

Populous: A tool for populating Templates for OWL ontologies

Simon Jupp¹, Matthew Horridge¹, Luigi Iannone¹, Julie Klein², Stuart Owen¹,
Joost Schanstra², Robert Stevens¹, and Katy Wolstencroft¹

¹ School of Computer Science, University of Manchester, Manchester, UK

² Institut National de la Santé et de la Recherche Médicale, Toulouse, France

Abstract. We present Populous, a tool for gathering content with which to populate an ontology. Domain experts need to add content, that is often repetitive in its form, but without having to tackle the underlying ontological representation. Populous presents users with a table based form in which columns are constrained to take values from particular ontologies; the user can select a concept from an ontology via its meaningful label to give a value for a given entity attribute. Populated tables are mapped to patterns that can then be used to automatically generate the ontology's content. Populous's contribution is in the knowledge gathering stage of ontology development. It separates knowledge gathering from the conceptualisation and also separates the user from the standard ontology authoring environments. As a result, Populous can allow knowledge to be gathered in a straight-forward manner that can then be used to do mass production of ontology content.

Keywords: Ontology, OWL, Spreadsheet, Template

1 Introduction

Ontologies are being developed to provide controlled vocabularies for the annotation of life science data. Annotating data with ontologies adds semantics to the data that can facilitate data integration and enrich data analysis [18], [5], [3]. For ontologies to have a faithful representation of a domain, experts from that domain must have input to the authoring process. There are barriers that prevent domain experts engaging in an ontology's development; in particular, the semantics of an ontology language, or the intricacies of the authoring tools. To address this issue we have developed Populous that allows users to contribute their knowledge by populating simple templates that are then transformed to statements in the underlying representation language.

When designing an ontology it is often the case that repeating patterns occur in the modelling. These patterns can be abstracted from the ontology and used to specify simple templates that could be populated by domain experts [7, 6, 2]. As an example of a pattern, consider an ontology about cells; eukaryotic cells can be classified as being either anucleate, mono-nucleate, binucleate or multinucleate.

We can abstract over this pattern to say that every cell can be classified by its nucleation. This pattern is repeated for all cell types; the only variables are the cell name and the value for its nucleation. We can now use this pattern to build a simple template that could be populated by a cytologist, without him or her ever knowing about the underlying ontological representation. This type of pattern is common in ontology development where you have one set of entities being described in terms of another set of entities [17].

The tabular layout provides a simple and intuitive form fill-in style of user interface for a user to populate such templates. Typically, each row corresponds to a set of related entities and each column represents the type of relationship. The intersection of row and column holds the ‘filler’ for the given entity’s relationship of that column’s type. By adopting templates ontology developers can separate the pattern from its population; this allows the domain expert to focus on the knowledge without the distraction of the knowledge representation language.

Templates are useful when data needs to be collected in a regular form. Applying constraints to the template reduces the amount of discrepancies in the input data. A common tool for collecting data in this form is the spreadsheet; spreadsheets provide a tabular interface, where columns and rows represent certain attributes, and individual cells capture the data. Tables help users to structure data in a logical way, that is useful for both its maintenance and processing. In ontology development spreadsheets can be used to gather and organise information about concepts and their relationships. Previous work in this area has focused on the transformation of data into ontologies, but little attention has been paid to supporting the population of the templates at the point of data entry and this is where Populous’s main contribution lies.

1.1 Previous work

Various tools are available to support the conversion of spreadsheet data into statements in a knowledge representation language. Excel2RDF³, Convert2RDF⁴ and RDF123 [1] are three tools that allow users to generate Resource Description Framework (RDF) statements from spreadsheets. Despite RDF being the reference syntax for the Web Ontology Language (OWL), its serialisation is complex and not intended for humans, making it inappropriate for defining higher level OWL construct in patterns.

