



# Populous: A tool for Populating OWL Ontologies from Templates

Simon Jupp  
Bio Health Informatics Group,  
School of Computer Science,  
University of Manchester, UK

10.12.2010 SWAT4LS 2010, Berlin.

# Ontology development barriers



- The underlying representation
- The art of modeling
- Community engagement
- The tool support

# Templates approach



- Shield the underlying technology
- Collect repetitive information consistently
- Put constraints on the input data
- Abstraction from any complex modeling

# Capturing the differentia

## Knowledge

All Eukaryotic Cells are either nucleated or anucleate, some cells are multinucleate

## Ontologically

'Eukaryotic Cells' *has\_nucleation* some 'Nucleation'  
'Nucleation' *subClassOf* {mononucleate , binucleate , polynucleate , anucleate}

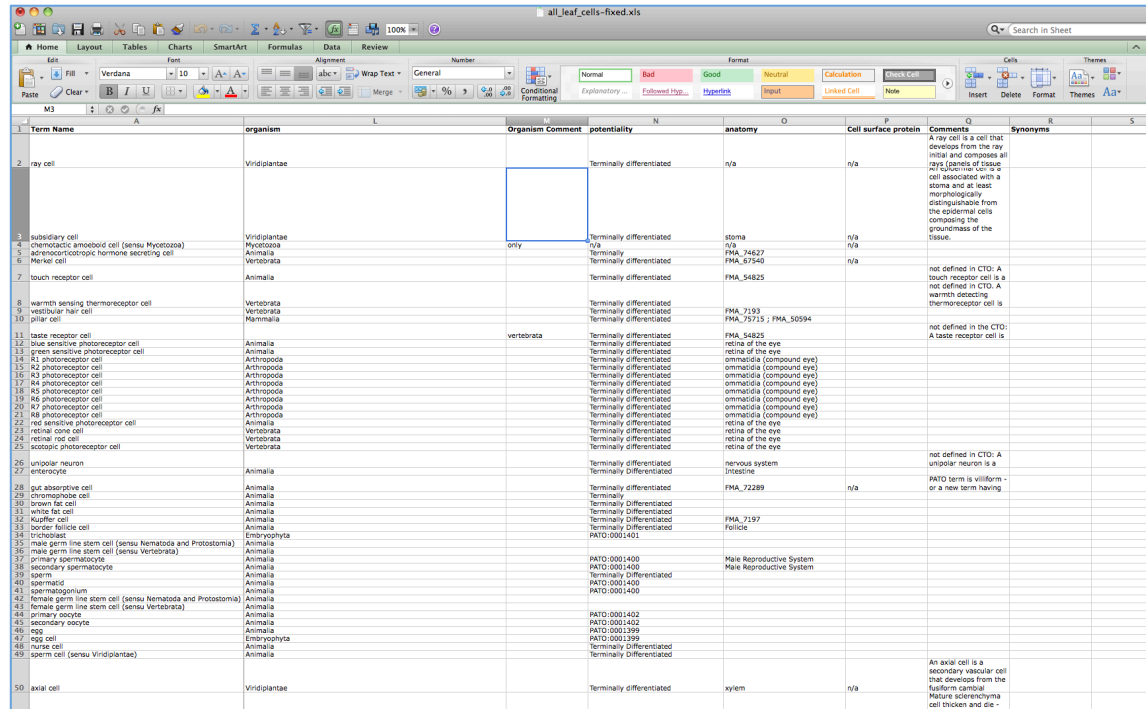
## Differentia

'Eukaryotic Cells' *has\_nucleation* some 'Nucleation'  
'Nucleation' *subClassOf* {mononucleate , binucleate , polynucleate , anucleate}

## Real Examples

'Eukaryotic Cells'	'Nucleation'
Mononuclear phagocyte	mononucleate
Flight Muscle cell	multinucleate
Red Blood cell	anucleate

