Modifiers** define the target structure and vocabulary to which is mapped

```
mp:BrainGeneToSmwGene
  a r2r:ClassMapping ;
  r2r:prefixDefinitions "brain: <http://brain-map.org/gene/0.1#> . smw: <http://www.example.org/smw/>";
  r2r:sourcePattern "?SUBJ a brain:gene" ;
  r2r:targetPattern "?SUBJ a smw:Gene" ;
  .

mp:BrainChromosomeToSmwChromosome
  a r2r:PropertyMapping ;
  r2r:classMappingRef mp:BrainGeneToSmwGene ;
  r2r:sourcePattern "?SUBJ brain:chromosome ?o" ;
  r2r:targetPattern "?SUBJ smw:ChromosomeNo ?o" ;
  .

(...)
```

Exemplary mapping script

SILK ENGINE: EXAMPLE

```

smw:ImportGraph_brain-map_05/20/2010-12:30:45 {
  <http://brain-map.org/mouse/brain/Ace.xml>      rdf:type          smw:Gene .
  <http://brain-map.org/mouse/brain/Ace.xml>      smw:EntrezId      "11421" .
  <http://brain-map.org/mouse/brain/Ace.xml>      smw:hasGeneExpression
  _:e                                             rdf:type          smw:GeneExpression .
  _:e                                             smw:ExprAvgDensity
  _:e                                             smw:ExprAvgLevel  "3.5|20085477829" .
  _:e                                             smw:ABARegion     "4.0068322420|202" .
  _:r                                             rdf:type          smw:Region .
  _:r                                             smw:ABAStructureID
  _:r                                             smw:ABAStructureLabel
  _:r                                             smw:ABAStructureName
  "85" .
  "CTX" .
  "Cerebral cortex" .
}

```



```

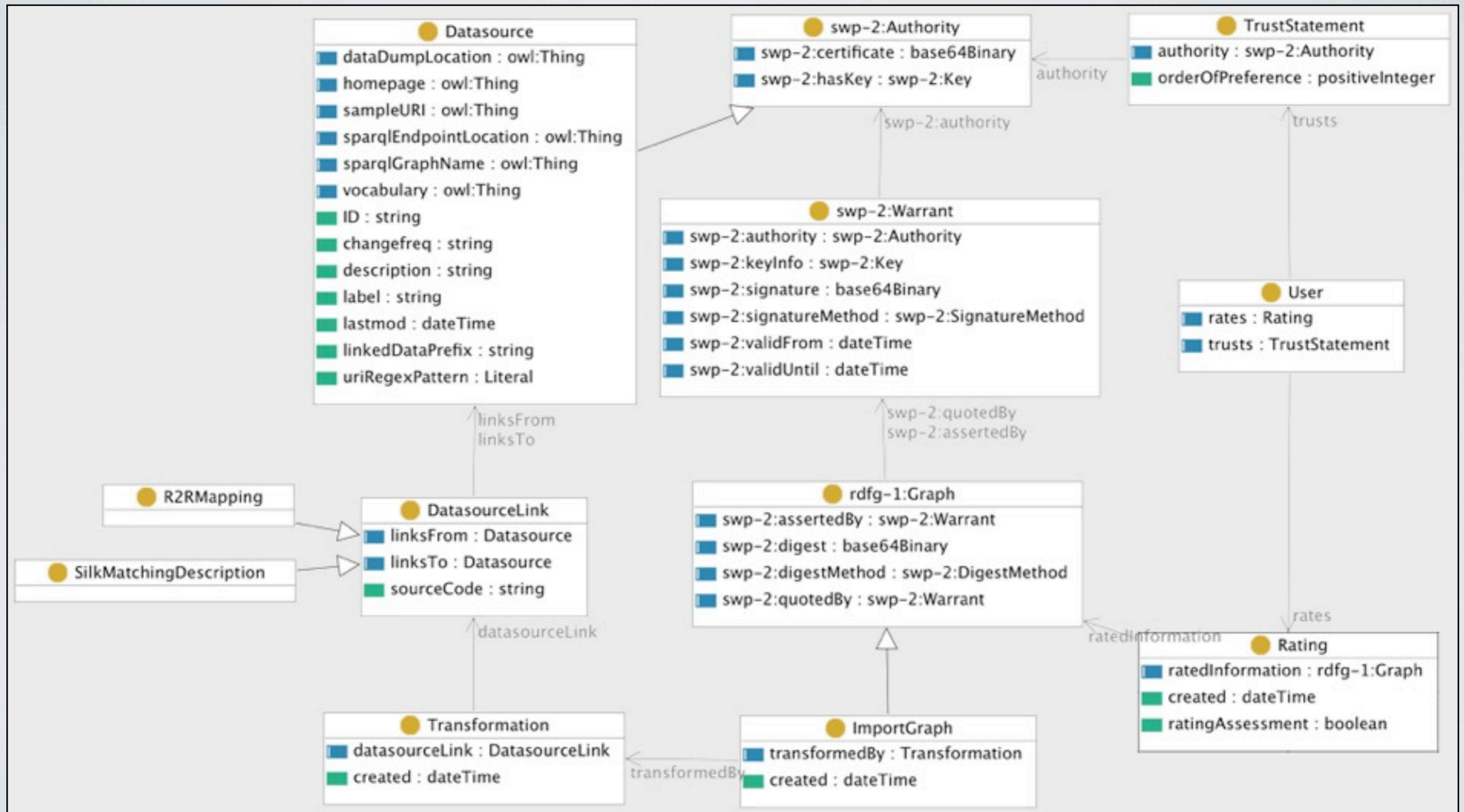
smw:ImportGraph_brain-map_05/20/2010-12:30:45 {
  smwGenes:ACE      rdf:type          smw:Gene .
  smwGenes:ACE      smw:EntrezId      "11421" .
  smwGenes:ACE      smw:hasGeneExpression
  smw:GeneExpression29837|
  smw:GeneExpression29837|
  smw:GeneExpression29837|
  smw:GeneExpression29837|
  smwRegions:CTX   rdf:type          smw:GeneExpression29837| .
  smwRegions:CTX   smw:GeneExpression
  smwRegions:CTX   smw:ExprAvgDensity
  smwRegions:CTX   smw:ExprAvgLevel  "3.5|20085477829" .
  smwGenes:ACE     smw:ABARegion     "4.0068322420|202" .
  smw:GeneExpression29837|
  smwRegions:CTX   smw:ABARegion     smwRegions:CTX .
  smwRegions:CTX   smw:ABAStructureID
  smwRegions:CTX   smw:ABAStructureLabel
  smwRegions:CTX   smw:ABAStructureName
  smwGenes:ACE     "85" .
  smw:GeneExpression29837|
  smwRegions:CTX   "CTX" .
  smwRegions:CTX   "Cerebral cortex" .
  smwGenes:ACE     <http://brain-map.org/mouse/brain/Ace.xml> .
  smw:GeneExpression29837|
  smwRegions:CTX   _:e .
  smwRegions:CTX   _:r .
}