The ExcelImporter plugin⁵ for Protégé 4.0 was a step up from these tools and enabled users to transform spreadsheets content directly into OWL axioms. It was, however, limited to only a small set of OWL constructs. The more recent tools to support template data and pattern instantiation include Mapping Master [14], OPPL 2 [11, 6] and the Protégé Matrix plugin⁶:

³ <http://www.mindswap.org/~rreck/excel2rdf.shtml>

⁴ <http://www.mindwap.org/~mhgoeve/convert/>

⁵ http://protegewiki.stanford.edu/wiki/Excel_Import

⁶ <http://protegewiki.stanford.edu/wiki/Matrix>

- The MappingMaster plugin for the Protégé 3.4 ontology editor is a more flexible tool for transforming arbitrary spreadsheet data into OWL. MappingMaster moves away from the row centric view of spreadsheets and has an expressive macro language called M² [15, 14] that can handle non-uniform and complex spreadsheets. M² combines a macro language for referring to cells in a spreadsheet with a human readable syntax for generating OWL expressions called the Manchester OWL Syntax [10]. MappingMaster and M² are primarily designed for the transformation of spreadsheet data to OWL, but provides little in the way of support for populating and validating the spreadsheet data.
- The Ontology Pre-Processing Language (OPPL) [6, 11] (version 2) is a scripting language similar to M². OPPL 2 is also Manchester OWL Syntax based and allows for the manipulation of OWL ontologies at the axiom level. OPPL 2 has support for the use of variables and the addition and removal of logical axioms from an ontology. OPPL 2 is a powerful scripting language for OWL and a user interface is provided via the OPPL plugin for Protégé 4.0. OPPL, however, does not currently support working with tabular data and is decoupled from any knowledge gathering.
- The MatrixPlugin for Protégé 4.0 allows users to specify simple OWL patterns in a tabular interface that can be used to populate repeating patterns with existing concepts from an ontology. This plugin is useful for ontology developers that have repetitive patterns to instantiate, and has the added benefit of cell validation and auto-completion at the point of data entry. The Matrix plugin is limited by the type of patterns that can be expressed along with the fact that it is tightly integrated with the Protégé interface, therefore, not suitable for all users. It does, however, combine knowledge gathering and axiom generation.

1.2 Requirements

All of the previous tools developed in this area tend to focus on the transformation from the template to the ontology. They provide little or no support for populating and validating template content. Furthermore, tools like ExcelImporter, OPPL and MappingMaster are integrated into the ontology development tool, so they are aimed at users that are already familiar with ontology development. A table based tool for ontology authoring should shield the user from the underlying ontology and help guide the user when populating the template. Providing validation at the time of authorship should significantly reduce the amount of time required to debug and process the data captured in the spreadsheet. Here we list some key requirements for an ontology based table editor:

1. Concepts may be described in terms of other concepts from other ontologies. In setting up Populous the users must be able to load and browse ontologies.
2. The contents in the column of a table need to be restricted to concepts from imported ontologies, or parts of imported ontologies.

3. To improve human comprehension the concept should be rendered using only the URI fragment, or optionally a human readable label from the ontology.
4. Each time a concept is added to a cell within the table Populous needs to check that the concept is valid according to the constraints resulting from requirement 2.
5. A cell might have multiple values; for example, when the concept being described has multiple parts.
6. Users should be free to suggest new concepts when an appropriate concept is not available.

2 Populous

Populous is an extension of RightField[21]; RightField is for creating Excel documents that contain ontology based restrictions on spreadsheets content. RightField enables users to upload Excel spreadsheets, along with ontologies from their local file systems, or from the BioPortal [13] (a repository of biological ontologies available at <http://bioportal.bioontology.org>). RightField supports OWL, OBO and RDFS ontologies. Using RightField, individual cells, or whole columns or rows can be marked with the required ranges of ontology terms. For example, they could include all subclasses from a chosen class, direct subclasses only, all individuals, or only direct individuals. Each spreadsheet can be annotated with terms from multiple ontologies. RightField is primarily designed for generating spreadsheet templates for data annotation; Populous extends RightField to support knowledge gathering and ontology generation. Populous and RightField are both open source cross platform Java application. They use the Apache-POI ⁷ for interacting with Microsoft documents and manipulating Excel spreadsheets. Populous is available for download from here <http://www.e-lico.eu/populous>.

Requirement 1 is already addressed using RightField functionality to upload both OWL and OBO ontologies. In order to better serve the life science community, users can also browse and load ontologies directly from BioPortal. Once the ontologies are loaded they are classified by a reasoner and the basic class hierarchy can be viewed.

