U LISBOA | UNIVERSIDADE DE LISBOA

Faculdade de Ciências
UNIVERSIDADE DE LISBOA

Towards visualizing the alignment of large biomedical ontologies

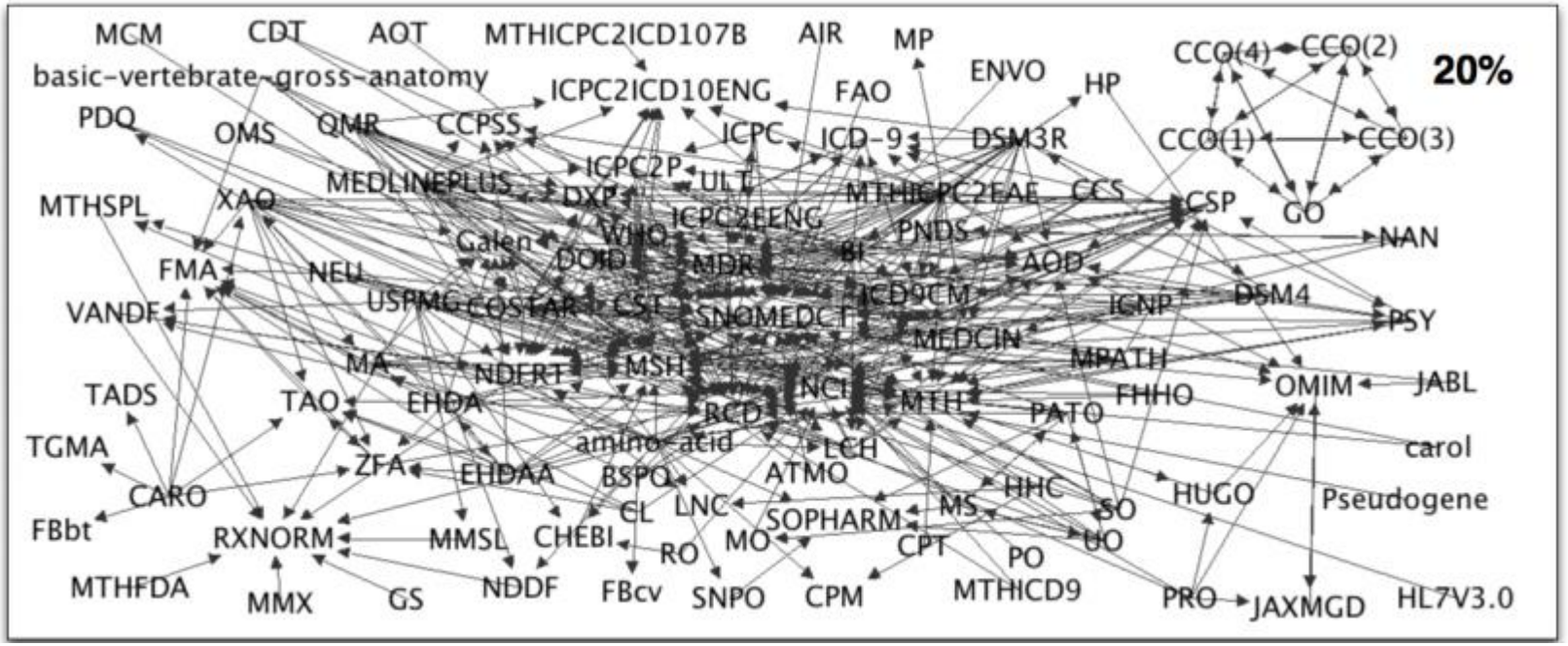
Catia Pesquita[1], Daniel Faria[1], Emanuel Santos[1],
Jean-Marc Neefs[2], Francisco M. Couto[1]

[1]LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2]Janssen Pharmaceutical Companies of Johnson & Johnson

cpesquita@di.fc.ul.pt

DILS 2014

Mappings between Biomedical Ontologies

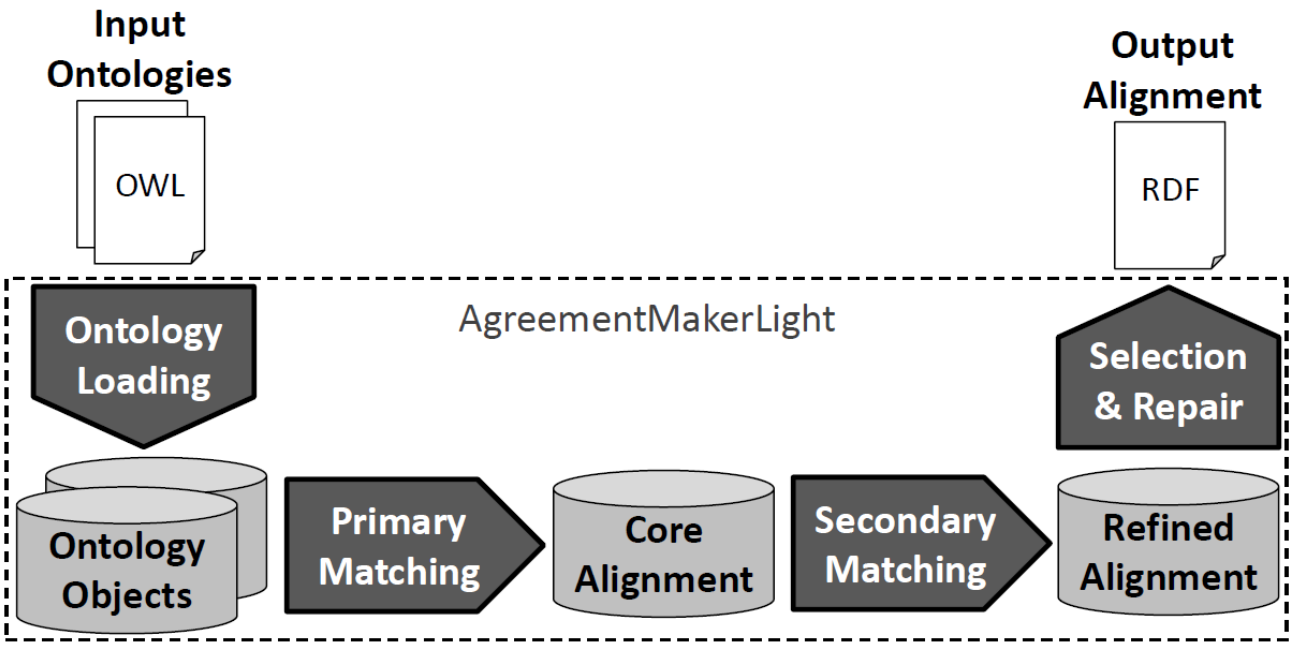


373 ontologies in BioPortal and about 13 million simple string mappings.

Image taken from: Ghazvinian, Amir, et al. *What four million mappings can tell you about two hundred ontologies*. Springer Berlin Heidelberg, 2009.

AgreementMakerLight (AML)

- Scalable and efficient
- Sophisticated string similarity
- Background knowledge use
- Logical consistency repair



AML @ OAEI 2013

System	Total Time (s)	Average			
		Prec.	F-m.	Rec.	Inc. Degree
YAM++	2,066	0.94	0.82	0.73	14%
AML-BK	1,814	0.91	0.79	0.71	44%
LogMap-BK	2,391	0.90	0.78	0.70	*0%*
AML-BK-R	1,987	0.92	0.78	0.69	*0%*
AML	1,681	0.93	0.78	0.68	43%
LogMap	2,485	0.91	0.78	0.69	*0%*
AML-R	1,821	0.94	0.78	0.67	*0%*
ServOMap	15,300	0.87	0.77	0.69	52%
$GOMMA_{2012}$	1,920	0.81	0.65	0.57	29%
LogMapLt	371	0.87	0.60	0.52	34%
SPHeRe	42,040	0.86	0.57	0.46	21%
IAMA	704	0.91	0.52	0.39	46%

Large biomedical

AML @ OAEI 2013

System	Total Time (s)	Average			
		Prec.	F-m.	Rec.	Inc. Degree
YAM++	2,066	0.94	0.82	0.73	14%
AML-BK	1,814	0.91	0.79	0.71	44%
LogMap-BK	2,391	0.90	0.78	0.70	*0%*
AML-BK-R	1,987	0.92	0.78	0.69	*0%*
AML					
LogMap					
AML-R					
ServOMap					
$GOMMA_{2012}$					
LogMapLt					
SPHeRe					
IAMA					

Large biomedic

Anatomy

Matcher	Runtime	Size	Precision	F-measure	Recall	Recall+	Coherent
AML-bk	43	1477	0.95	0.94	0.93	0.82	√
GOMMA-bk*	11	1534	0.92	0.92	0.93	0.81	-
YAM++	62**	1395	0.94	0.90	0.87	0.66	-
AML	15	1315	0.95	0.89	0.83	0.54	√
LogMap	13	1398	0.92	0.88	0.85	0.59	√
GOMMA*	9	1264	0.96	0.87	0.80	0.47	-
StringsAuto	1444	1314	0.90	0.83	0.78	0.43	-
LogMapLite	7	1148	0.96	0.83	0.73	0.29	-
MapSSS	2040	1296	0.90	0.83	0.77	0.44	-
ODGOMS	1212	1102	0.98	0.82	0.71	0.24	-
WikiMatch	19366	1027	0.99	0.80	0.67	0.15	-
HotMatch	300	989	0.98	0.77	0.64	0.14	-
StringEquiv	-	946	*1.00*	0.77	0.62	0.00	-
XMapSig	393	1192	0.86	0.75	0.67	0.13	-
ServOMap	43	975	0.96	0.75	0.62	0.10	-
XMapGen	403	1304	0.81	0.75	0.69	0.19	-

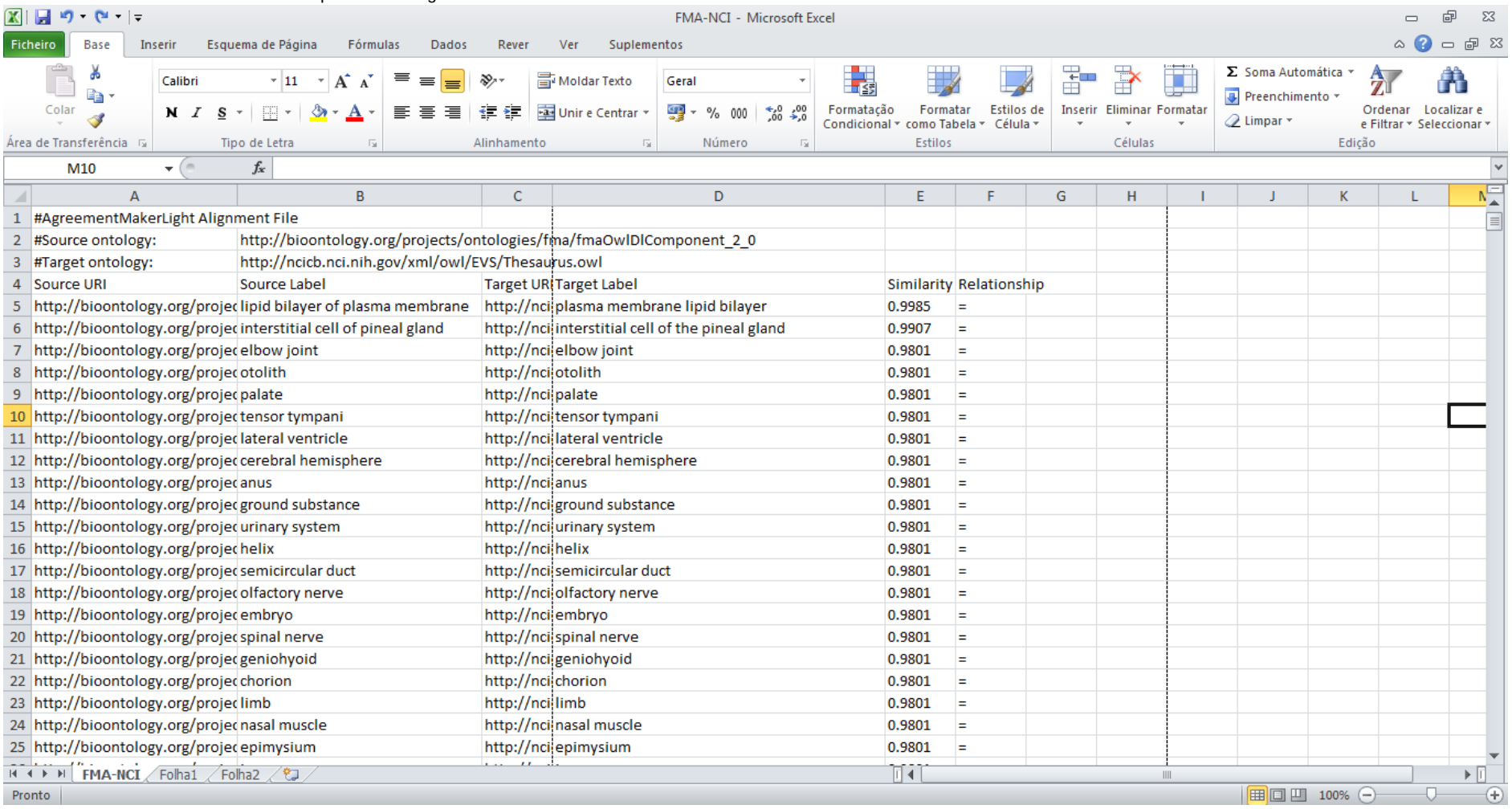
AML output

```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF xmlns="http://knowledgeweb.semanticweb.org/heterogeneity/alignment"
                xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
                xmlns:xsd="http://www.w3.org/2001/XMLSchema#">
<Alignment>
<xml>yes</xml>
<level>0</level>
<type>??</type>
<map>
                <Cell>
                                <entity1 rdf:resource="http://mouse.owl#MA_0002401"/>
                                <entity2 rdf:resource="http://human.owl#NCI_C52561"/>
                                <measure rdf:datatype="xsd:float">1.0</measure>
                                <relation>=</relation>
                </Cell>
</map>
<map>
                <Cell>
                                <entity1 rdf:resource="http://mouse.owl#MA_0000270"/>
                                <entity2 rdf:resource="http://human.owl#NCI_C33736"/>
                                <measure rdf:datatype="xsd:float">1.0</measure>
                                <relation>=</relation>
                </Cell>
</map>
<map>
                <Cell>
                                <entity1 rdf:resource="http://mouse.owl#MA_0001951"/>
                                <entity2 rdf:resource="http://human.owl#NCI_C12715"/>
                                <measure rdf:datatype="xsd:float">1.0</measure>
                                <relation>=</relation>
                </Cell>
</map>
<map>
                <Cell>
                                <entity1 rdf:resource="http://mouse.owl#MA_0002303"/>
                                <entity2 rdf:resource="http://human.owl#NCI_C52701"/>
                                <measure rdf:datatype="xsd:float">1.0</measure>
                                <relation>=</relation>
                </Cell>
</map>
```