```

SILK ENGINE: EXAMPLE

```
<Silk>
  (...)
  <Prefix id="smw" namespace="http://www.example.org/smw/" />
  <Prefix id="brain" namespace="http://brain-map.org/gene/0.1#" />
  <DataSource id="smw">
    <EndpointURI>http://localhost/sparql</EndpointURI>
    <DoCache>1</DoCache>
    <RetryCount>100</RetryCount>
    <RetryTime>5000</RetryTime>
  </DataSource>
  <DataSource id="brain-map">
    <EndpointURI>http://localhost/sparql</EndpointURI>
    <Graph>http://smw/ImportGraph_brain-map_05/20/2010-12:30:45</Graph>
    <DoCache>1</DoCache>
    <RetryCount>100</RetryCount>
    <RetryTime>5000</RetryTime>
  </DataSource>
  <LinkSpec id="gene">
    <LinkType>owl:sameAs</LinkType>
    <SourceDataset dataSource="smw" var="a">
      <RestrictTo>?a rdf:type smw:Gene</RestrictTo>
    </SourceDataset>
    <TargetDataset dataSource="brain-map" var="b">
      <RestrictTo>?b rdf:type brain:gene</RestrictTo>
    </TargetDataset>
    <LinkCondition>
      <MAX>
        <Compare metric="stringEquality">
          <Param name="str1" path="?a/smw:EntrezId" />
          <Param name="str2" path="?b/brain:entrezgeneid" />
        </Compare>
      </MAX>
    </LinkCondition>
    <Thresholds accept="0.99" verify="0.9" />
    <LinkLimit max="1" method="metric_value" />
    <Output acceptedLinks="brain_map_smw_accepted_links.n3" verifyLinks="brain_map_smw_verify_links.n3" mode="truncate" />
  </LinkSpec>
</Silk>
```

METADATA MODEL



METADATA MODEL: PROVENANCE

smwGraphs:ProvenanceGraph

Updated by LODImporter as well as by SMW+

```

smwGraphs:ProvenanceGraph {
  smwGraphs:ImportGraph_brain-map_05/20/2010-12:30:45
    a smw-lde:ImportGraph ;
    swp:assertedBy [
      a swp:Warrant ;
      swp:authority smwDatasources:BrainAtlas;
    ];
    sme-lde:transformedBy [
      a smw-lde:Transformation ;
      smw-lde:datasourceLink smwDatasourceLinks:brain_atlas_r2r_mapping ;
      smw-lde:created "2010-05-19T08:33:19"^^xsd:dateTime .
    ];
    smw-lde:transformedBy [
      a smw-lde:Transformation ;
      smw-lde:datasourceLink smwDatasourceLinks:brain_atlas_silk_matching_description ;
      smw-lde:created "2010-05-19T08:33:30"^^xsd:dateTime .
    ];
  smwGraphs:WikiGraph_Genes_ACE_826
    a swp:Graph;
    swp:assertedBy [
      a swp:Warrant ;
      swp:authority smwDatasources:Wiki ;
    ];
    smw-lde:created "2010-05-19T08:34:19"^^xsd:dateTime ;
    smw-lde:revisionNo 826 ;
    smw-lde:lastChangedBy smwusers:ChrisBizer .
  smw:RatingTripleGraph_7164
    a swp:Graph ;
    smw-lde:created "2010-05-19T08:40:38"^^xsd:dateTime ;
    swp:assertedBy [
      a swp:Warrant ;
      swp:authority smw-lde:RatingTool ;
    ];
}

```

REFERENCES

- R2R: <http://www4.wiwiss.fu-berlin.de/bizer/r2r/>
- Silk: <http://www4.wiwiss.fu-berlin.de/bizer/silk/>