# Spreadsheets

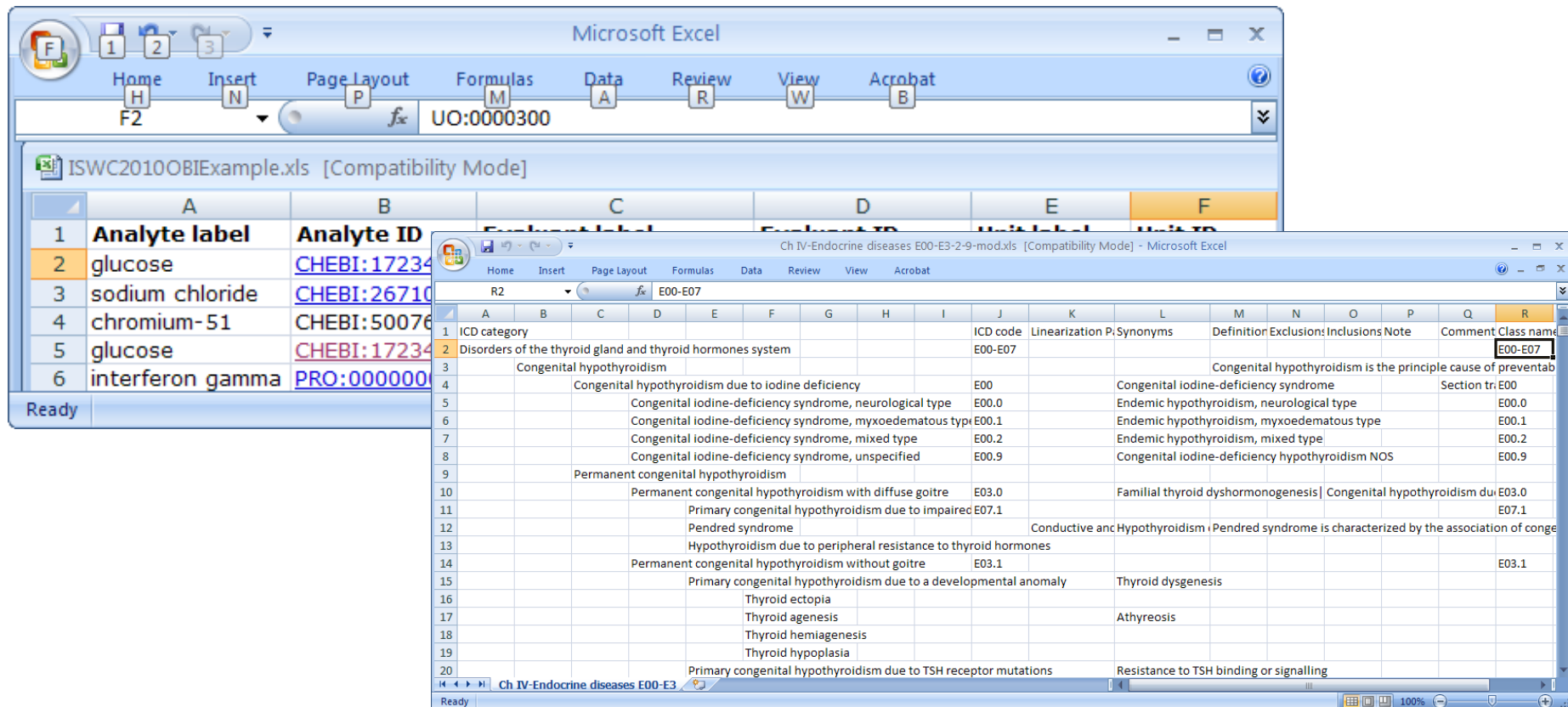


Term Name	organism	Organism Comment	potentiality	anatomy	Cell surface protein	Comments	Synonyms
2 ray cell	Viridiplantae		Terminally differentiated	n/a	n/a	A ray cell is a cell that develops from the ray initial and composes all rays (panels of tissue) in a leaf. It is a cell associated with a stoma and at least morphologically distinguishable from the epidermal cells composing the groundmass of the tissue.	
3 subsidiary cell	Viridiplantae	only	Terminally differentiated	stoma	n/a		
4 chemotactile amoeboid cell (sensu Mycetozoa)	Mycetozoa		n/a	n/a	n/a		
5 adrenocorticotrophic hormone secreting cell	Animalia		Terminally differentiated	FMA_24627	n/a		
6 Merkel cell	Vertebrata		Terminally differentiated	FMA_57540	n/a	not defined in CTD: A touch receptor cell is a	
7 touch receptor cell	Animalia		Terminally differentiated	FMA_54825		not defined in CTD: A touch receptor cell is a	
8 warmth sensing thermoreceptor cell	Vertebrata		Terminally differentiated	FMA_7193		not defined in CTD: A warmth detecting thermoreceptor cell is	
9 vestibular hair cell	Vertebrata		Terminally differentiated	FMA_75715 ; FMA_50594			
10 pillar cell	Mammalia		Terminally differentiated			not defined in the CTD: A taste receptor cell is	
11 taste receptor cell	Animalia	vertebrata	Terminally differentiated	FMA_54825			
12 blue sensitive photoreceptor cell	Animalia		Terminally differentiated	retina of the eye			
13 green sensitive photoreceptor cell	Animalia		Terminally differentiated	retina of the eye			
14 R1 photoreceptor cell	Arthropoda		Terminally differentiated	ommatidia (compound eye)			
15 R2 photoreceptor cell	Arthropoda		Terminally differentiated	ommatidia (compound eye)			
16 R3 photoreceptor cell	Arthropoda		Terminally differentiated	ommatidia (compound eye)			
17 R4 photoreceptor cell	Arthropoda		Terminally differentiated	ommatidia (compound eye)			
18 R5 photoreceptor cell	Arthropoda		Terminally differentiated	ommatidia (compound eye)			
19 R6 photoreceptor cell	Arthropoda		Terminally differentiated	ommatidia (compound eye)			
20 R7 photoreceptor cell	Arthropoda		Terminally differentiated	ommatidia (compound eye)			
21 R8 photoreceptor cell	Arthropoda		Terminally differentiated	ommatidia (compound eye)			
22 rod sensitive photoreceptor cell	Animalia		Terminally differentiated	retina of the eye			
23 retinal cone cell	Vertebrata		Terminally differentiated	retina of the eye			
24 retinal rod cell	Vertebrata		Terminally differentiated	retina of the eye			
25 rodless photoreceptor cell	Vertebrata		Terminally differentiated	retina of the eye			
26 unipolar neuron	Animalia		Terminally differentiated	nervous system		not defined in CTD: A unipolar neuron is a	
27 enterocyte	Animalia		Terminally Differentiated	intestine			
28 gut absorptive cell	Animalia		Terminally differentiated	FMA_72289	n/a	PATO term is villiform - or a new term having	
29 chromophobe cell	Animalia		Terminally differentiated				
30 brown fat cell	Animalia		Terminally Differentiated				
31 white fat cell	Animalia		Terminally Differentiated				
32 Kupffer cell	Animalia		Terminally Differentiated	FMA_7197			
33 border follicle cell	Animalia		Terminally Differentiated	Follicle			
34 trichoblast	Embryophyta		Terminally Differentiated				
35 male germ line stem cell (sensu Nematoda and Protostomia)	Animalia		PATO:0001401				
36 male germ line stem cell (sensu Vertebrata)	Animalia		PATO:0001400	Male Reproductive System			
37 primary spermatocyte	Animalia		Terminally Differentiated	Male Reproductive System			
38 secondary spermatocyte	Animalia		Terminally Differentiated	Male Reproductive System			
39 sperm	Animalia		PATO:0001400				
40 spermaticid	Animalia		PATO:0001400				
41 spermatogonium	Animalia		PATO:0001400				
42 female germ line stem cell (sensu Nematoda and Protostomia)	Animalia		PATO:0001402				
43 female germ line stem cell (sensu Vertebrata)	Animalia		PATO:0001402				
44 primary oocyte	Animalia		PATO:0001399				
45 secondary oocyte	Animalia		PATO:0001399				
46 egg	Animalia		PATO:0001399				
47 egg cell	Embryophyta		PATO:0001399				
48 nurse cell	Animalia		Terminally Differentiated				
49 sperm cell (sensu Viridiplantae)	Animalia		Terminally Differentiated				
50 axial cell	Viridiplantae		Terminally differentiated	xylem	n/a	An axial cell is a secondary vascular cell that develops from the fusiform cambium. Mature sclerenchyma cell thicken and die.	

- A popular tool for data management
- Good at collecting regular data
- Users from many domains

# Related work

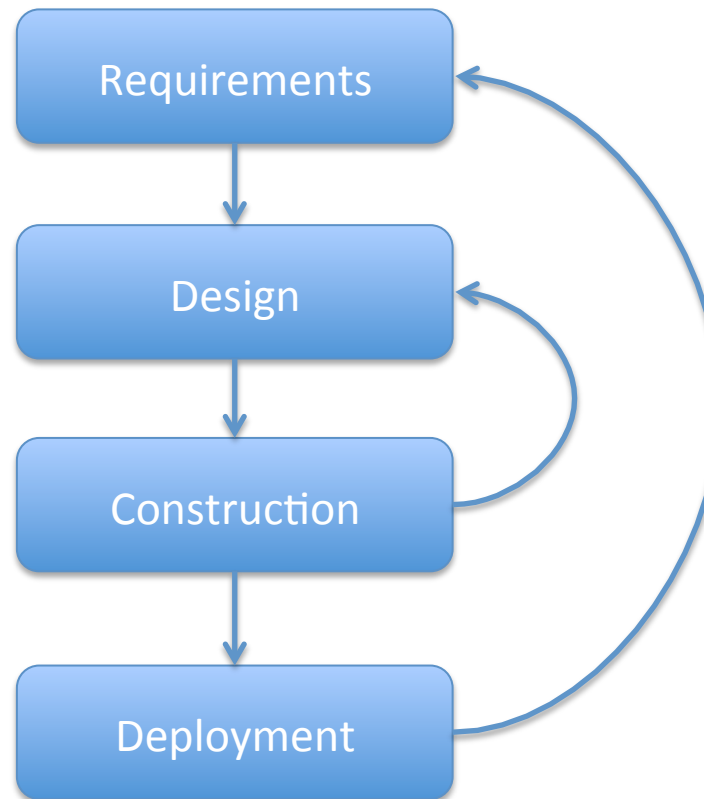
- Cell type ontology normalisation
- OBI and Quick Term Templates
- ICD-11



The image displays two Microsoft Excel spreadsheets. The top spreadsheet, titled 'ISWC2010OBIExample.xls', shows a table with two columns: 'Analyte label' and 'Analyte ID'. The bottom spreadsheet, titled 'Ch IV-Endocrine diseases E00-E3-2-9-mod.xls', shows a detailed ICD-11 classification for endocrine diseases, with columns for 'ICD category', 'ICD code', 'Linearization P', 'Synonyms', 'Definition', 'Exclusions', 'Inclusions', 'Note', 'Comment', and 'Class name'.

ICD category	ICD code	Linearization P	Synonyms	Definition	Exclusions	Inclusions	Note	Comment	Class name
Disorders of the thyroid gland and thyroid hormones system	E00-E07								E00-E07
Congenital hypothyroidism				Congenital hypothyroidism is the principle cause of preventab					
Congenital hypothyroidism due to iodine deficiency	E00			Congenital iodine-deficiency syndrome					
Congenital iodine-deficiency syndrome, neurological type	E00.0			Endemic hypothyroidism, neurological type					
Congenital iodine-deficiency syndrome, myxoedematous type	E00.1			Endemic hypothyroidism, myxoedematous type					
Congenital iodine-deficiency syndrome, mixed type	E00.2			Endemic hypothyroidism, mixed type					
Congenital iodine-deficiency syndrome, unspecified	E00.9			Congenital iodine-deficiency hypothyroidism NOS					
Permanent congenital hypothyroidism									
Permanent congenital hypothyroidism with diffuse goitre	E03.0			Familial thyroid dysmorphogenesis   Congenital hypothyroidism du					
Primary congenital hypothyroidism due to impaired	E07.1								
Pendred syndrome				Conductive anc Hypothyroidism   Pendred syndrome is characterized by the association of conge					
Hypothyroidism due to peripheral resistance to thyroid hormones									
Permanent congenital hypothyroidism without goitre	E03.1								
Primary congenital hypothyroidism due to a developmental anomaly				Thyroid dysgenesis					
Thyroid ectopia									
Thyroid agenesis				Athyreosis					
Thyroid hemiagenesis									
Thyroid hypoplasia									
Primary congenital hypothyroidism due to TSH receptor mutations				Resistance to TSH binding or signalling					