U LISBOA | UNIVERSIDADE DE LISBOA

AML output

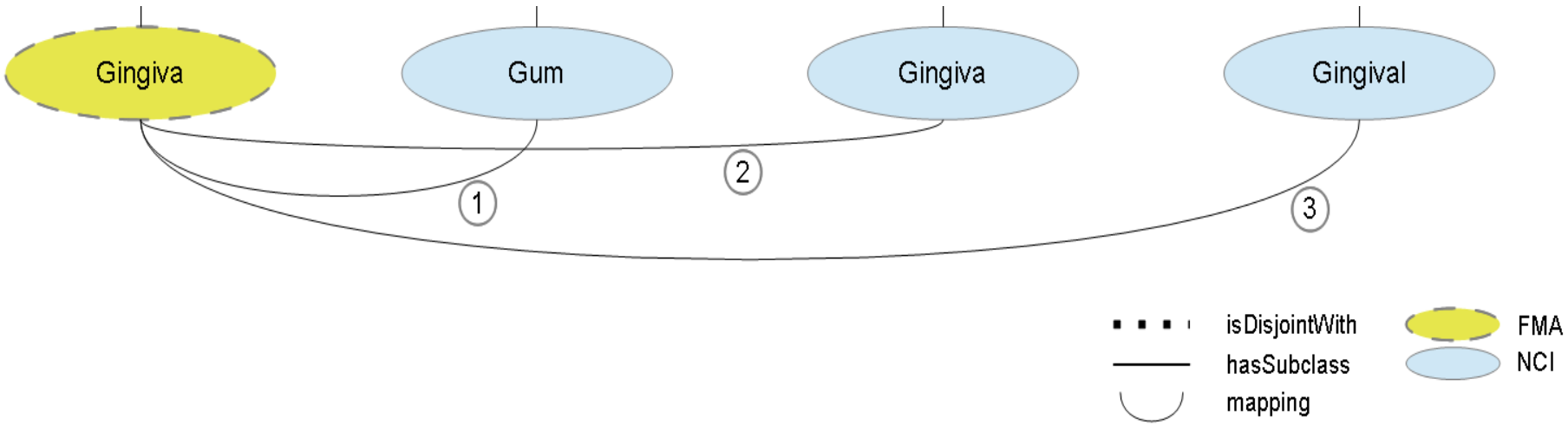
```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF xmlns="http://knowledgeweb.semanticweb.org/heterogeneity/alignment"
                    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
                    xmlns:xsd="http://www.w3.org/2001/XMLSchema#">
```



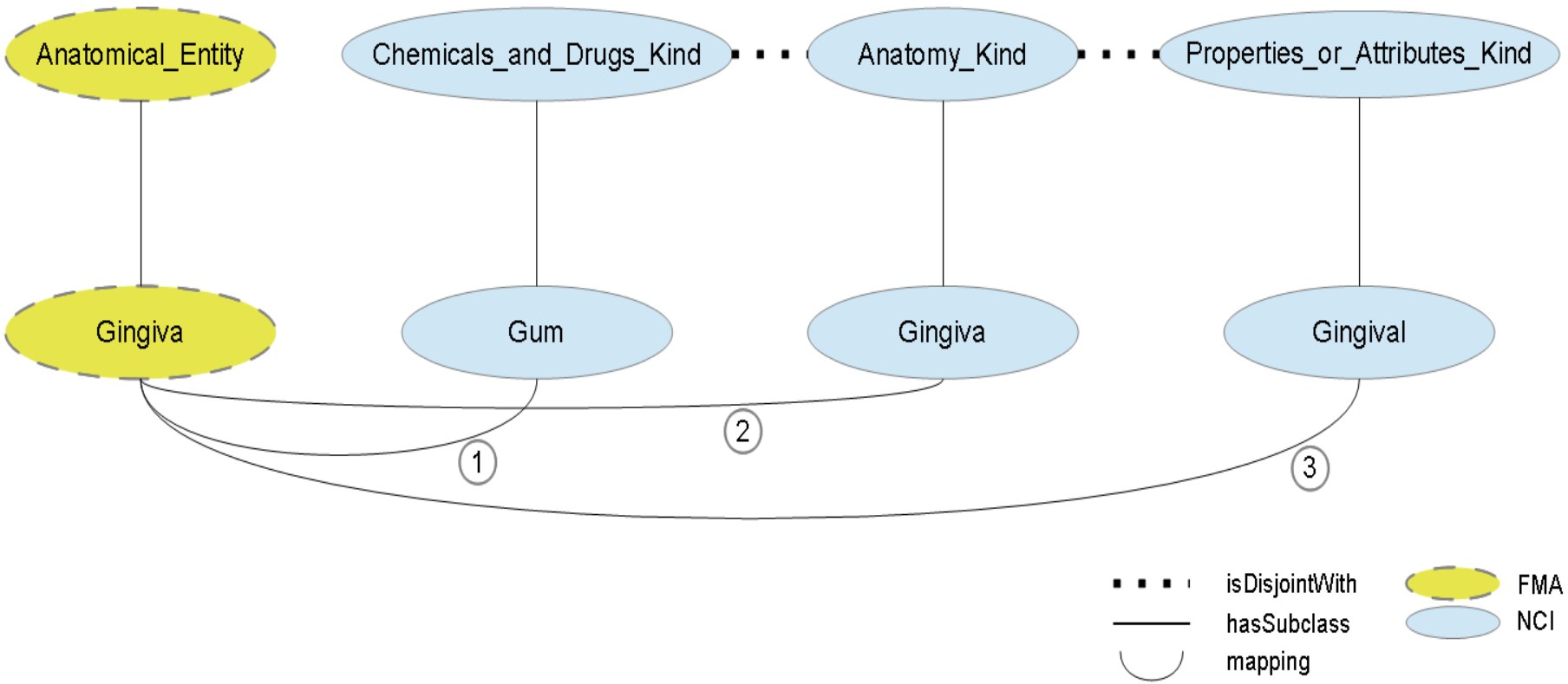
FMA-NCI - Microsoft Excel

	A	B	C	D	E	F
1	#AgreementMakerLight Alignment File					
2	#Source ontology:	http://bioontology.org/projects/ontologies/fma/fmaOwlDlComponent_2_0				
3	#Target ontology:	http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl				
4	Source URI	Source Label	Target URI	Target Label	Similarity	Relationship
5	http://bioontology.org/projec	lipid bilayer of plasma membrane	http://nci	plasma membrane lipid bilayer	0.9985	=
6	http://bioontology.org/projec	interstitial cell of pineal gland	http://nci	interstitial cell of the pineal gland	0.9907	=
7	http://bioontology.org/projec	elbow joint	http://nci	elbow joint	0.9801	=
8	http://bioontology.org/projec	otolith	http://nci	otolith	0.9801	=
9	http://bioontology.org/projec	palate	http://nci	palate	0.9801	=
10	http://bioontology.org/projec	tensor tympani	http://nci	tensor tympani	0.9801	=
11	http://bioontology.org/projec	lateral ventricle	http://nci	lateral ventricle	0.9801	=
12	http://bioontology.org/projec	cerebral hemisphere	http://nci	cerebral hemisphere	0.9801	=
13	http://bioontology.org/projec	anus	http://nci	anus	0.9801	=
14	http://bioontology.org/projec	ground substance	http://nci	ground substance	0.9801	=
15	http://bioontology.org/projec	urinary system	http://nci	urinary system	0.9801	=
16	http://bioontology.org/projec	helix	http://nci	helix	0.9801	=
17	http://bioontology.org/projec	semicircular duct	http://nci	semicircular duct	0.9801	=
18	http://bioontology.org/projec	olfactory nerve	http://nci	olfactory nerve	0.9801	=
19	http://bioontology.org/projec	embryo	http://nci	embryo	0.9801	=
20	http://bioontology.org/projec	spinal nerve	http://nci	spinal nerve	0.9801	=
21	http://bioontology.org/projec	geniohyoid	http://nci	geniohyoid	0.9801	=
22	http://bioontology.org/projec	chorion	http://nci	chorion	0.9801	=
23	http://bioontology.org/projec	limb	http://nci	limb	0.9801	=
24	http://bioontology.org/projec	nasal muscle	http://nci	nasal muscle	0.9801	=
25	http://bioontology.org/projec	epimysium	http://nci	epimysium	0.9801	=

Mappings need a context



Mappings need a context



U LISBOA | UNIVERSIDADE DE LISBOA

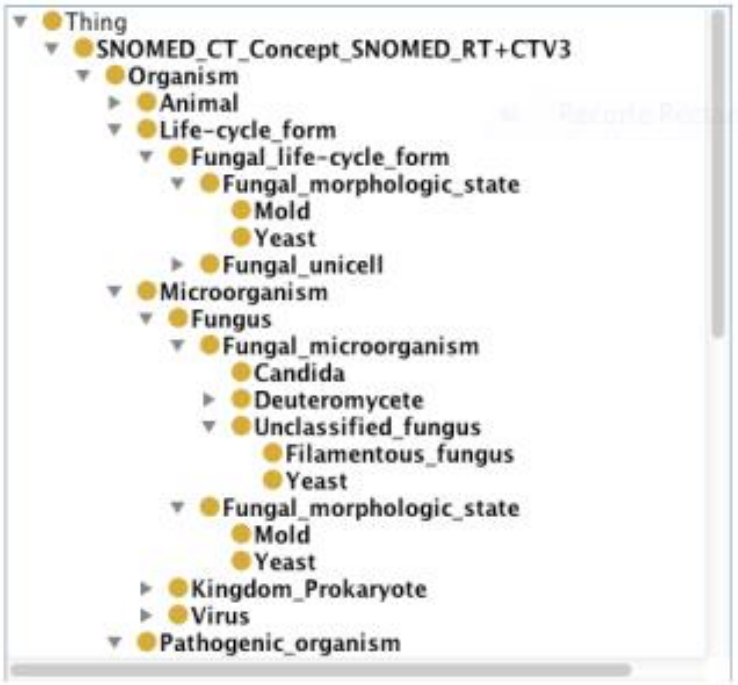
Visualizing Biomedical Ontologies Alignment

- Navigation and inspection of mappings
- Interactive matching, correct or add new mappings