Requirement 2 is met by the ability to select terms from the ontology to create validation sets. A data validation restricts the set of values that are valid for a particular cell in the table. Validations can span multiple rows and columns and be composed of classes, properties or individuals from the ontology. These data validations are stored in hidden worksheets along with additional information such as the full URI for the term, a label and the source ontology URI. These templates can also be exported as Microsoft Excel documents, which include the data validations on cells.

We address requirement 3 by allowing users to populate cells using ontology labels. Once data has been entered the default will be to render the ontology term using its label; if no label is specified the URI fragment is used. RightField

⁷ <http://poi.apache.org>

already supports reading Microsoft Excel workbooks so users are free to populate the templates in external tools before importing them into Populous for validation and transformation.

By using Populous directly users will benefit from having instant validation of the input data, satisfying requirement 4, along with some advanced features such as regular expression based auto-completion as they type into annotated cells. Additionally Populous supports the addition of multiple values into a single cell that are validated individually according to requirement 5. This can be particularly useful for certain kinds of patterns where a conjunction of variables is required to construct the axiom (see Section 3 for example). Populous also allows the addition of free text values, even if the cell has an associated validation range, thus satisfying requirement 6. These values are highlighted to the user in red and can act as placeholders for new or suggested terms when no suitable candidate could be found in the validation set.

Populous supports the use of OPPL patterns in order to generate new OWL axioms from the populated template. OPPL scripts can be written directly in Populous's design mode or imported from scripts generated in the OPPL plugin. Variables from the OPPL pattern must be mapped to columns from the table using the column name. A pattern Wizard guides the user through the generation and execution of the OPPL scripts. When the template is processed new identifiers for unknown terms can be auto-generated and exported from Populous.

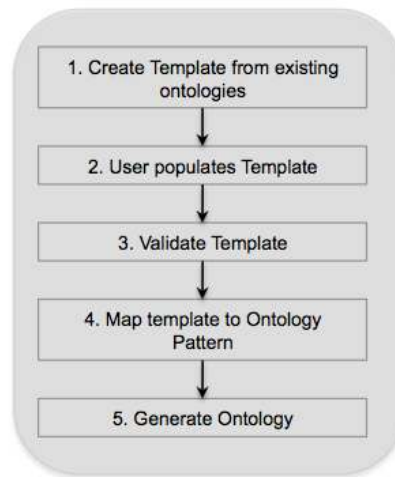


Fig. 1. Populous workflow

2.1 Building an Ontology with Populous

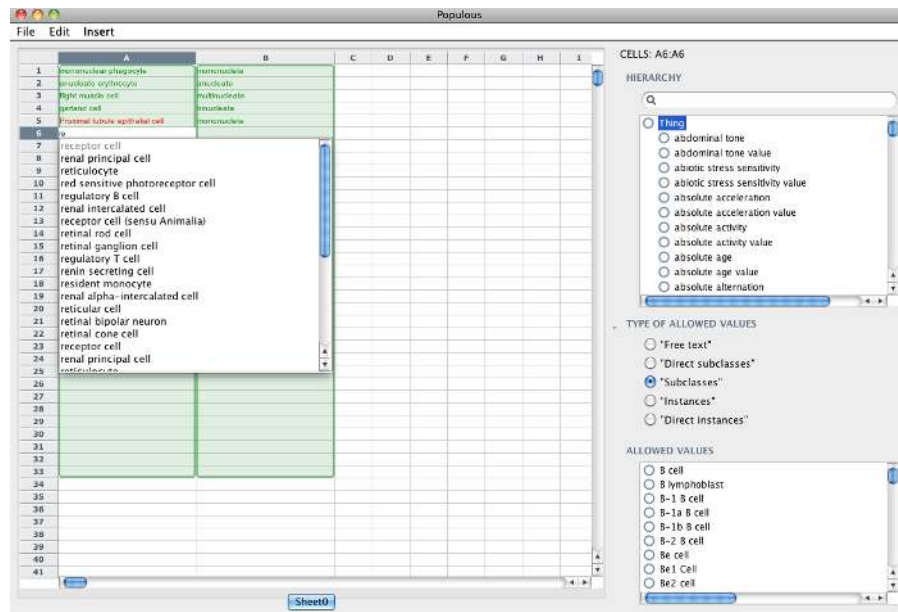


Fig. 2. Screenshot of Populous showing template population for cell types and nucleation

We can demonstrate Populous in building a simple ontology about cell types as described in Section 1. The pattern in the ontology states that every cell must have a nucleation. We need to create a template with two columns, column A is for cell type concepts, whilst column B is for nucleation concepts. Ontologies describing cells and their nucleation already exist that we can import into Populous. By connecting to BioPortal we can load the Cell Type Ontology (CTO) [4] and Phenotype and Trait Ontology (PATO) [8]. In order to restrict column A to terms from the CTO, we highlight all the cells in column A and restrict them to all subclasses of the root class. Column B is restricted subclasses of the *nucleation* concept from PATO. The template is now ready to be populated by the domain expert.