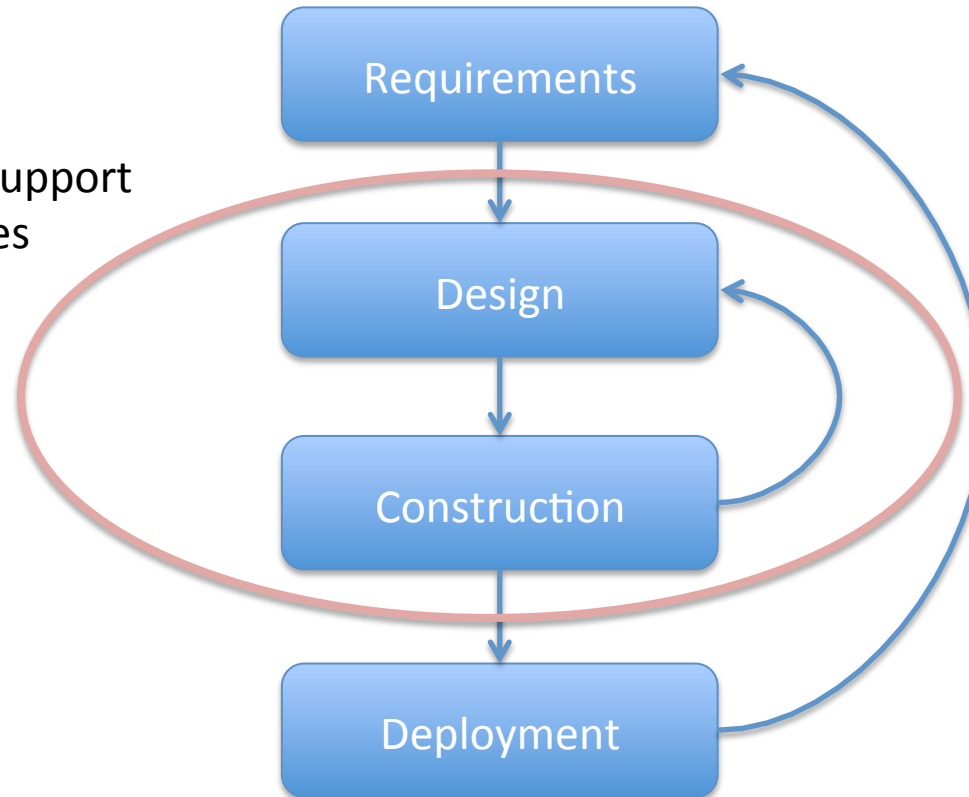
# Typical Stages in Ontology development



# Typical Stages in Ontology development



Populous support  
these stages



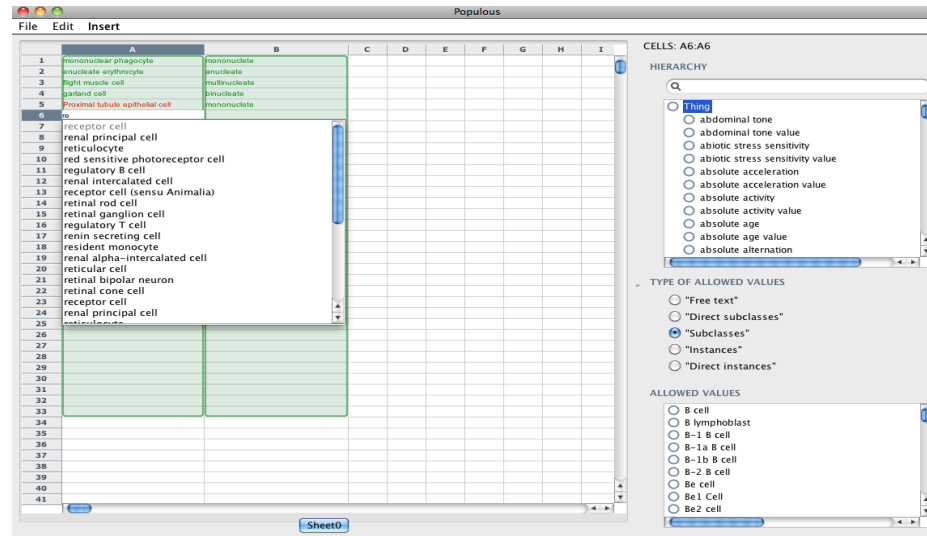




# Previous work

- RDF based tools
  - Excel2RDF, Convert2RDF, RDF123
- Protégé 4.0 plugins
  - Excel Importer, Matrix Plugin
- Protégé 3.4
  - Mapping Master (Protégé 3)

# Populous



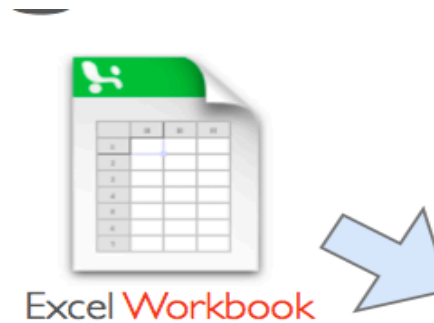
<http://www.e-lico.eu/populous>

- Generic tool for populating ontology templates
- Spreadsheet style interface
- Supports validation at the point of data entry
- Expressive Pattern language for OWL Ontology generation

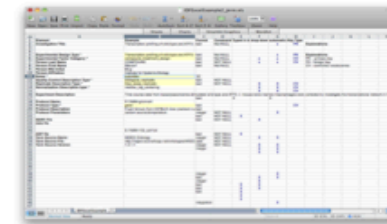
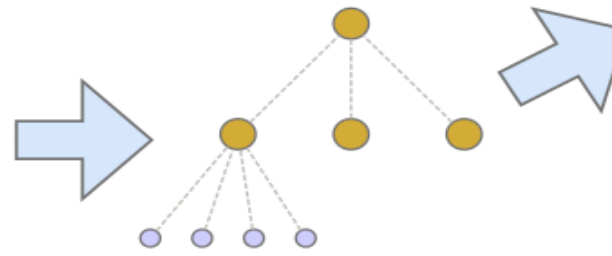
# Built on top of RightField

<http://www.sysmo-db.org/rightfield>

- Semantic Annotation by Stealth



Ontology



Terms Embedded into  
Excel Workbook



# Export to Excel

IDFExcelExample2\_jerm.xls

New Open Save Print Import Copy Paste Format Undo Redo AutoSum Sort A-Z Sort Z-A Gallery Toolbox Zoom Help

	A	B	C	D	E	F	G	H	I
	Element	Example	Format	Constraint	Typed in b	drop-down	automatic	Key	Type
	Investigation Title	Transcription profiling of wild-type and ATF3 -	text	Not NULL			x	PK	Explanations
5	Experimental Design Type *	Transcription profiling of wild-type and ATF3 -	text	Not NULL			x	PK	Explanations
6	Experimental Factor Category *	compound_treatment_design	text	Not NULL		x	x	CV	PK - primary k
7	Person Last Name	COMPOUND	text	NOT NULL		x	x	CV	FK - foreign ke
8	Person First Name	Marzolf	text	Not NULL					CV - controlled
9	Person Mid Initial	Bruz							
10	Person Affiliation	Institute for Systems Biology							
11	Roles *	curator							
12	Quality Control Description Type *	biological_replicate	text	NOT NULL					
13	Replicate Description Type *	Dye_swap_replicate	text	NOT NULL		x	x	CV	
14	Normalization Description type *	median_log_centering				x	x	CV	
15						x	x	CV	
16	Experiment Description	Time course data from lipopolysaccharide-stimulated wild-type and ATF3 -/- mouse bone marrow macrophages were collected to inv	text		X		X		
17									
18	Protocol Name	P-TABM-gilchrist1							
19	Protocol Type *	grow	text						CV
20	Protocol Description	Flush femurs from C57BL/6 mice (Jackson L	text						
21	Protocol Parameters	carbon source,temperature	integer	NOT NULL				X	
22			text	NOT NULL	X				
23	SDRF File		text	NOT NULL		X			
24	data file								
25									
26		E-TABM-102_sdrf.txt							
27	ADF file		text	NOT NULL	X				
28	Term Source Name	MGED Ontology	integer	NOT NULL		X	X		
29	Term Source File	http://mged.sourceforge.net/ontologies/MGED	date	NOT NULL					
30	Term Source Version	1.3.1.1	integer	NOT NULL		X	X		
31			integer			X	X		
32						X	X		
33									
34									
35									
36									
37									
38			integer				X		
39			text			X	X		
40			integer			X			

Normal View Ready Sum=0 SCRL

# Excel Validations

The screenshot shows an Excel spreadsheet titled 'IDFExcelExample2\_jerm.xls'. The spreadsheet has columns A through G and rows 1 through 31. The 'Roles' column (B) is highlighted, and a dropdown menu is open, showing a list of roles. The roles listed are: array\_manufacturer, biomaterial\_provider, biosequence\_provider, consortium\_member, consultant, curator, data\_analyst, data\_coder, funder, hardware\_manufacturer, institution, investigator, software\_manufacturer, and submitter (which is checked). The spreadsheet also shows other columns like 'Example', 'Format', and 'Constraint'.