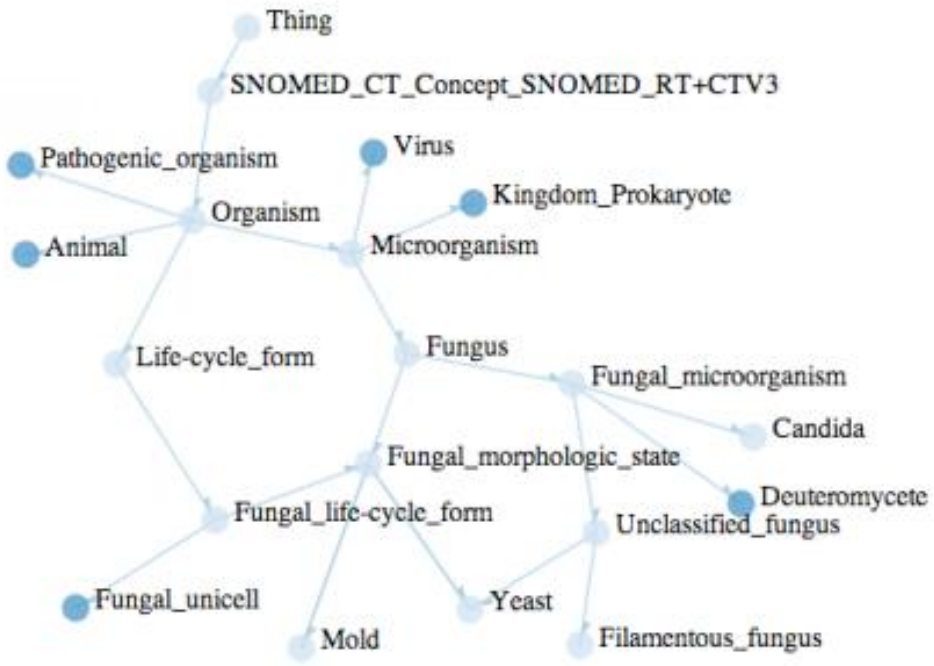
- Visualize ontologies
- Visualize mappings

Visualizing Biomedical Ontologies

Indented Tree



Graph



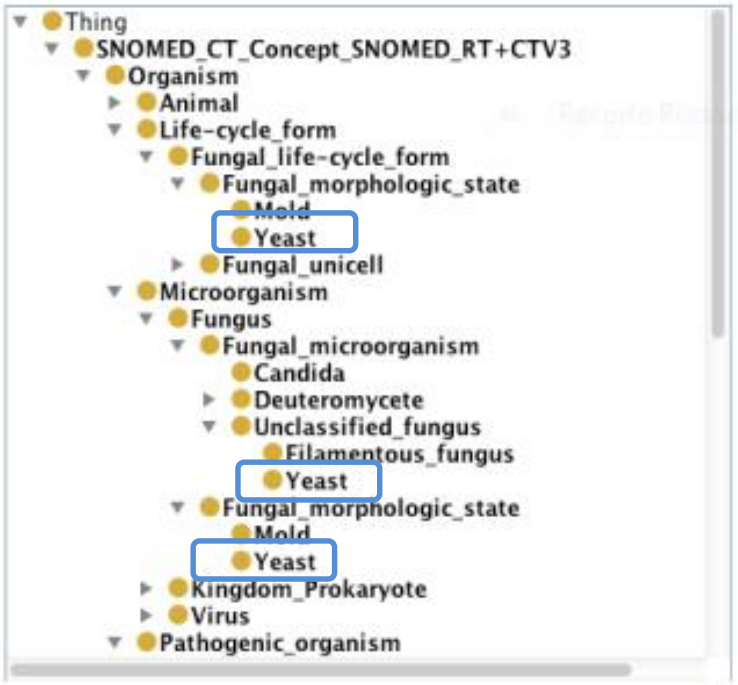
Better for evaluating mappings

Better for creating mappings

Adapted from: Fu, Bo, Natalya F. Noy, and Margaret-Anne Storey. "Indented Tree or Graph? A Usability Study of Ontology Visualization Techniques in the Context of Class Mapping Evaluation." I*SWC 2013*.

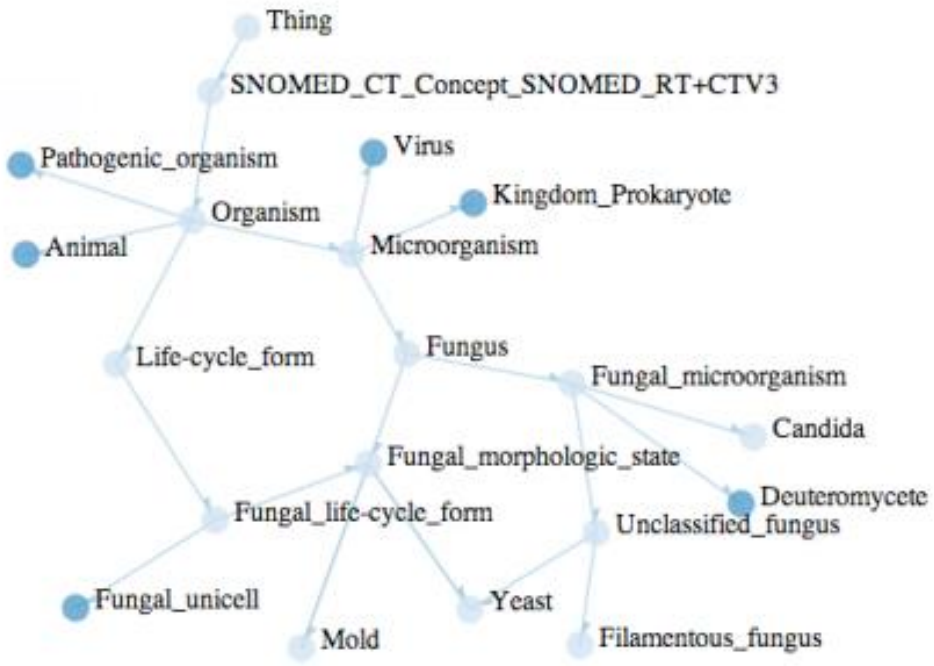
Visualizing Biomedical Ontologies

Indented Tree



Better for evaluating mappings

Graph

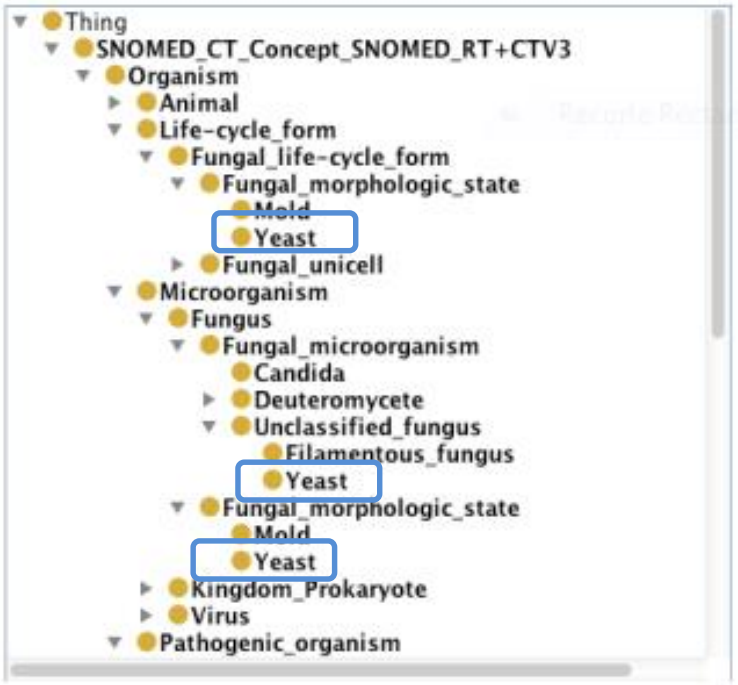


Better for creating mappings

Adapted from: Fu, Bo, Natalya F. Noy, and Margaret-Anne Storey. "Indented Tree or Graph? A Usability Study of Ontology Visualization Techniques in the Context of Class Mapping Evaluation." I*SWC 2013*.

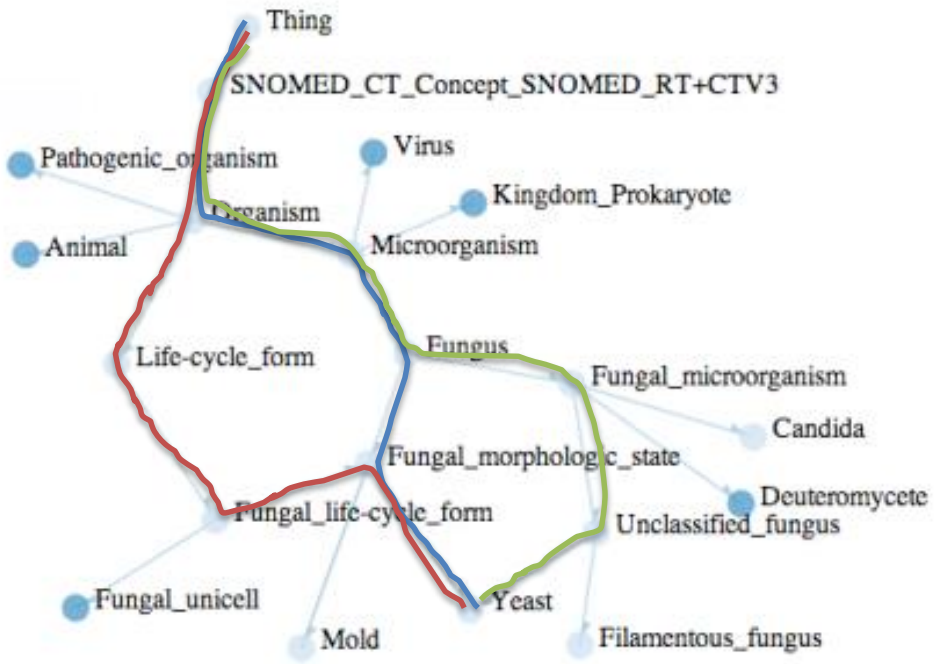
Visualizing Biomedical Ontologies

Indented Tree



Better for evaluating mappings

Graph

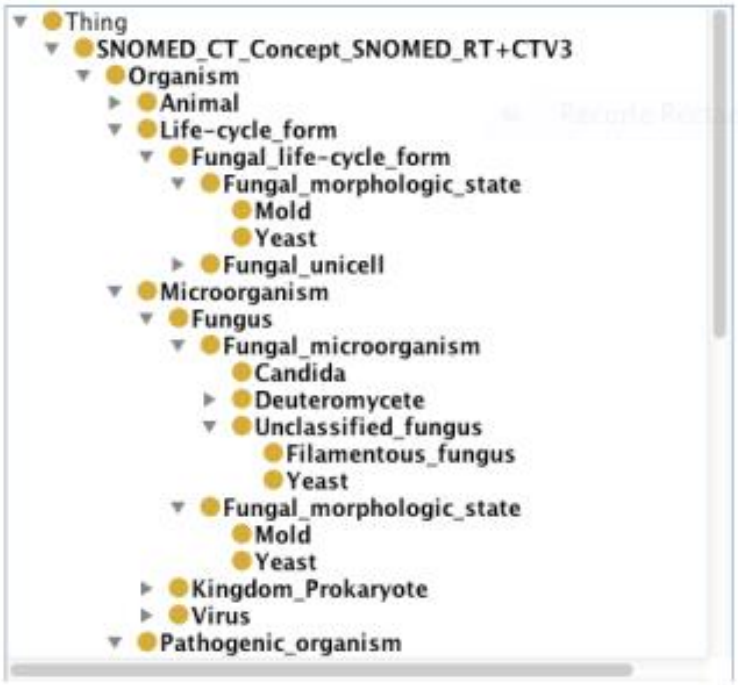


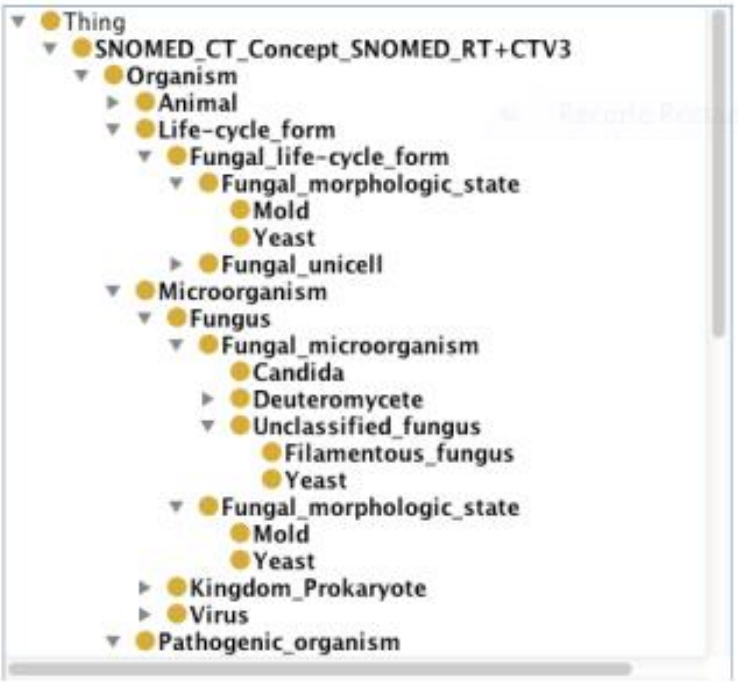
Better for creating mappings

Adapted from: Fu, Bo, Natalya F. Noy, and Margaret-Anne Storey. "Indented Tree or Graph? A Usability Study of Ontology Visualization Techniques in the Context of Class Mapping Evaluation." I*SWC 2013*.