Figure 2 shows a part populated template. The terms in green indicate a valid term has been entered into the cell. The term in Column A5, *Proximal tubule epithelial cell* is red because it is not a valid term from the CTO. Cell A6 is in the process of being edited with the auto-completer offering valid suggestion for input.

The populated spreadsheet can now be transformed into an ontology. This can be done using the pattern wizard in Populous (Figure 3). The first step in

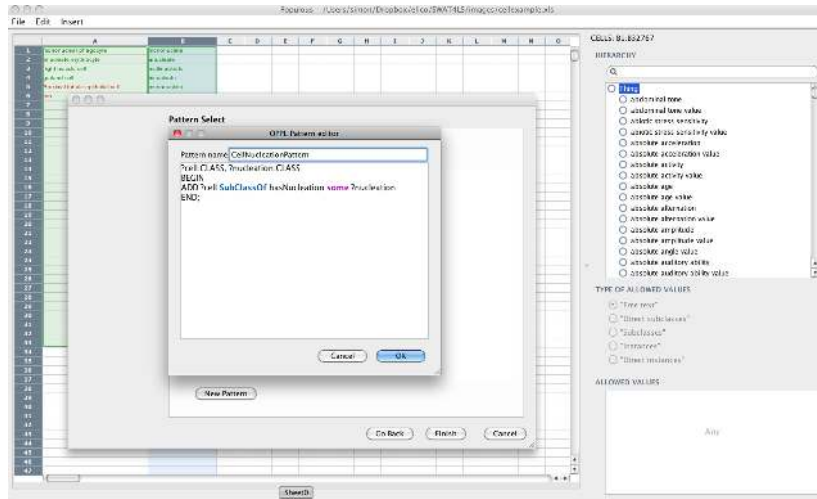


Fig. 3. Screenshot of Populous Pattern Wizard showing the OPPL script editor

the pattern wizard asks the user to select the columns and rows that contain populated data. In this example the pattern creates a restriction on each cell stating that all cells have a relationship, called *hasNucleation*, to an instance of the class nucleation. This pattern can be expressed in OPPL 2 with the syntax shown in Figure 4.

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

Fig. 4. OPPL 2 pattern for cells and nucleation

There are two variables in the pattern, `?cell` and `?nucleation`. These variables are mapped to column A and B respectively. The pattern is to be instantiated using data from rows one to six that must be specified in the Wizard. The next step involves validating the pattern, given that *Proximal tubule epithelial cell* is unknown by the validator, the user is given the option to assign a new URI for this concept. The final step generated the full OPPL script for applying this pattern. The Manchester OWL syntax generated for row one is shown in Figure 5. A complete grammar for the OPPL 2 syntax is available here⁸.

⁸ <http://oppl2.sourceforge.net>

```
Class: cto:CL.0000113
SubClassOf:
hasNucleation some pato:PATO.0001407
```

Fig. 5. Mononuclear Phagocyte described in Manchester OWL syntax generated from the OPPL 2 pattern in figure 4 (PATO.0001407 is the identifier for *mononucleate*)

3 Use case and evaluation

In order to evaluate Populous in a real ontology building scenario, it has been used to populate a template for gathering knowledge about the kidney and urinary system. The kidney is a complex organ composed of several distinct anatomical compartments that together enable the filtration of waste from the blood in the form of urine. Each of the kidney compartments is formed from a wide variety of cell types, and the specificity of the compartments relies on these specialised cell functions. The Kidney and Urinary Pathway Ontology (KUPO) [12] describes kidney cells, their function and their anatomical locations. KUPO is being built to annotate and integrate a variety of KUP related data held in the Kidney and Urinary Pathway Knowledge Base (KUPKB)⁹.