Row	Column A	Column B	Column C	Column D	Column E	Column F	Column G
1	Element	Example	Format	Constraint	Typed in b	drop-down	automatic
2	Investigation Title	Transcription profiling of wild-type and ATF3 -	text	Not NULL			x
3							
4							
5	Experimental Design Type *	Transcription profiling of wild-type and ATF3 -	text	Not NULL			x
6	Experimental Factor Category *	compound_treatment_design	text	Not NULL		x	x
7	Person Last Name	COMPOUND	text	NOT NULL		x	x
8	Person First Name	Marzolf	text	Not NULL			
9	Person Mid Initial						
10	Person Affiliation	Institute for Systems Biology					
11	Roles *	submitter					
12	Quality Control Description Type *		text	NOT NULL			
13	Replicate Description Type *		text	NOT NULL			
14	Normalization Description type *					x	x
15						x	x
16	Experiment Description						
17							
18	Protocol Name						
19	Protocol Type *		text				
20	Protocol Description		text				
21	Protocol Parameters						
22			integer	NOT NULL			x
23	SDRF File		text	NOT NULL			
24	data file		text	NOT NULL		x	
25							
26							
27	ADF file		text	NOT NULL		x	
28	Term Source Name		integer	NOT NULL		x	x
29	Term Source File		date	NOT NULL			
30	Term Source Version		integer	NOT NULL		x	x
31			integer			x	x
32						x	x
33						x	x
34						x	x
35						x	x
36						x	x
37						x	x
38						x	x
39						x	x
40						x	x
41						x	x
42						x	x
43						x	x
44						x	x
45						x	x
46						x	x
47						x	x
48						x	x
49						x	x
50						x	x
51						x	x
52						x	x
53						x	x
54						x	x
55						x	x

# Creating Templates

RightField

File Edit Tools Sheet

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1																				
2																				
3																				
4																				
5																				
6																				
7																				
8																				
9																				
10																				
11																				
12																				
13																				
14																				
15																				
16																				
17																				
18																				
19																				
20																				
21																				
22																				
23																				
24																				
25																				
26																				
27																				
28																				
29																				
30																				
31																				
32																				
33																				
34																				
35																				
36																				
37																				
38																				
39																				
40																				
41																				
42																				
43																				
44																				
45																				
46																				
47																				

Sheet0

HIERARCHY

Q

No ontologies loaded

TYPE OF ALLOWED VALUES

- "Free text"
- "Direct subclasses"
- "Subclasses"
- "Instances"
- "Direct instances"
- "All classes"

Select ontology... ▾

Apply

ALLOWED VALUES

Any

# Creating Templates

RightField

File Edit Tools Sheet

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47

A B C D E F G H I J K L M N O P Q R S

*Load from file or directly from BioPortal*

Open from BioPortal repository

- C. elegans development (OBO)
- C. elegans gross anatomy (OBO)
- C. elegans phenotype (OWL)
- Cancer Research and Management ACGT Master Ontology (OBO)
- Cardiac Electrophysiology Ontology (OWL-DL)
- Cell Behavior Ontology (null)
- Cell Cycle Ontology (OBO)
- Cell line ontology (OWL-DL)
- Cell Line Ontology (LEXGRID-XML)
- Cell line ontology (OWL)
- Cell type (OBO)**
- Cereal plant development (OBO)
- Cereal plant gross anatomy (OBO)
- Cereal plant trait (OBO)
- Chemical entities of biological interest (OWL)

Filter:

Cancel OK

HIERARCHY

No ontologies loaded

TYPE OF ALLOWED VALUES

- "Free text"
- "Direct subclasses"
- "Subclasses"
- "Instances"
- "Direct instances"
- "All classes"

Select ontology... ▾

Apply

ALLOWED VALUES

Any

Sheet0

# Creating Templates

RightField

File Edit Tools Sheet

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1																			
2																			
3																			
4																			
5																			
6																			
7																			
8																			
9																			
10																			
11																			
12																			
13																			
14																			
15																			
16																			
17																			
18																			
19																			
20																			
21																			
22																			
23																			
24																			
25																			
26																			
27																			
28																			
29																			
30																			
31																			
32																			
33																			
34																			
35																			
36																			
37																			
38																			
39																			
40																			
41																			
42																			
43																			
44																			
45																			
46																			
47																			

Sheet0

*Ontology browser*

HIERARCHY

- Thing
  - adrenal medulla cell
  - cell
  - DbXref
  - Definition
  - dental papilla cell
  - glial cell
  - hair papilla cell
  - mesangial cell
  - mesenchyme condensation cell
  - muscle precursor cell
  - neuroepithelial cell
  - neuron associated cell
  - neuron associated cell (sensory/vegetative)

TYPE OF ALLOWED VALUES

- "Free text"
- "Direct subclasses"
- "Subclasses"
- "Instances"
- "Direct instances"
- "All classes"

Select ontology...

Apply

ALLOWED VALUES

Any



# Creating Templates

RightField

File Edit Tools Sheet

CELLS: A1:A32767

HIERARCHY

Thing

- adrenal medulla cell
- cell
- DbXref
- Definition
- dental papilla cell
- glial cell
- hair papilla cell
- mesangial cell
- mesenchyme condensation cell
- muscle precursor cell
- neuroepithelial cell
- neuron.associated cell

TYPE OF ALLOWED VALUES

- "Free text"
- "Direct subclasses"
- "Subclasses"
- "Instances"
- "Direct instances"
- "All classes"

Select ontology...