U LISBOA | UNIVERSIDADE DE LISBOA

Visualizing Biomedical Ontologies Alignment

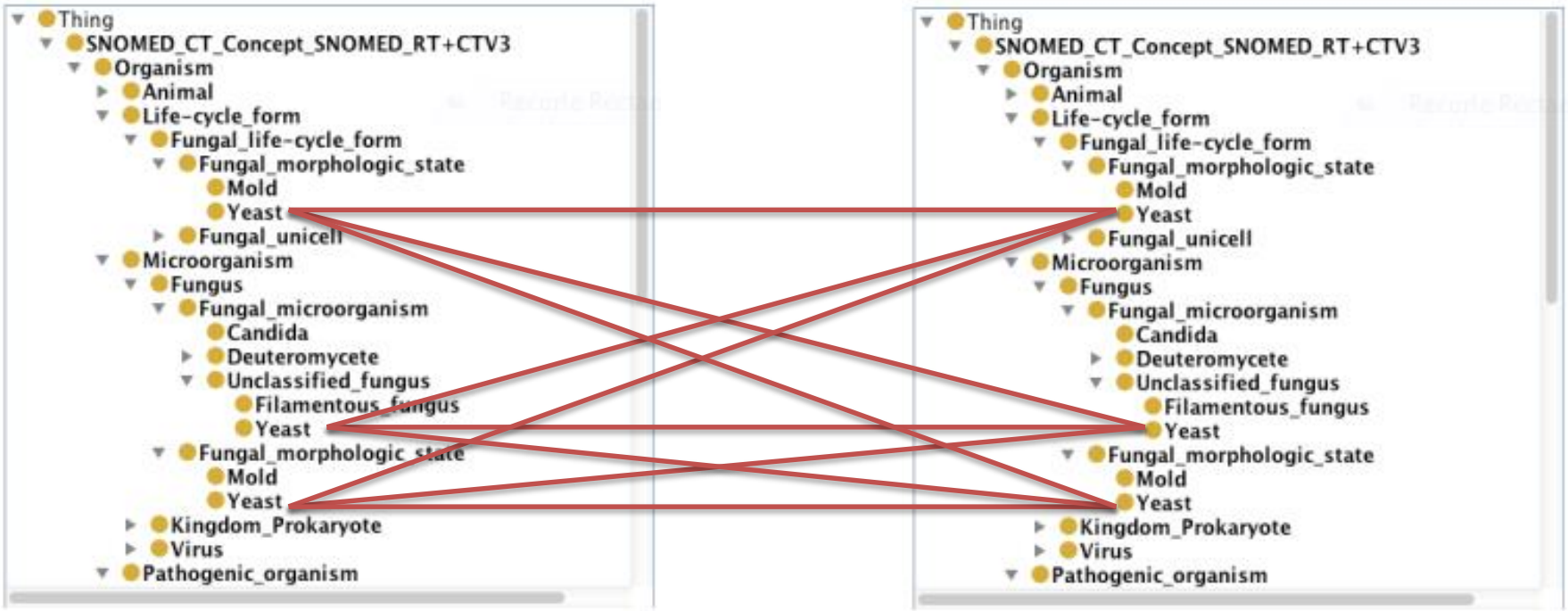




Adapted from: Fu, Bo, Natalya F. Noy, and Margaret-Anne Storey. "Indented Tree or Graph? A Usability Study of Ontology Visualization Techniques in the Context of Class Mapping Evaluation." I*SWC 2013*.

U LISBOA | UNIVERSIDADE DE LISBOA

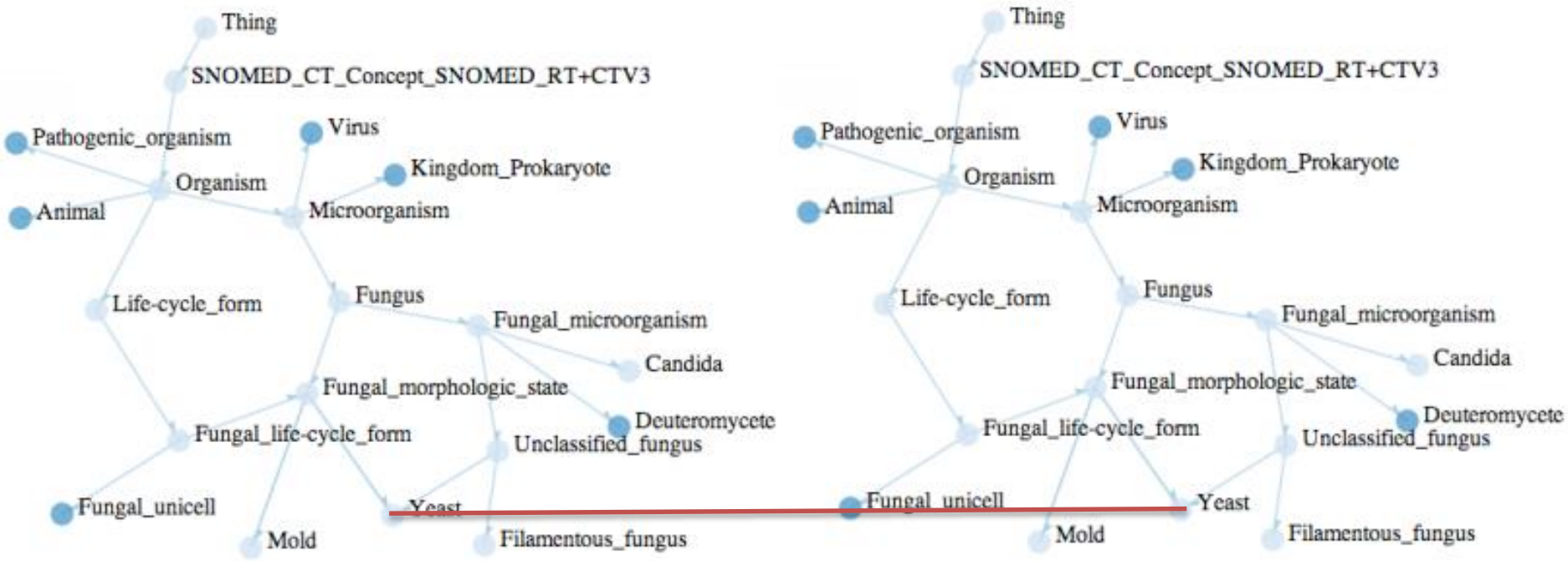
Visualizing Biomedical Ontologies Alignment



Adapted from: Fu, Bo, Natalya F. Noy, and Margaret-Anne Storey. "Indented Tree or Graph? A Usability Study of Ontology Visualization Techniques in the Context of Class Mapping Evaluation." I*SWC 2013*.

U LISBOA | UNIVERSIDADE DE LISBOA

Visualizing Biomedical Ontologies Alignment



Adapted from: Fu, Bo, Natalya F. Noy, and Margaret-Anne Storey. "Indented Tree or Graph? A Usability Study of Ontology Visualization Techniques in the Context of Class Mapping Evaluation." I*SWC 2013*.

Visualizing alignments using indented trees

- Focused on taxonomical hierarchy
- Loss of other relations
- Model distortion due to multiple inheritance
- Potentially difficult to see neighbourhood mappings
- Easier to navigate hierarchy

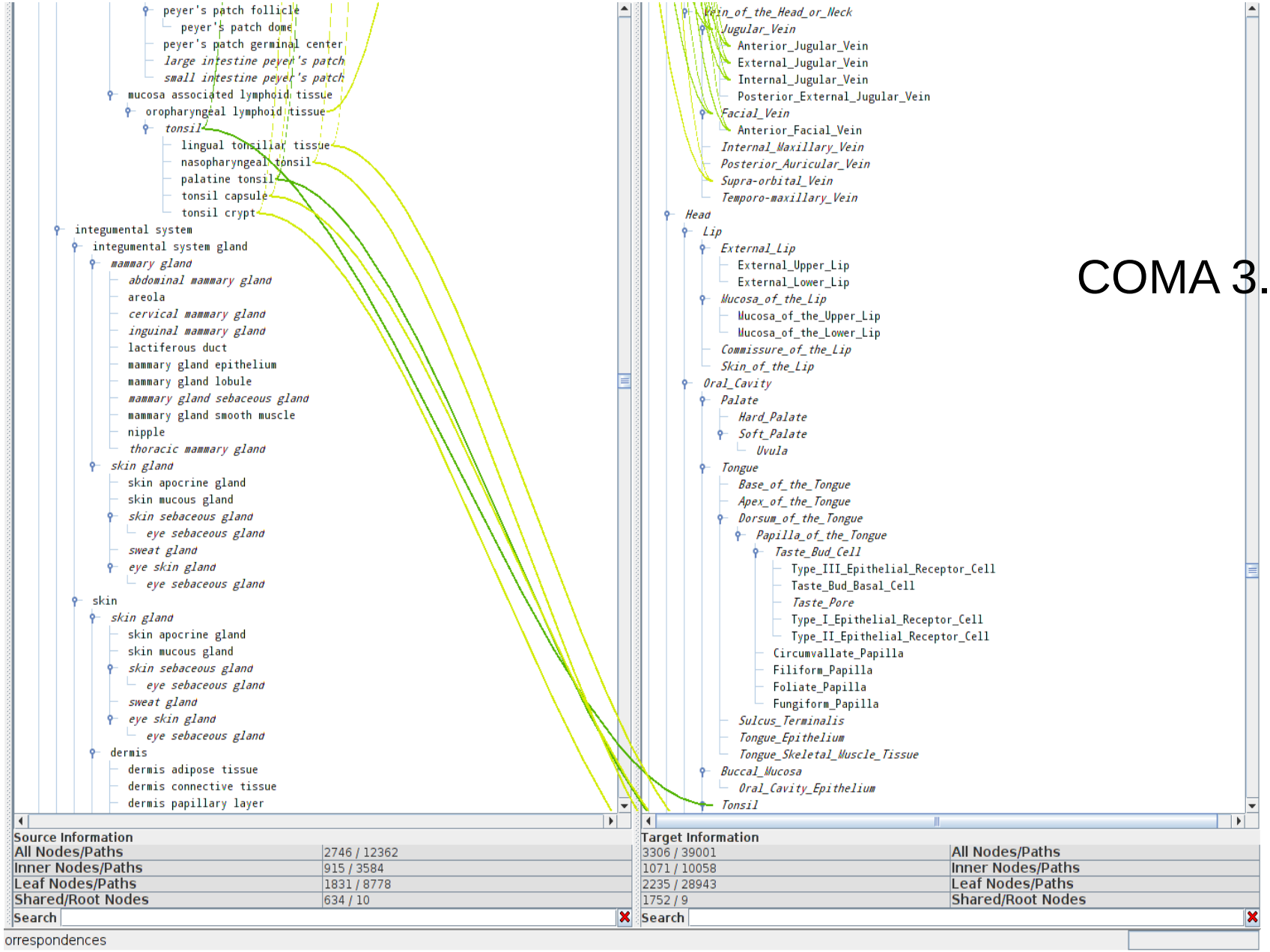
Visualizing alignments using graphs

- Can be focused on mappings
- Good at representing different relations and multiple inheritance
- Potential for cluttered display
  - Pan and zoom
  - Click nodes to expand
  - Subgraph