A simple template was designed for experts from the KUP domain to capture the relationships between cell types, their anatomical location and their biological functions. The template has three main columns; column A is for entering cell type terms, column C is for anatomy terms and column D for biological process terms. Populous was used to constrain the allowable values in columns A, C and D to concepts from the Open Biomedical Ontology Cell Type Ontology [4], subclasses or part of the *Kidney* or *Urinary system* concepts from the Mouse Adult Gross Anatomy Ontology [9], and all subclasses of the *Biological Process* concept from the Gene Ontology [20], respectively. The experts were instructed that the relationship between concepts in column A and C was *part of*, and the relationship between column A and D, *participates in*. For concepts that were related to multiple concepts they were allowed to list concepts in a cell separated by a comma. Figure 6 is a screen shot of Populous populated with data from the domain experts.

In order to transform the tabular data into an OWL representation the OPPL pattern in Figure 7 was created by the ontology engineers. This pattern states that a cell type is equivalent to a cell that is *part of* an anatomy term and a subclass of cells that participate in biological processes.¹⁰ For both restrictions the existential (*some*) quantification is used. The two differentia in this pattern for a cell genus are the anatomical location and the biological process, which is retrieved from column A, B and C respectively in the template. The entire KUP ontology is generated from the template data combined with the ontology

⁹ <http://www.e-lico.eu/kupkb>

¹⁰ Where possible we use the relationships from [19].

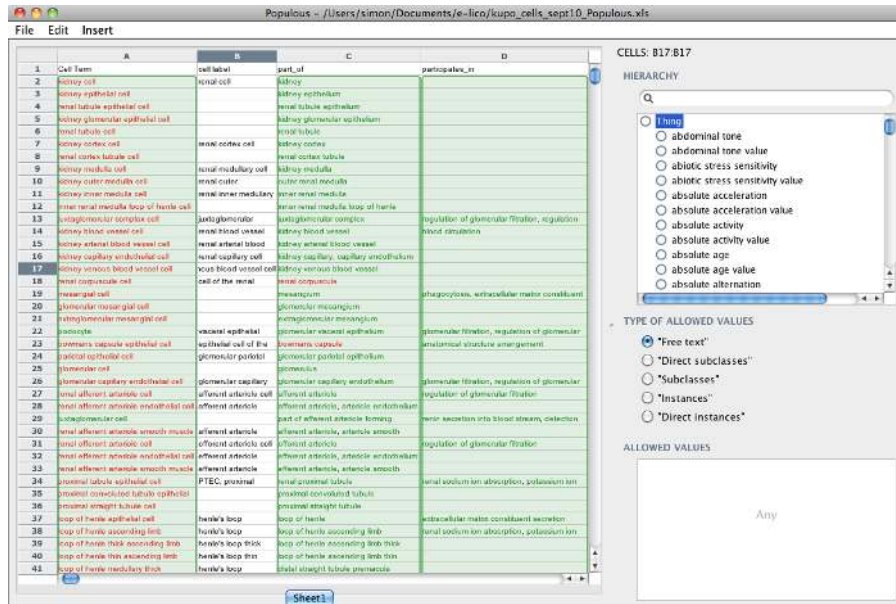


Fig. 6. Screenshot of Populous showing template population for KUP ontology

pattern. Figure 8 shows the Manchester OWL syntax generated from the pattern and data from row 13 for the *Juxtaglomerular complex cell*.

Using this template approach the domain experts described over 140 cell types, many of which are absent from the current CTO. Figure 9 shows the inferred hierarchy after classifying the ontology in Protégé 4.0. Note the asserted class hierarchy is simply a flat list of cell types, the parthood of the mouse anatomy is used to drive inferences about super/sub class relationships between cell types. Leaving the reasoner to compute the class hierarchy allows the domain experts to inspect the ontology for missing or incorrect inferences. These often indicate some missing information in the template, or an error in the imported ontologies. This methodology was successful in engaging the domain experts to both contribute to the KUP ontology and generate new term requests for the imported ontologies.