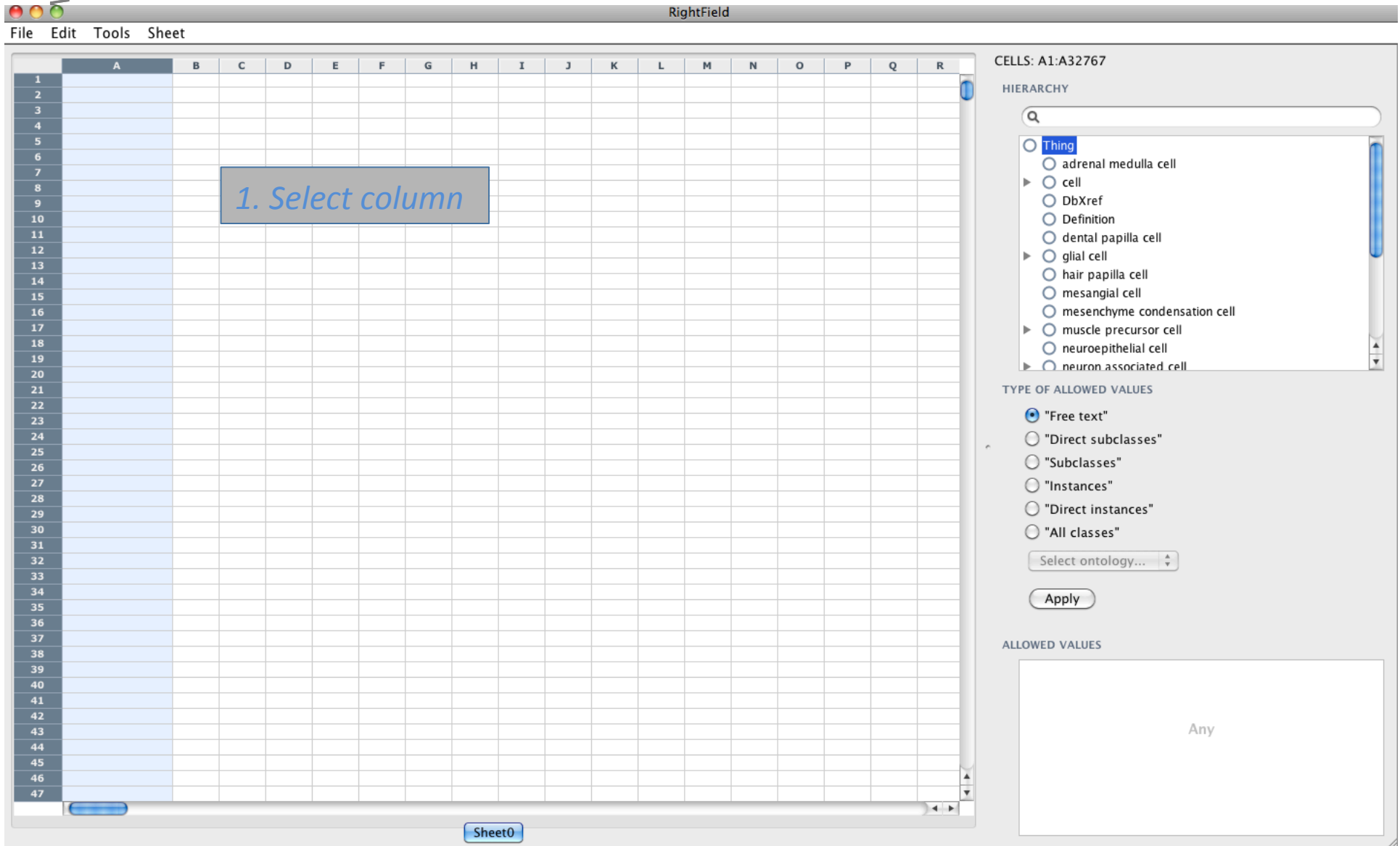
Apply

ALLOWED VALUES

Any

Sheet0

1. Select column



# Creating Templates

RightField

File Edit Tools Sheet

CELLS: A1:A32767

HIERARCHY

Thing

- adrenal medulla cell
- cell
- DbXref
- Definition
- dental papilla cell
- glial cell
- hair papilla cell
- mesangial cell
- mesenchyme condensation cell
- muscle precursor cell
- neuroepithelial cell
- neuron.associated cell

TYPE OF ALLOWED VALUES

- "Free text"
- "Direct subclasses"
- "Subclasses"
- "Instances"
- "Direct instances"
- "All classes"

Select ontology...

Apply

ALLOWED VALUES

Any

Sheet0

1. Select column

2. Select Class in Ontology

# Creating Templates

RightField

File Edit Tools Sheet

CELLS: A1:A32767

HIERARCHY

○ Thing  
○ adrenal medulla cell  
○ cell  
○ DbXref  
○ Definition  
○ dental papilla cell  
○ dial cell

TYPE OF ALLOWED VALUES

"Free text"  
 "Direct subclasses"  
 "Subclasses"  
 "Instances"  
 "Direct instances"  
 "All classes"

Select ontology... ▾

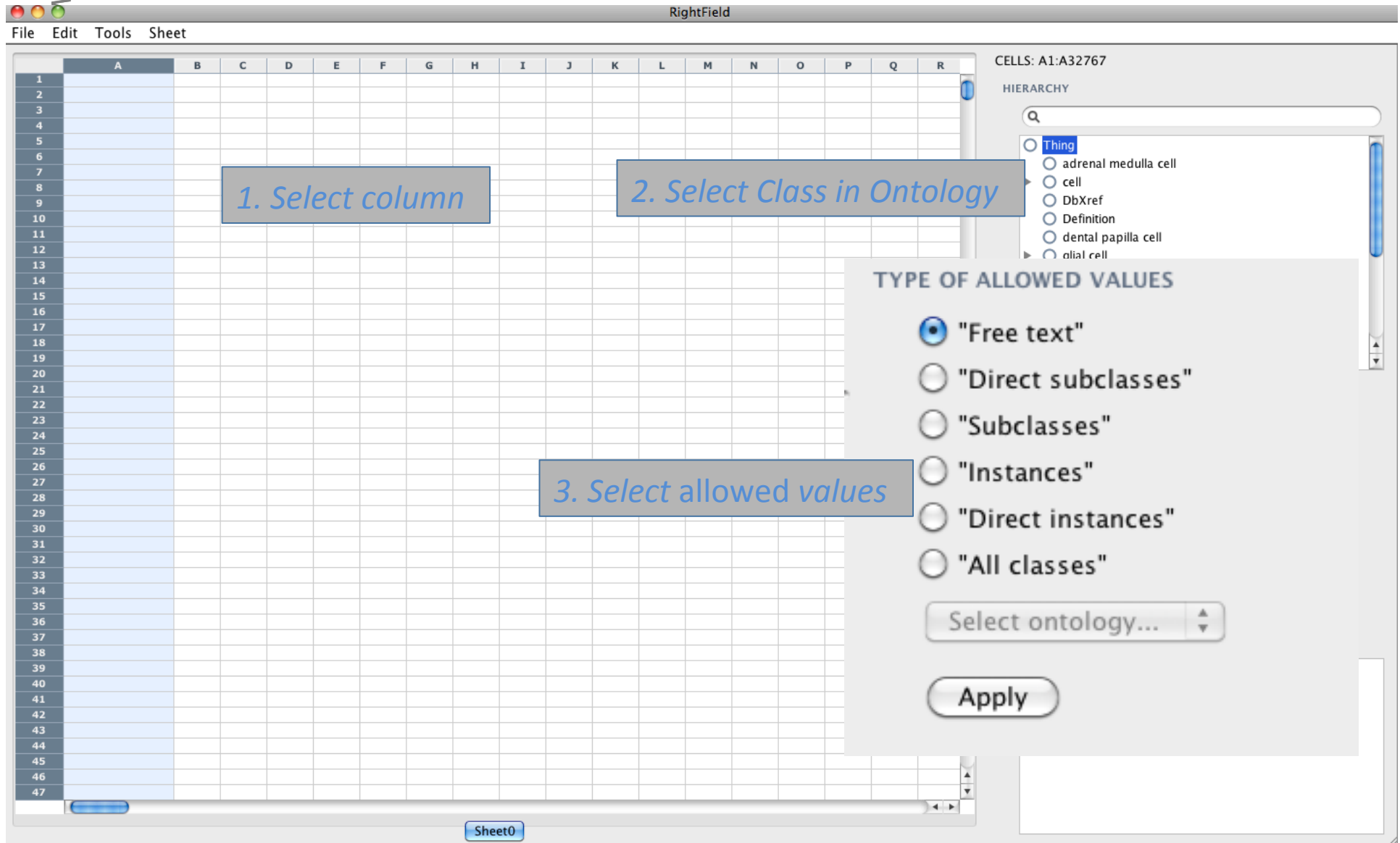
Apply

Sheet0

1. Select column

2. Select Class in Ontology

3. Select allowed values



# Creating Templates

RightField

File Edit Tools Sheet

CELLS: A1:A32767

HIERARCHY

Q

- Thing
  - adrenal medulla cell
  - cell
  - DbXref
  - Definition
  - dental papilla cell
  - glial cell
  - hair papilla cell
  - mesangial cell
  - mesenchyme condensation cell
  - muscle precursor cell
  - neuroepithelial cell
  - neuron.associated cell

TYPE OF ALLOWED VALUES

- "Free text"
- "Direct subclasses"
- "Subclasses"
- "Instances"
- "Direct instances"
- "All classes"

Select ontology...