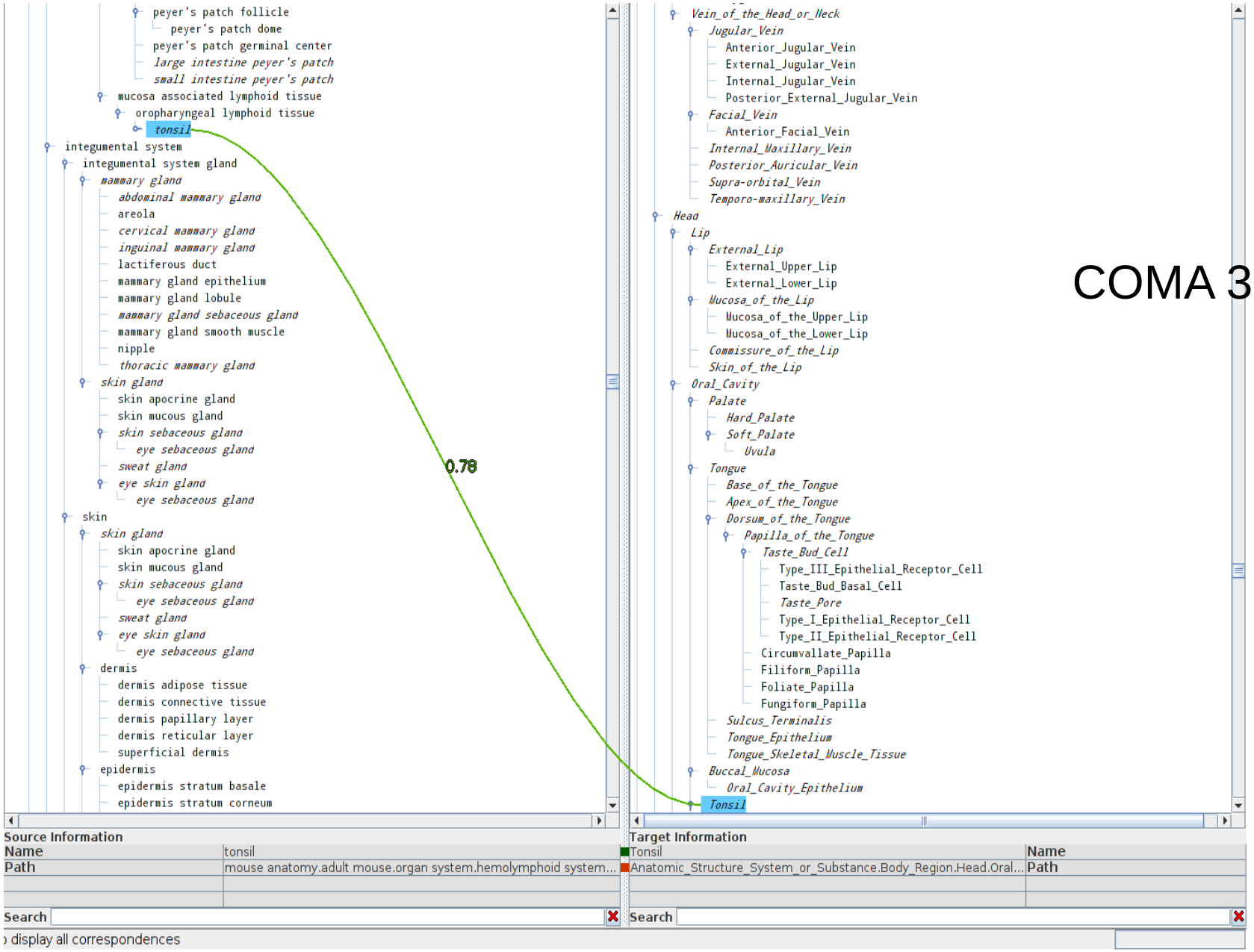
Visualizing Biomedical Ontologies Alignment

- COMA 3.0
- AgreementMaker

- AgreementMakerLight


peyer's patch follicle
peyer's patch dome
peyer's patch germinal center
large intestine peyer's patch
small intestine peyer's patch
mucosa associated lymphoid tissue
oropharyngeal lymphoid tissue
tonsil
lingual tonsillar tissue
nasopharyngeal tonsil
palatine tonsil
tonsil capsule
tonsil crypt
integumental system
integumental system gland
mammary gland
abdominal mammary gland
areola
cervical mammary gland
inguinal mammary gland
lactiferous duct
mammary gland epithelium
mammary gland lobule
mammary gland sebaceous gland
mammary gland smooth muscle
nipple
thoracic mammary gland
skin gland
skin apocrine gland
skin mucous gland
skin sebaceous gland
eye sebaceous gland
sweat gland
eye skin gland
eye sebaceous gland
skin
skin gland
skin apocrine gland
skin mucous gland
skin sebaceous gland
eye sebaceous gland
sweat gland
eye skin gland
eye sebaceous gland
dermis
dermis adipose tissue
dermis connective tissue
dermis papillary layer
Vein_of_the_Head_or_Neck
Jugular_Vein
Anterior_Jugular_Vein
External_Jugular_Vein
Internal_Jugular_Vein
Posterior_External_Jugular_Vein
Facial_Vein
Anterior_Facial_Vein
Internal_Maxillary_Vein
Posterior_Auricular_Vein
Supra-orbital_Vein
Temporo-maxillary_Vein
Head
Lip
External_Lip
External_Upper_Lip
External_Lower_Lip
Mucosa_of_the_Lip
Mucosa_of_the_Upper_Lip
Mucosa_of_the_Lower_Lip
Commissure_of_the_Lip
Skin_of_the_Lip
Oral_Cavity
Palate
Hard_Palate
Soft_Palate
Uvula
Tongue
Base_of_the_Tongue
Apex_of_the_Tongue
Dorsum_of_the_Tongue
Papilla_of_the_Tongue
Taste_Bud_Cell
Type_III_Epithelial_Receptor_Cell
Taste_Bud_Basal_Cell
Taste_Pore
Type_I_Epithelial_Receptor_Cell
Type_II_Epithelial_Receptor_Cell
Circumvallate_Papilla
Filiform_Papilla
Foliate_Papilla
Fungiform_Papilla
Sulcus_Terminalis
Tongue_Epithelium
Tongue_Skeletal_Muscle_Tissue
Buccal_Mucosa
Oral_Cavity_Epithelium
Tonsil
Source Information
All Nodes/Paths 2746 / 12362
Inner Nodes/Paths 915 / 3584
Leaf Nodes/Paths 1831 / 8778
Shared/Root Nodes 634 / 10
Search
Target Information
3306 / 39001 All Nodes/Paths
1071 / 10058 Inner Nodes/Paths
2235 / 28943 Leaf Nodes/Paths
1752 / 9 Shared/Root Nodes
Search
orrespondences

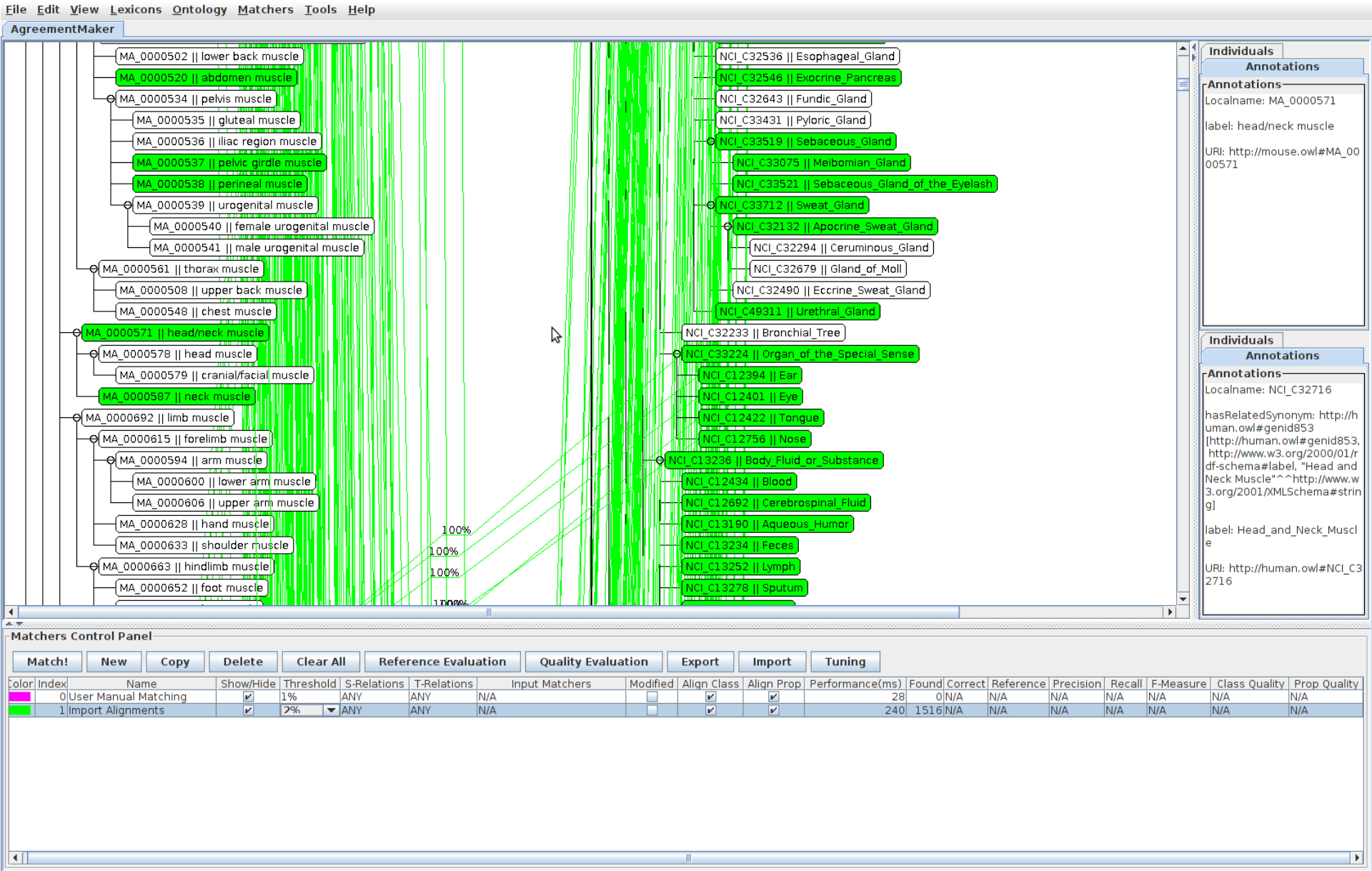
COMA 3.0


peyer's patch follicle
peyer's patch dome
peyer's patch germinal center
large intestine peyer's patch
small intestine peyer's patch
mucosa associated lymphoid tissue
oropharyngeal lymphoid tissue
tonsil
integumental system
integumental system gland
mammary gland
abdominal mammary gland
areola
cervical mammary gland
inguinal mammary gland
lactiferous duct
mammary gland epithelium
mammary gland lobule
mammary gland sebaceous gland
mammary gland smooth muscle
nipple
thoracic mammary gland
skin gland
skin apocrine gland
skin mucous gland
skin sebaceous gland
eye sebaceous gland
sweat gland
eye skin gland
eye sebaceous gland
skin
skin gland
skin apocrine gland
skin mucous gland
skin sebaceous gland
eye sebaceous gland
sweat gland
eye skin gland
eye sebaceous gland
dermis
dermis adipose tissue
dermis connective tissue
dermis papillary layer
dermis reticular layer
superficial dermis
epidermis
epidermis stratum basale
epidermis stratum corneum
0.78
Vein_of_the_Head_or_Neck
Jugular_Vein
Anterior_Jugular_Vein
External_Jugular_Vein
Internal_Jugular_Vein
Posterior_External_Jugular_Vein
Facial_Vein
Anterior_Facial_Vein
Internal_Maxillary_Vein
Posterior_Auricular_Vein
Supra-orbital_Vein
Temporo-maxillary_Vein
Head
Lip
External_Lip
External_Upper_Lip
External_Lower_Lip
Mucosa_of_the_Lip
Mucosa_of_the_Upper_Lip
Mucosa_of_the_Lower_Lip
Commissure_of_the_Lip
Skin_of_the_Lip
Oral_Cavity
Palate
Hard_Palate
Soft_Palate
Uvula
Tongue
Base_of_the_Tongue
Apex_of_the_Tongue
Dorsum_of_the_Tongue
Papilla_of_the_Tongue
Taste_Bud_Cell
Type_III_Epithelial_Receptor_Cell
Taste_Bud_Basal_Cell
Taste_Pore
Type_I_Epithelial_Receptor_Cell
Type_II_Epithelial_Receptor_Cell
Circumvallate_Papilla
Filiform_Papilla
Foliate_Papilla
Fungiform_Papilla
Sulcus_Terminalis
Tongue_Epithelium
Tongue_Skeletal_Muscle_Tissue
Buccal_Mucosa
Oral_Cavity_Epithelium
Tonsil
Source Information
Name
tonsil
Path
mouse anatomy.adult mouse.organ system.hemolymphoid system...
Target Information
Tonsil
Name
Anatomic_Structure_System_or_Substance.Body_Region.Head.Oral...
Path
Search
Search
display all correspondences

COMA 3.0

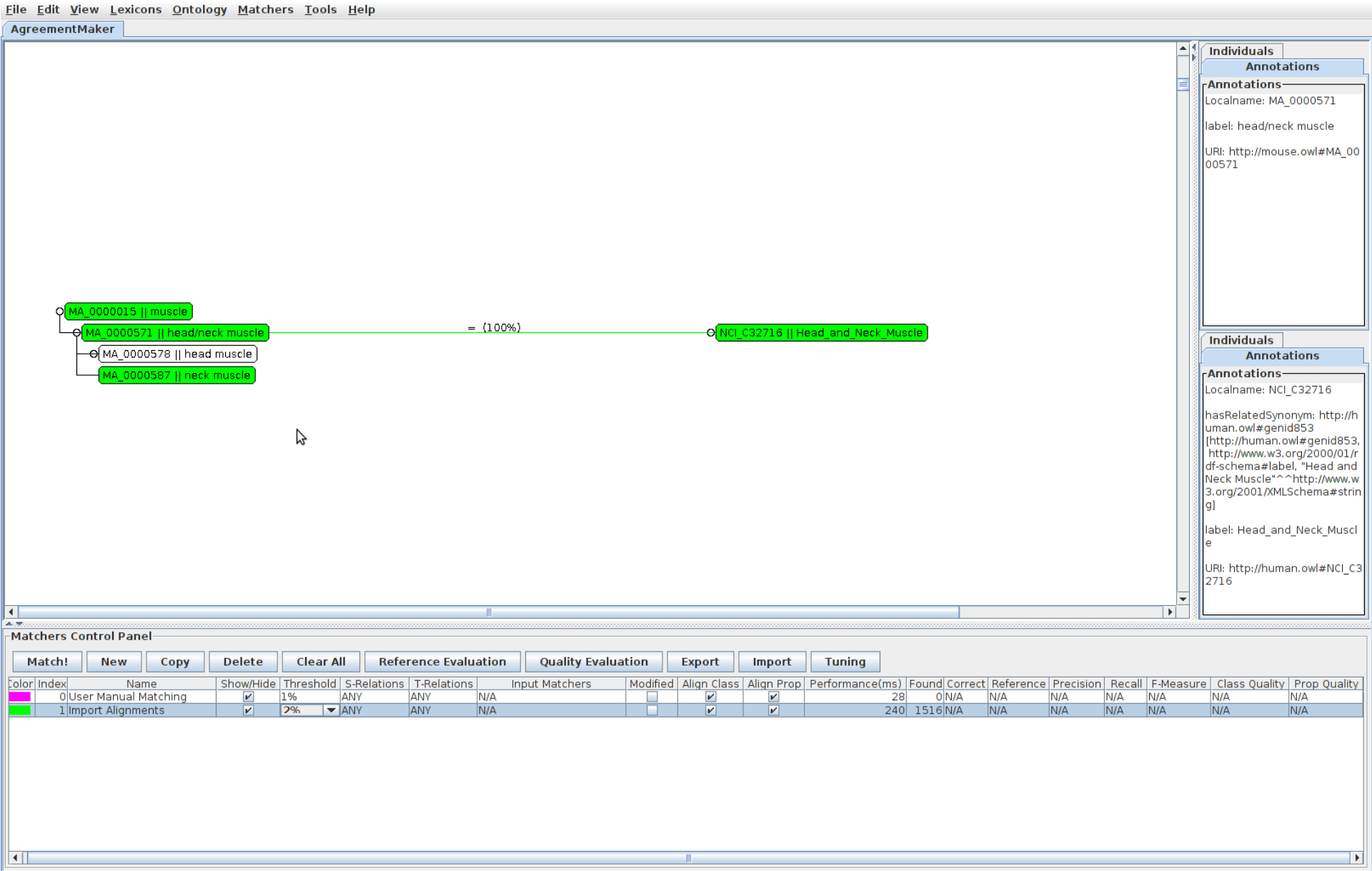
U LISBOA | UNIVERSIDADE DE LISBOA

AgreementMaker



U LISBOA | UNIVERSIDADE DE LISBOA

AgreementMaker

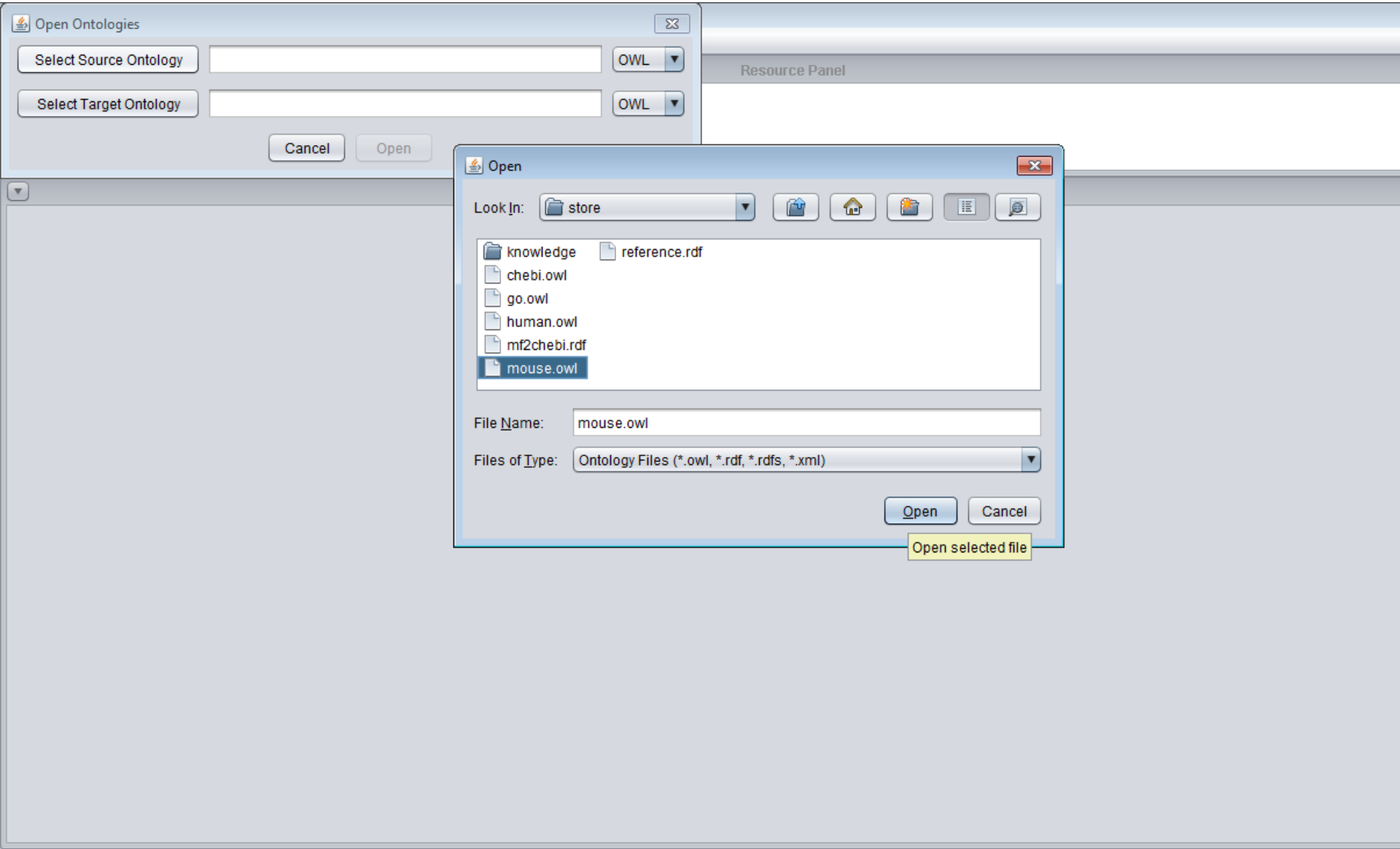


Visualizing alignments in AML

- A **single graph**
- Focused on **mappings**
  - Search for a mapped class
  - Browse mappings sequentially
- Represents **all types of relations**
  - Labeled and directed
- Shows only **subgraph** of the neighbourhood of the mapping
  - Size set by user