Apply

ALLOWED VALUES

- B cell
- B-1 B cell
- B-1a B cell
- B-1b B cell
- B-2 B cell
- Be cell
- Be1 Cell
- Be2 cell
- Bm1 B cell

Sheet0

# Editing Templates

Populous

File Edit Insert

	A	B	C	D	E	F	G	H	I
1	mononuclear phagocyte	mononucleate							
2	enucleate erythrocyte	anucleate							
3	flight muscle cell	multinucleate							
4	garland cell	binucleate							
5	Proximal tubule epithelial cell	mononucleate							
6	receptor cell								
7	renal principal cell								
8	reticulocyte								
9	red sensitive photoreceptor cell								
10	regulatory B cell								
11	renal intercalated cell								
12	receptor cell (sensu Animalia)								
13	retinal rod cell								
14	retinal ganglion cell								
15	regulatory T cell								
16	renin secreting cell								
17	resident monocyte								
18	renal alpha-intercalated cell								
19	reticular cell								
20	retinal bipolar neuron								
21	retinal cone cell								
22	receptor cell								
23	renal principal cell								
24	reticulocyte								
25									
26									
27									
28									
29									
30									
31									
32									
33									
34									
35									
36									
37									
38									
39									
40									
41									

CELLS: A6:A6

HIERARCHY

Thing

- abdominal tone
- abdominal tone value
- abiotic stress sensitivity
- abiotic stress sensitivity value
- absolute acceleration
- absolute acceleration value
- absolute activity
- absolute activity value
- absolute age
- absolute age value
- absolute alternation

TYPE OF ALLOWED VALUES

- "Free text"
- "Direct subclasses"
- "Subclasses"
- "Instances"
- "Direct instances"

ALLOWED VALUES

- B cell
- B lymphoblast
- B-1 B cell
- B-1a B cell
- B-1b B cell
- B-2 B cell
- Be cell
- Be1 Cell
- Be2 cell

Sheet0

Label rendering

Tab completion

Syntax Highlighting

Multi-value cells

# Editing Templates

Populous

File Edit Insert

	A	B	C	D	E	F	G	H	I
1	mononuclear phagocyte	mononucleate							
2	enucleate erythrocyte	anucleate							
3	flight muscle cell	multinucleate							
4	garland cell	binucleate							
5	Proximal tubule epithelial cell	mononucleate							
6	receptor cell								
7	renal principal cell								
8	reticulocyte								
9	red sensitive photoreceptor cell								
10	regulatory B cell								
11	renal intercalated cell								
12	receptor cell (sensu Animalia)								
13	retinal rod cell								
14	retinal ganglion cell								
15	regulatory T cell								
16	renin secreting cell								
17	resident monocyte								
18	renal alpha-intercalated cell								
19	reticular cell								
20	retinal bipolar neuron								
21	retinal cone cell								
22	receptor cell								
23	renal principal cell								
24	reticulocyte								
25									
26									
27									
28									
29									
30									
31									
32									
33									
34									
35									
36									
37									
38									
39									
40									
41									

CELLS: A6:A6

HIERARCHY

Thing

- abdominal tone
- abdominal tone value
- abiotic stress sensitivity
- abiotic stress sensitivity value
- absolute acceleration
- absolute acceleration value
- absolute activity
- absolute activity value
- absolute age
- absolute age value
- absolute alternation

TYPE OF ALLOWED VALUES

- free text"
- direct subclasses"
- subclasses"
- instances"
- direct instances"

VALUES

- cell
- lymphoblast
- 1 B cell
- 1a B cell
- B-1b B cell
- B-2 B cell
- Be cell
- Be1 Cell
- Be2 cell

Label rendering

Tab completion

Syntax Highlighting

Multi-value cells

glomerular capillary endothelium

afferent arteriole

afferent arteriole | arteriole endothelium

part of afferent arteriole forming juxtaglomerular complex

afferent arteriole | arteriole smooth muscle

efferent arteriole

efferent arteriole | arteriole endothelium

efferent arteriole | arteriole smooth muscle

renal proximal tubule

Sheet0

# Ontology Pre-Processing Language

Pattern

```
A cell type is equivalent to a cell type  
that is part of some anatomy
```

# Ontology Pre-Processing Language

Pattern

A cell type is equivalent to a cell type  
that is part of some anatomy

OPPL Script

```
?cell:CLASS,  
?anatomyPart:CLASS,  
?anatomy:CLASS =  
    (CL:0000000 part_of some ?anatomyPart)  
  
BEGIN  
ADD ?cell equivalentTo ?anatomy  
END;
```

Variables

Create axioms



# Ontology Pre-Processing Language

Pattern

```
A cell type is equivalent to a cell type
that is part of some anatomy
```

Variable mapper

```
?cell -> 'Kidney Cell' [CL:0003523]
?anatomyPart -> 'Kidney' [FMA:629093]
```

OPPL Script

```
?cell:CLASS,
?anatomyPart:CLASS,
?anatomy:CLASS =
    (CL:0000000 part_of some ?anatomyPart)

BEGIN
ADD ?cell equivalentTo ?anatomy
END;
```



# Resulting OWL axioms

## Example

A 'Kidney Cell' is equivalent to a cell that is part of the 'Kidney'

## Generated OWL (Manchester Syntax)

```
Class: CL:0003523
```

```
Annotation:
```

```
rdfs:label 'Kidney Cell'
```

```
EquivalentTo:
```

```
CL:0000000 and OBO_REL:part_of some FMA:629093
```

# OPPL Wizzard

Populous - /Users/simon/Dropbox/elico/SWAT4LS/images/cellexample.xls

File Edit Insert

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	mononuclear phagocyte	mononucleate													
2	anucleate erythrocyte	anucleate													
3	flight muscle cell	multinucleate													
4	garland cell	binucleate													
5	Proximal tubule epithelial cell	mononucleate													
6	ren														
7															
8															
9															
10															
11															
12															
13															
14															
15															
16															
17															
18															
19															
20															
21															
22															
23															
24															
25															
26															
27															
28															
29															
30															
31															
32															
33															
34															
35															
36															
37															
38															
39															
40															
41															
42															
43															
44															
45															
46															
47															

CELLS: B1:B32767

HIERARCHY

Thing

- abdominal tone
- abdominal tone value
- abiotic stress sensitivity
- abiotic stress sensitivity value
- absolute acceleration
- absolute acceleration value
- absolute activity
- absolute activity value
- absolute age
- absolute age value
- absolute alternation
- absolute alternation value
- absolute amplitude
- absolute amplitude value
- absolute angle value
- absolute auditory ability
- absolute auditory ability value

TYPE OF ALLOWED VALUES

- Free text
- Direct subclasses
- Subclasses
- Instances
- Direct instances

ALLOWED VALUES

Any

Pattern Select

OPPL Pattern editor

Pattern name: CellNucleationPattern

```
?cell:CLASS, ?nucleation:CLASS
BEGIN
ADD ?cell SubClassOf hasNucleation some ?nucleation
END;
```

Cancel OK

New Pattern

Go Back Finish Cancel

Sheet0

# Variable mapping

```
?cell:CLASS,
?anatomyPart:CLASS,
?partOfRestriction:CLASS = CL_000000 and part_of some ?anatomyPart,
?anatomyIntersection:CLASS = createIntersection(?
partOfRestriction.VALUES)
BEGIN
ADD ?cell equivalentTo ?anatomyIntersection
END;
```