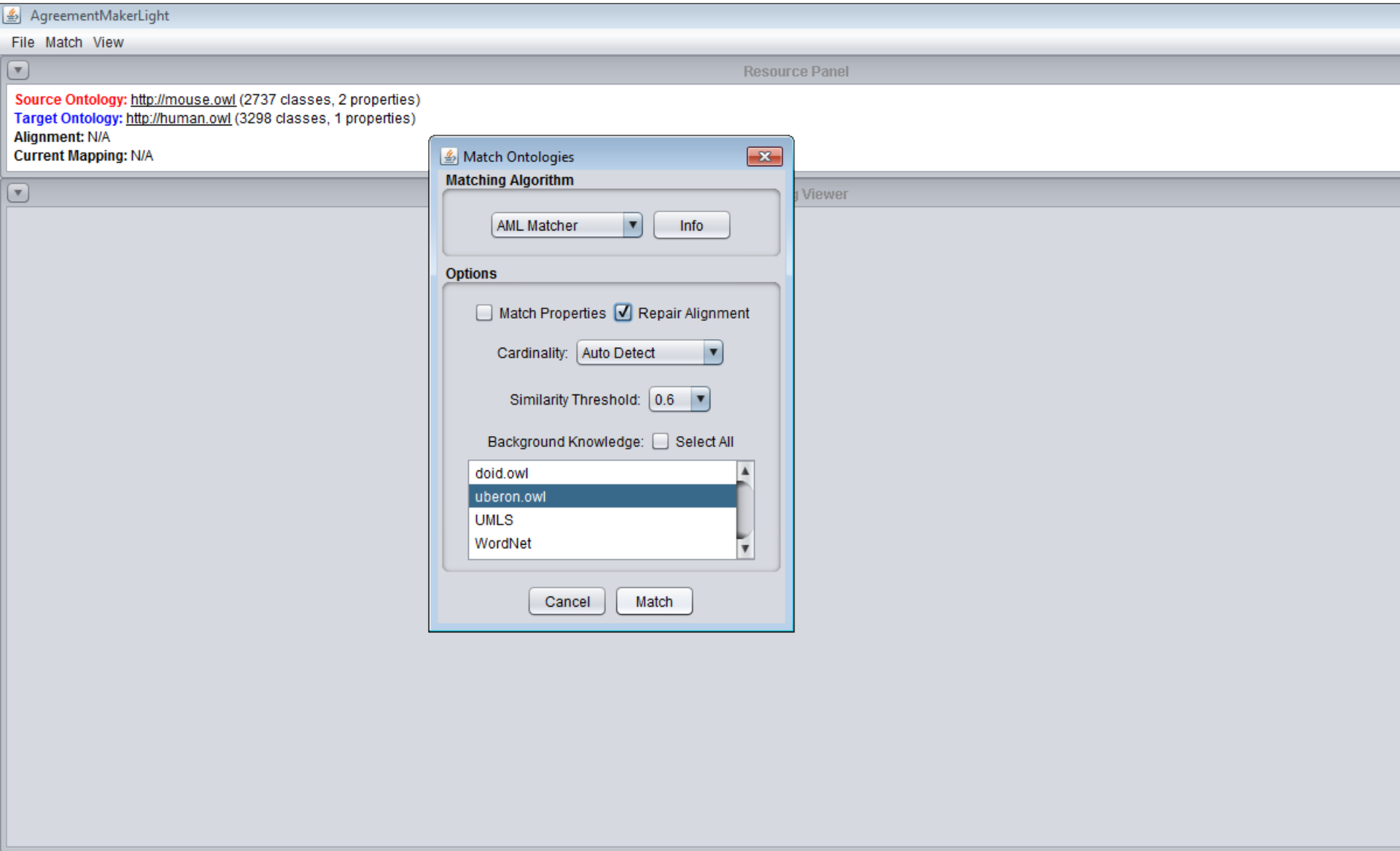
U LISBOA | UNIVERSIDADE DE LISBOA

AML



U LISBOA | UNIVERSIDADE DE LISBOA

AML



U LISBOA | UNIVERSIDADE DE LISBOA

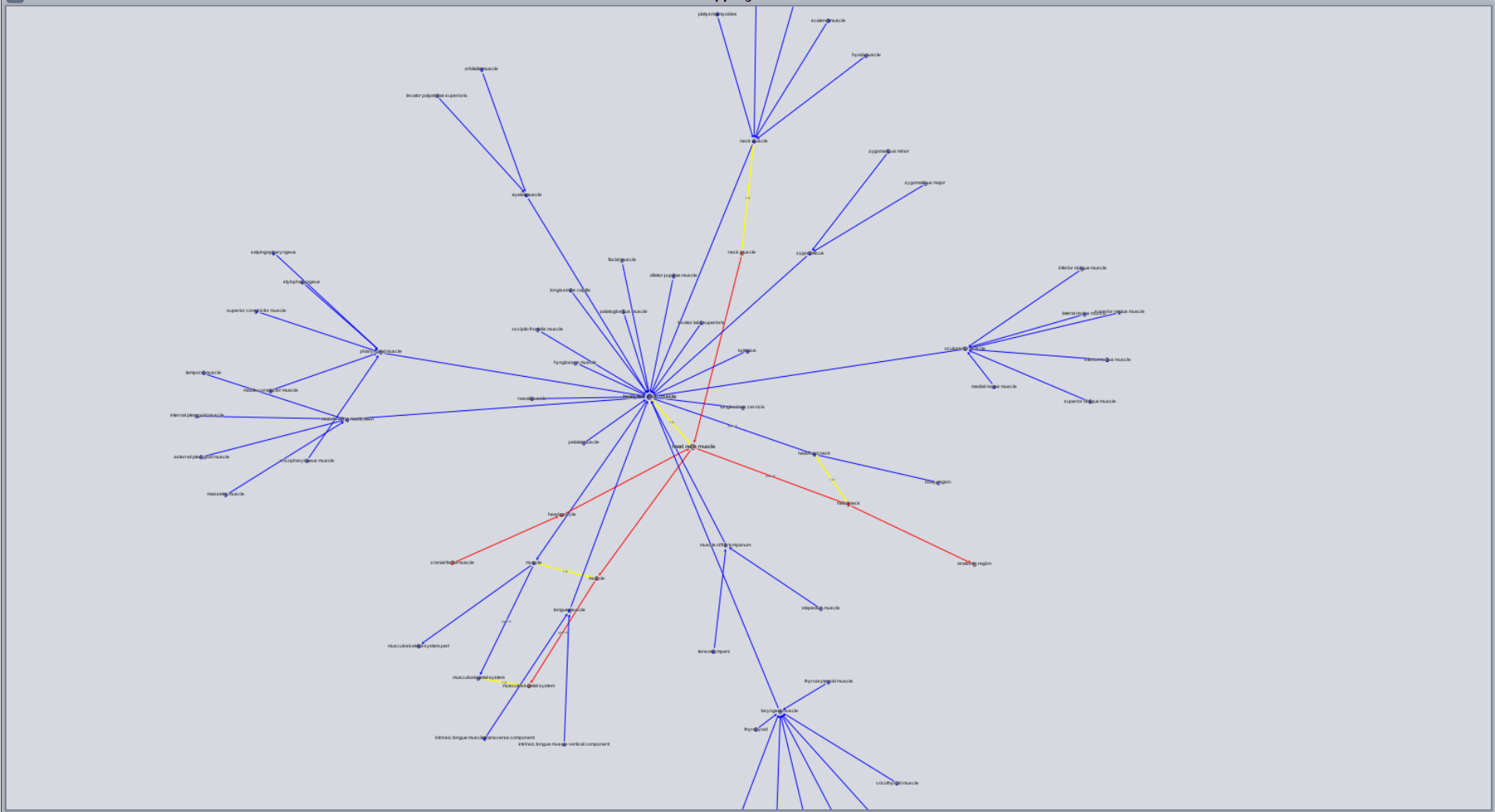
AML

File Match View

Resource Panel

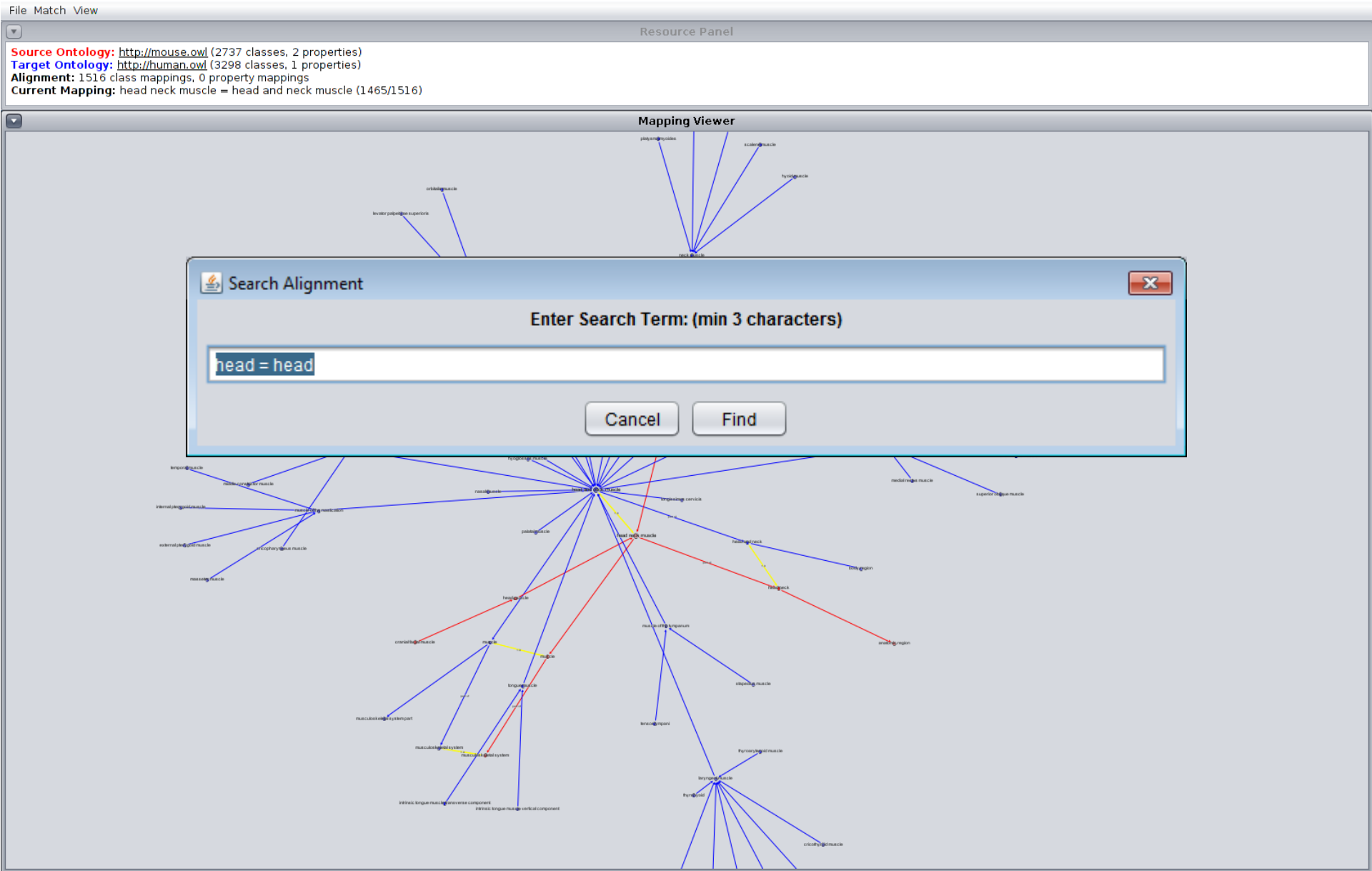
Source Ontology: http://mouse.owl (2737 classes, 2 properties)
Target Ontology: http://human.owl (3298 classes, 1 properties)
Alignment: 1516 class mappings, 0 property mappings
Current Mapping: head neck muscle = head and neck muscle (1465/1516)