Populous - /Users/simon/Documents/e-lico/kupo\_cells\_sept10\_Populous.xls

File Edit Insert

	A	B	C	D
1	Cell Term	cell label	part_of	participates_in
2	kidney cell	renal cell	kidney	
3	kidney epithelial cell		kidney epithelium	
4	renal tubule epithelial cell		renal tubule epithelium	
5	kidney glomerular epithelial cell		kidney glomerular epithelium	
6	renal tubule cell		renal tubule	
7	kidney cortex cell	renal cortex cell	kidney cortex	
8	renal cortex tubule cell		renal cortex tubule	
9	kidney medulla cell	renal medullary cell	kidney medulla	
10	kidney outer medulla cell	renal outer	outer renal medulla	
11	kidney inner medulla cell	renal inner medullary	inner renal medulla	
12	inner renal medulla loop of henle cell		inner renal medulla loop of henle	
13	juxtaglomerular complex cell	juxtaglomerular	juxtaglomerular complex	regulation of glomerular filtration, regulation
14	kidney blood vessel cell	renal blood vessel	kidney blood vessel	blood circulation
15	kidney arterial blood vessel cell	renal arterial blood	kidney arterial blood vessel	
16	kidney capillary endothelial cell	renal capillary cell	kidney capillary, capillary endothelium	
17	kidney venous blood vessel cell	ous blood vessel cell	kidney venous blood vessel	
18	renal corpuscle cell	cell of the renal	renal corpuscle	phagocytosis, extracellular matrix constituent
19	mesangial cell		mesangium	
20	glomerular mesangial cell		glomerular mesangium	
21	extraglomerular mesangial cell		extraglomerular mesangium	
22	podocyte	visceral epithelial	glomerular visceral epithelium	glomerular filtration, regulation of glomerular
23	bowmans capsule epithelial cell	epithelial cell of the	bowmans capsule	anatomical structure arrangement
24	parietal epithelial cell	glomerular parietal	glomerular parietal epithelium	
25	glomerular cell		glomerulus	
26	glomerular capillary endothelial cell	glomerular capillary	glomerular capillary endothelium	glomerular filtration, regulation of glomerular
27	renal afferent arteriole cell	afferent arteriole cell	afferent arteriole	regulation of glomerular filtration
28	renal afferent arteriole endothelial cell	afferent arteriole	afferent arteriole, arteriole endothelium	
29	juxtaglomerular cell		part of afferent arteriole forming	renin secretion into blood stream, detection
30	renal afferent arteriole smooth muscle	afferent arteriole	afferent arteriole, arteriole smooth	
31	renal efferent arteriole cell	efferent arteriole cell	efferent arteriole	regulation of glomerular filtration
32	renal efferent arteriole endothelial cell	efferent arteriole	efferent arteriole, arteriole endothelium	
33	renal efferent arteriole smooth muscle	efferent arteriole	efferent arteriole, arteriole smooth	
34	proximal tubule epithelial cell	PTEC, proximal	renal proximal tubule	renal sodium ion absorption, potassium ion
35	proximal convoluted tubule epithelial		proximal convoluted tubule	
36	proximal straight tubule cell		proximal straight tubule	
37	loop of henle epithelial cell	henle's loop	loop of henle	extracellular matrix constituent secretion
38	loop of henle ascending limb	henle's loop	loop of henle ascending limb	renal sodium ion absorption, potassium ion
39	loop of henle thick ascending limb	henle's loop thick	loop of henle ascending limb thick	
40	loop of henle thin ascending limb	henle's loop thin	loop of henle ascending limb thin	
41	loop of henle medullary thick	henle's loop	distal straight tubule pramacula	

CELLS: B17:B17

HIERARCHY

Thing

- abdominal tone
- abdominal tone value
- abiotic stress sensitivity
- abiotic stress sensitivity value
- absolute acceleration
- absolute acceleration value
- absolute activity
- absolute activity value
- absolute age
- absolute age value
- absolute alternation

TYPE OF ALLOWED VALUES

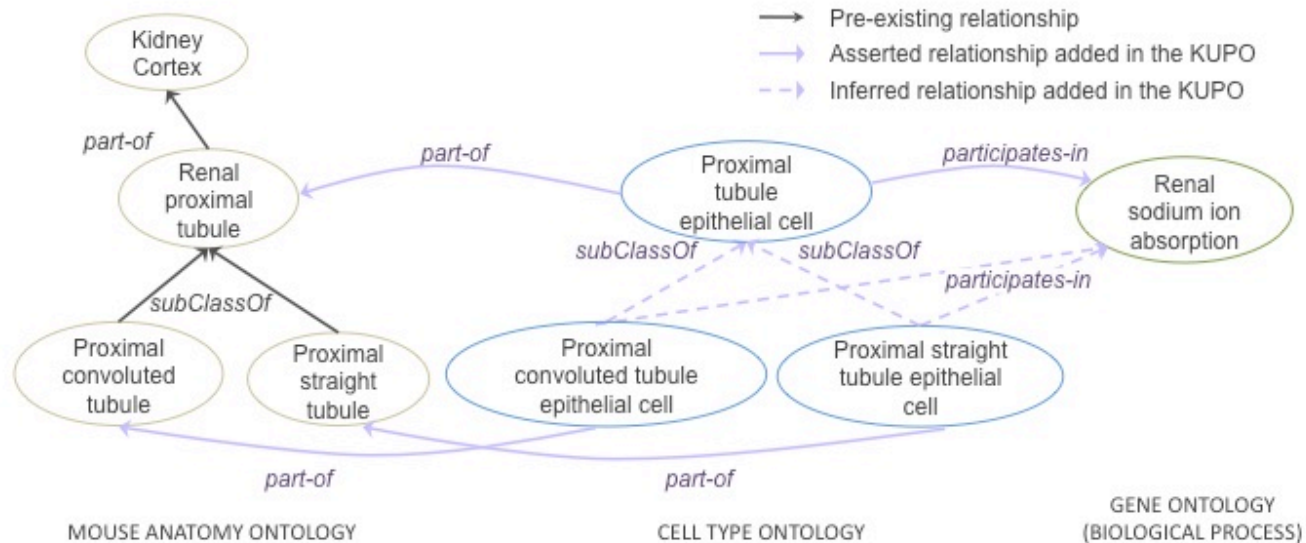
- "Free text"
- "Direct subclasses"
- "Subclasses"
- "Instances"
- "Direct instances"

ALLOWED VALUES

Any

Sheet1

# Use case – Kidney and Urinary Pathway Ontology (KUPO)



# KUPO in Populous

Cell type Ontology

Mouse anatomy Ontology

Biological Process (GO)

Populous - /Users/simon/Documents/e-lico/kupo\_cells\_sept10\_Populous.xls

File Edit Insert

	A	B	C	D
1	Cell Term	cell label	part_of	participates_in
2	kidney cell	renal cell	kidney	
3	kidney epithelial cell		kidney epithelium	
4	renal tubule epithelial cell		renal tubule epithelium	
5	kidney glomerular epithelial cell		kidney glomerular epithelium	
6	renal tubule cell		renal tubule	
7	kidney cortex cell	renal cortex cell	kidney cortex	
8	renal cortex tubule cell		renal cortex tubule	
9	kidney medulla cell	renal medullary cell	kidney medulla	
10	kidney outer medulla cell	renal outer	outer renal medulla	
11	kidney inner medulla cell	renal inner medullary	inner renal medulla	
12	inner renal medulla loop of henle cell		inner renal medulla loop of henle	
13	juxtglomerular complex cell	juxtglomerular	juxtglomerular complex	regulation of glomerular filtration, regulation of blood circulation
14	kidney blood vessel cell	renal blood vessel	kidney blood vessel	
15	kidney arterial blood vessel cell	renal arterial blood	kidney arterial blood vessel	
16	kidney capillary endothelial cell	renal capillary cell	kidney capillary, capillary endothelium	
17	kidney venous blood vessel cell	renal venous blood vessel cell of the renal	kidney venous blood vessel, renal corpuscle	
18	renal corpuscle cell		renal corpuscle	
19	mesangial cell		mesangium	phagocytosis, extracellular matrix constituent
20	glomerular mesangial cell		glomerular mesangium	
21	extraglomerular mesangial cell		extraglomerular mesangium	
22	podocyte	visceral epithelial	glomerular visceral epithelium	glomerular filtration, regulation of glomerular anatomical structure arrangement
23	bowmans capsule epithelial cell	epithelial cell of the	bowmans capsule	
24	parietal epithelial cell	glomerular parietal	glomerular parietal epithelium	
25	glomerular cell		glomerulus	
26	glomerular capillary endothelial cell	glomerular capillary	glomerular capillary endothelium	glomerular filtration, regulation of glomerular anatomical structure arrangement
27	renal afferent arteriole cell	afferent arteriole cell	afferent arteriole	regulation of glomerular filtration
28	renal afferent arteriole endothelial cell	afferent arteriole	afferent arteriole, arteriole endothelium	
29	juxtglomerular cell		part of afferent arteriole forming	renin secretion into blood stream, detection
30	renal afferent arteriole smooth muscle	afferent arteriole	afferent arteriole, arteriole smooth	
31	renal efferent arteriole cell	efferent arteriole cell	efferent arteriole	regulation of glomerular filtration
32	renal efferent arteriole endothelial cell	efferent arteriole	efferent arteriole, arteriole endothelium	
33	renal efferent arteriole smooth muscle	efferent arteriole	efferent arteriole, arteriole smooth	
34	proximal tubule epithelial cell	PTEC, proximal	renal proximal tubule	renal sodium ion absorption, potassium ion
35	proximal convoluted tubule epithelial		proximal convoluted tubule	
36	proximal straight tubule cell		proximal straight tubule	
37	loop of henle epithelial cell	henle's loop	loop of henle	extracellular matrix constituent secretion
38	loop of henle ascending limb	henle's loop	loop of henle ascending limb	renal sodium ion absorption, potassium ion
39	loop of henle thick ascending limb	henle's loop thick	loop of henle ascending limb thick	
40	loop of henle thin ascending limb	henle's loop thin	loop of henle ascending limb thin	
41	loop of henle medullary thick	henle's loop	distal straight tubule premacula	

CELLS: B17:B17

HIERARCHY

Thing

- abdominal tone
- abdominal tone value
- abiotic stress sensitivity
- abiotic stress sensitivity value
- absolute acceleration
- absolute acceleration value
- absolute activity
- absolute activity value
- absolute age
- absolute age value
- absolute alternation

TYPE OF ALLOWED VALUES

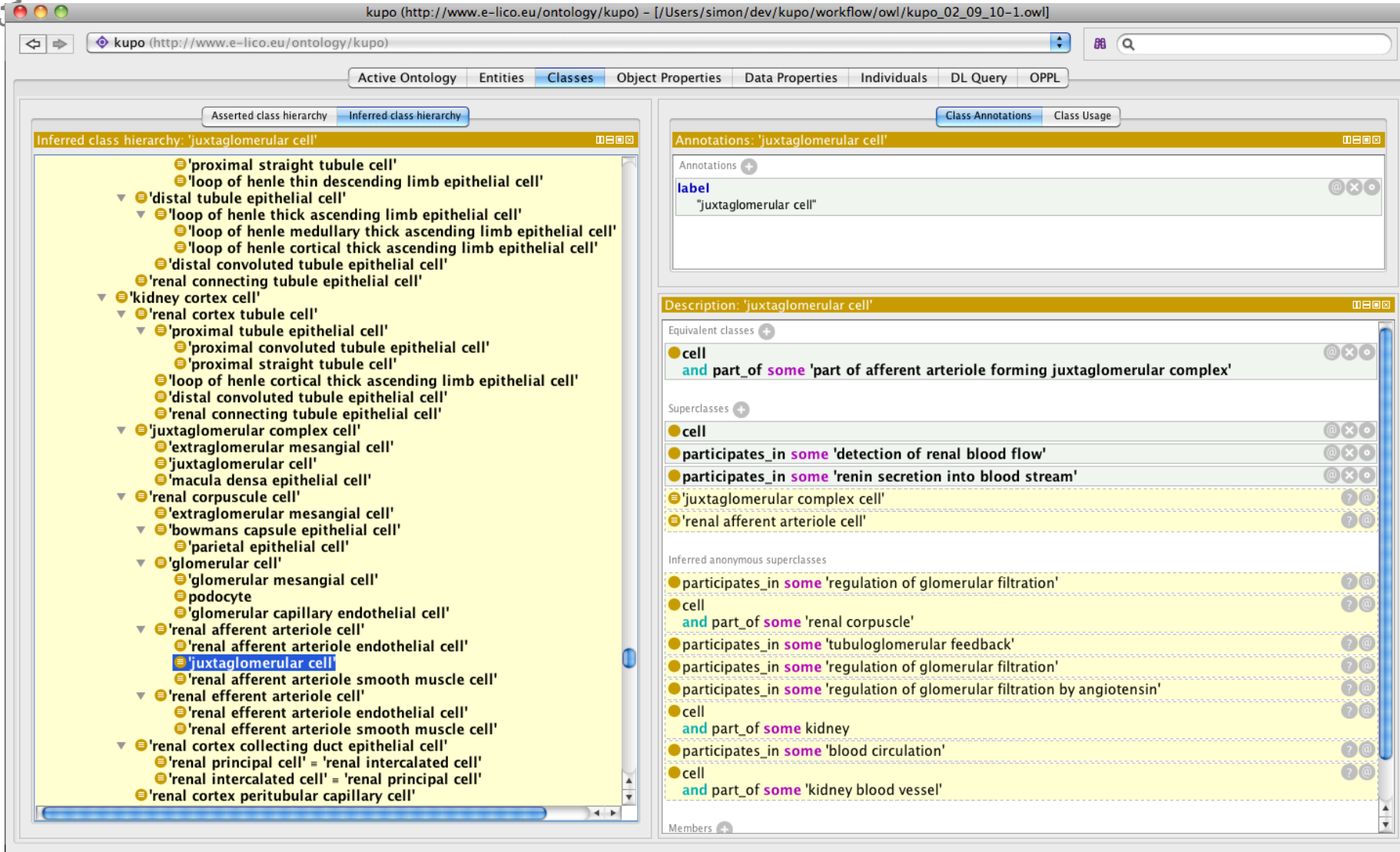
- "Free text"
- "Direct subclasses"
- "Subclasses"
- "Instances"
- "Direct instances"

ALLOWED VALUES

Any

Sheet1

# KUPO in OWL



The screenshot displays the Kupo ontology editor interface. The main window shows the 'Inferred class hierarchy' for the class 'juxtglomerular cell'. The hierarchy is as follows:

- 'proximal straight tubule cell'
- 'loop of henle thin descending limb epithelial cell'
- 'distal tubule epithelial cell'
  - 'loop of henle thick ascending limb epithelial cell'
    - 'loop of henle medullary thick ascending limb epithelial cell'
    - 'loop of henle cortical thick ascending limb epithelial cell'
  - 'distal convoluted tubule epithelial cell'
- 'renal connecting tubule epithelial cell'
- 'kidney cortex cell'
  - 'renal cortex tubule cell'
    - 'proximal tubule epithelial cell'
      - 'proximal convoluted tubule epithelial cell'
      - 'proximal straight tubule cell'
    - 'loop of henle cortical thick ascending limb epithelial cell'
    - 'distal convoluted tubule epithelial cell'
    - 'renal connecting tubule epithelial cell'
  - 'juxtglomerular complex cell'
    - 'extraglomerular mesangial cell'
    - 'juxtglomerular cell'
    - 'macula densa epithelial cell'
  - 'renal corpuscle cell'
    - 'extraglomerular mesangial cell'
  - 'bowmans capsule epithelial cell'
    - 'parietal epithelial cell'
  - 'glomerular cell'
    - 'glomerular mesangial cell'
    - 'podocyte'
    - 'glomerular capillary endothelial cell'
  - 'renal afferent arteriole cell'
    - 'renal afferent arteriole endothelial cell'
    - 'juxtglomerular cell'
    - 'renal afferent arteriole smooth muscle cell'
  - 'renal efferent arteriole cell'
    - 'renal efferent arteriole endothelial cell'
    - 'renal efferent arteriole smooth muscle cell'
  - 'renal cortex collecting duct epithelial cell'
    - 'renal principal cell' = 'renal intercalated cell'
    - 'renal intercalated cell' = 'renal principal cell'
  - 'renal cortex peritubular capillary cell'

The right-hand pane shows the 'Class Annotations' for 'juxtglomerular cell', including a 'label' annotation with the value 'juxtglomerular cell'. Below this, the 'Description' pane shows the class definition: 'cell and part\_of some 'part of afferent arteriole forming juxtglomerular complex''. The 'Superclasses' pane lists several classes, including 'cell', 'participates\_in some 'detection of renal blood flow'', and 'participates\_in some 'renin secretion into blood stream''. The 'Inferred anonymous superclasses' pane shows a complex set of logical expressions involving 'participates\_in some' and 'and part\_of some' relationships.



# Limitations of approach

- Assumes regular patterns
  - Developing patterns is hard
  - Not always sensible
- Handling exceptions to patterns
  - Need special syntax, but how far do you go...
- Regular data (entity-per-row assumption)





# Future plans

- Mapping Master Support
- Additional syntax
- Ontology Alignment
- Template creation from Protégé OPPL patterns plugin.
- Building KUP knowledge base



# Summary

- Populous is for populating templates
- Focus on supporting domain experts
- Expressive pattern language for modeling OWL
- Abstract knowledge from the modeling
- Engaging the experts by stealth!



# Acknowledgments

- **RightField**  
Matthew Horridge, Katy Wolstencroft, Stuart Owen,  
Carole Goble
- **OPPL**  
Luigi Iannone, Mikel Aranguren, Alan Rector, Robert Stevens
- **KUPO**  
Robert Stevens, Julie Klein, Joost Schanstra
- **e-LICO**  
EU-FP7 Collaborative Project (2009-2012)  
Theme ICT-4.4: Intelligent Content and Semantics