Mapping Viewer



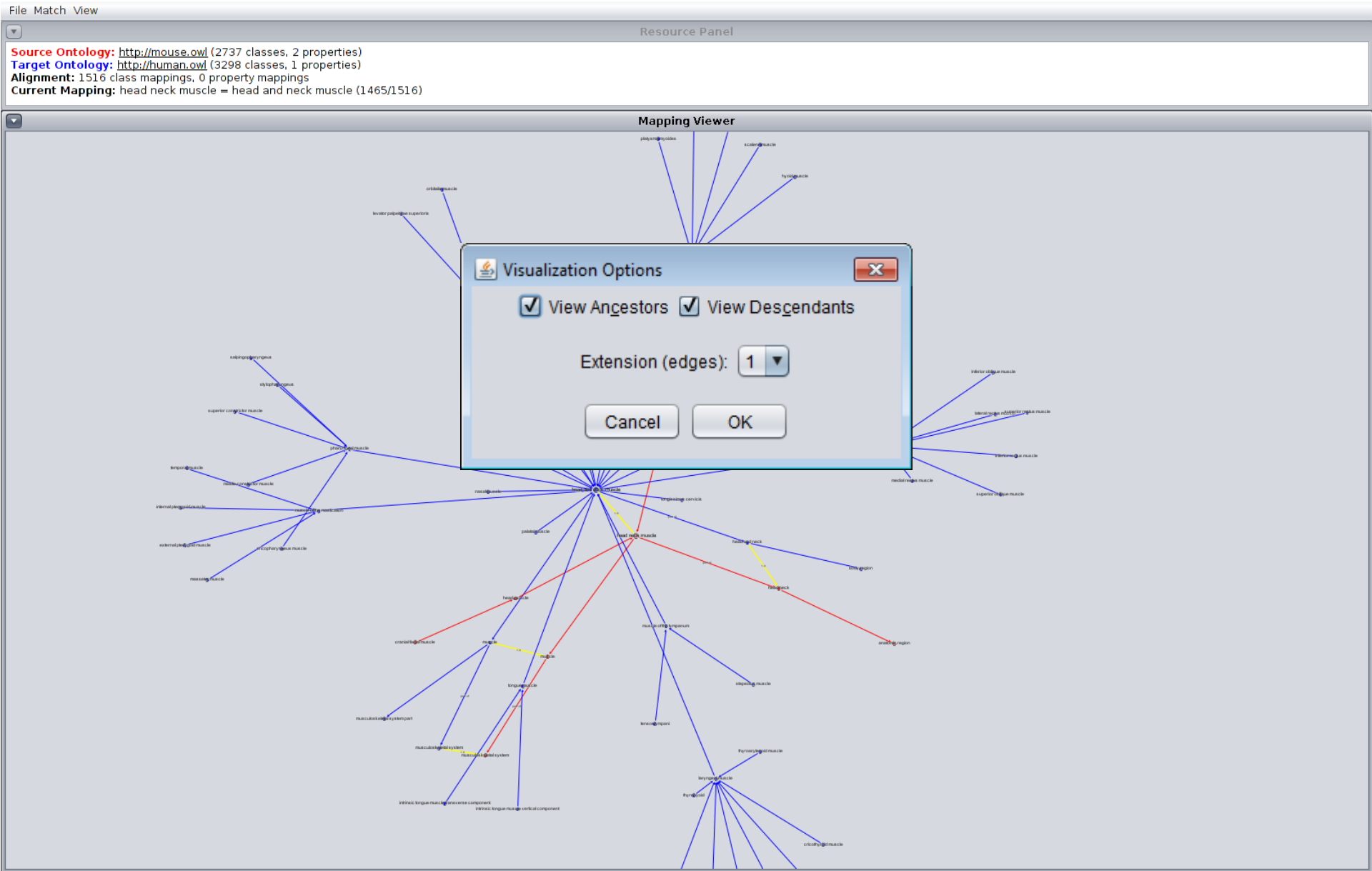
U LISBOA | UNIVERSIDADE DE LISBOA

AML



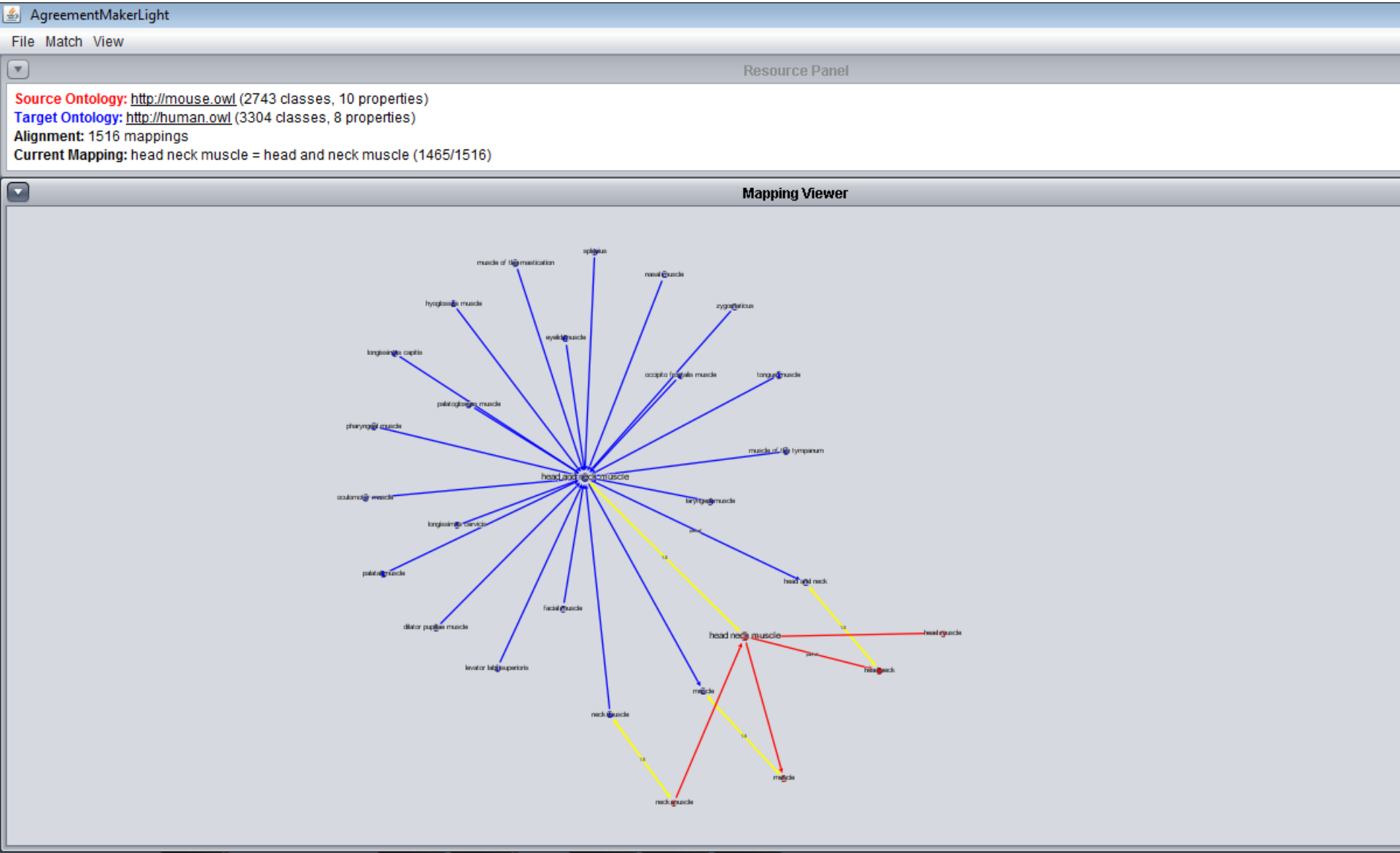
U LISBOA | UNIVERSIDADE DE LISBOA

AML



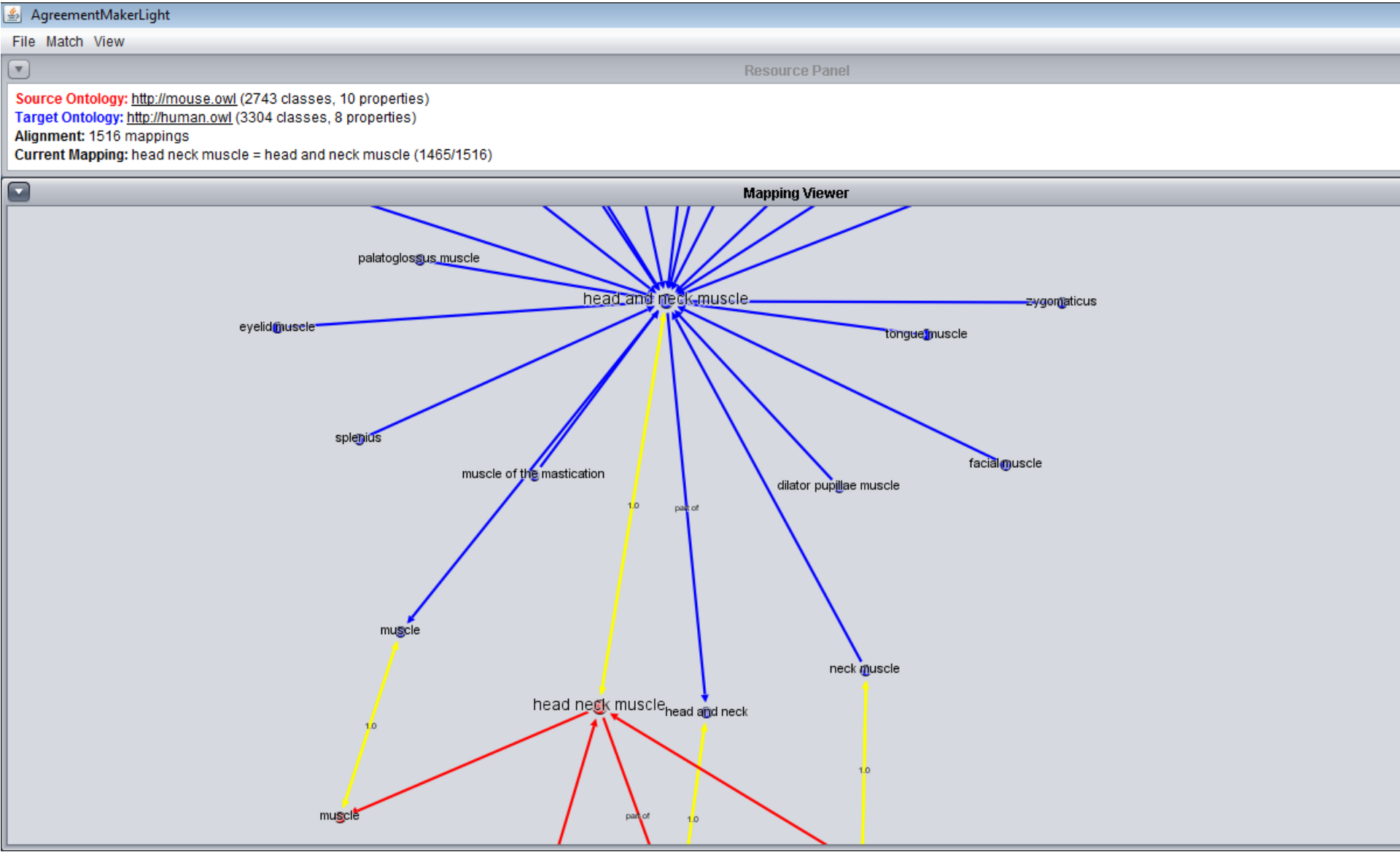
U LISBOA | UNIVERSIDADE DE LISBOA

AML



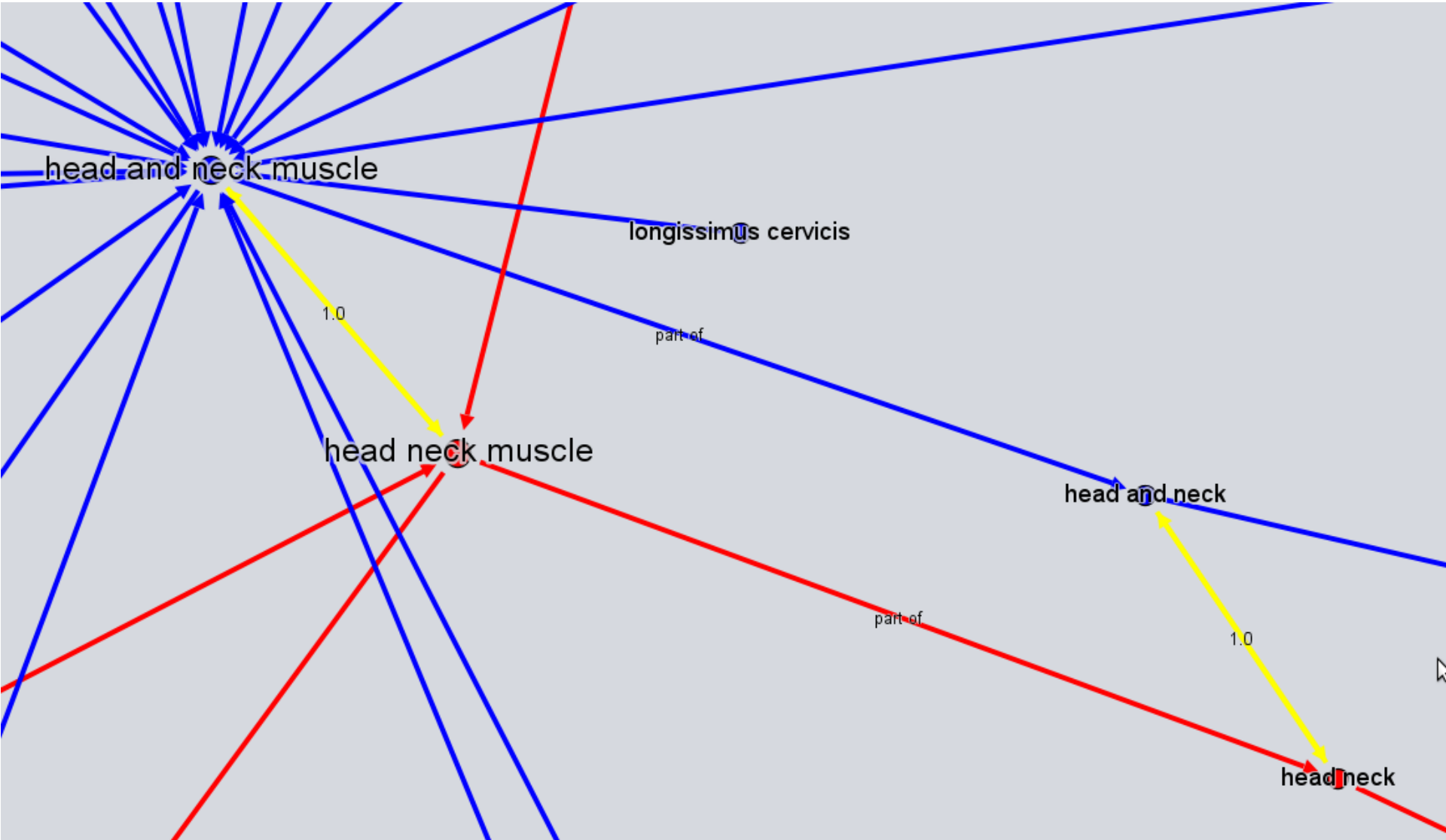
U LISBOA | UNIVERSIDADE DE LISBOA

AML



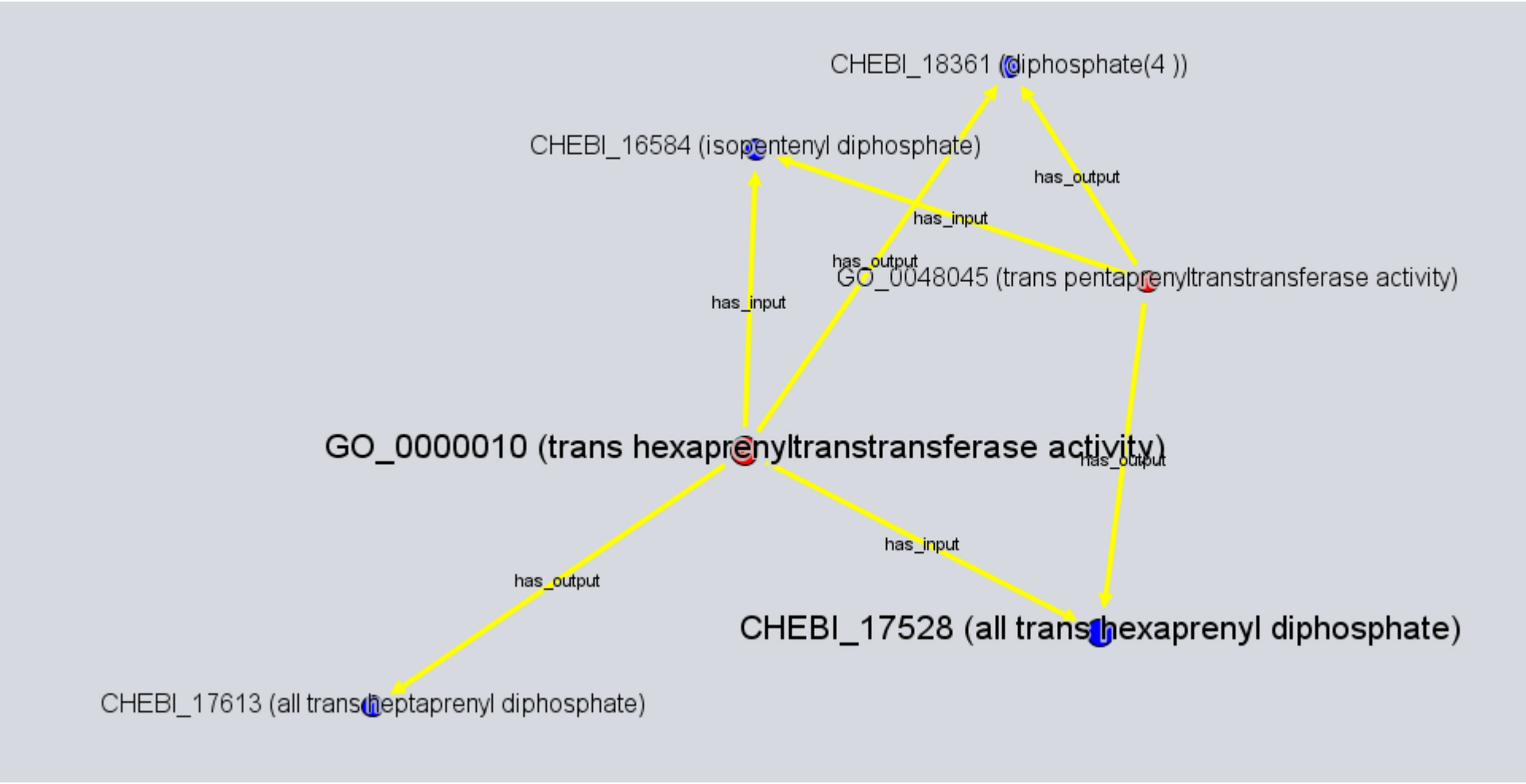
U LISBOA | UNIVERSIDADE DE LISBOA

AML



U LISBOA | UNIVERSIDADE DE LISBOA

Loading external mappings in AML



Conclusions

- Trees are better for browsing
- Graphs are better for context

- Future work:
  - Add dynamic graphs, graph color customization
  - Add inspection of classes properties
  - **Investigate integrating both trees and graphs**

U LISBOA | UNIVERSIDADE DE LISBOA

Faculdade de Ciências
UNIVERSIDADE DE LISBOA

Acknowledgements

- SOMER team at LASIGE-UL
- Funding:
  - –SOMER project (PTDC/EIA-EIA/119119/2010)
  - –FCT multi-annual funding program to LASIGE.

- Visit us at http://somer.fc.ul.pt

- Get AML at GitHub – Open source and completely free to use, adapt or fork!

https://github.com/AgreementMakerLight/AML-Project
https://github.com/AgreementMakerLight/AML-Jar

SOMER

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA E DO ENSINO SUPERIOR