4 Discussion

Populous is designed for domain experts to gather knowledge that can be subsequently used to populate ontologies. Whilst previous tools have provided support for transforming templates into ontologies, they lacked basic support to help the user at the point of data entry. Populous was designed to fill this niche and meet the requirements outlined in section 1. The simple tabular interface used

```

Pattern 1
?cell:CLASS,
?anatomyPart:CLASS,
?partOfRestriction:CLASS = cell and part_of some ?anatomyPart,
?anatomyIntersection:CLASS = createIntersection(?partOfRestriction.VALUES)
BEGIN
ADD ?cell equivalentTo ?anatomyIntersection
END;

```

```

Pattern 2
?participant:CLASS,
?participatesRestriction:CLASS = ?cell and participates_in some
?participant,
?participatesIntersection:CLASS = createIntersection(?participatesRestriction.VALUES)
BEGIN
ADD ?cell SubClassOf ?participatesIntersection
END;

```

Fig. 7. Two OPPL 2 patterns for describing cell types in KUPO

```

Class: kupo_000027
EquivalentTo:
cell:CL_0000000
and (ro:part_of some MA:MA_0002580)
SubClassOf:
cell:CL_0000000,
ro:participates_in some gene_ontology:GO_0002000,
ro:participates_in some gene_ontology:GO_0002001

```

Fig. 8. Manchester OWL syntax for Juxtaglomerular cell (MA_0002580 = ‘part of afferent arteriole forming juxtaglomerular complex’, GO_0002000 = ‘detection of renal blood flow’ and GO_0002001 = ‘renin secretion into blood stream’)

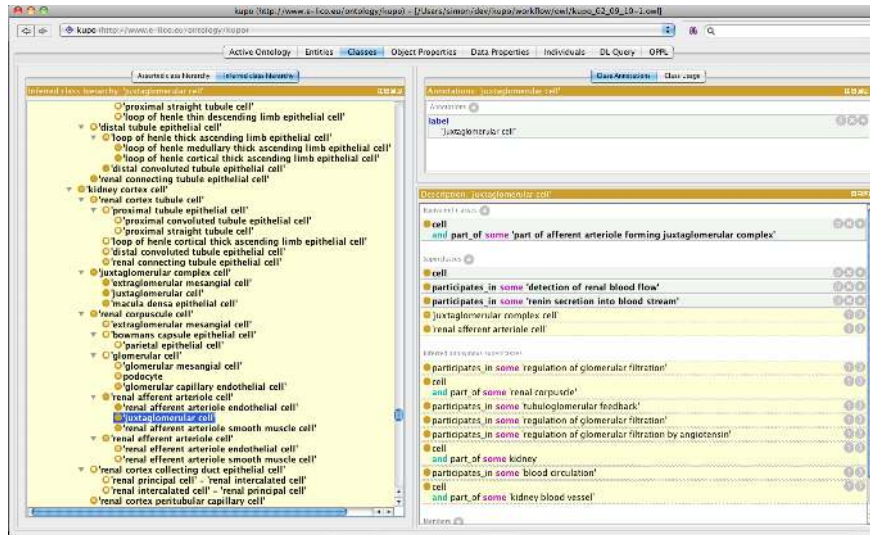


Fig. 9. Screenshot of KUPLO loaded into Protégé 4.0 showing inferred class hierarchy for Juxtaglomerular cell

in Populous is familiar to users who have already used a spreadsheet application. Populous should lower the entry requirements for domain experts to contribute to ontology development projects.

The release of Populous as presented is an early version; there remains many possible additions, some of which are:

1. Populous can handle multiple values in a cell that maps to conjunctions of properties. Extensions to Populous such that some minor syntax can be used to extend the ability to use OWL's syntax would be useful. In particular, being able to specify numbers for cardinality and numbers and other literals for datatype properties.
2. As already described, Populous uses a row centric model. We aim to use M^2 to enable more variety in how tables are mapped to templates or patterns. For example, only portions of columns may be required to be mapped to certain axiom patterns and M^2 enables this sort of mapping. OPPL and M^2 together should cover our mapping needs.
3. Populous currently gathers domain knowledge for the ontology, but not about the ontology. We aim to extend Populous to support various metadata such as editorial metadata and definitional metadata etc.
4. Populous is a single user application. Making Populous collaborative such that contributors may collectively add material to the same spreadsheet.
5. Feedback from the generated ontology to fix or extend data in Populous is currently *ad hoc*. A tighter coupling of this feedback cycle, without having

to go into an axiom based editor, will increase the quality assurance aspects of Populous.

We have demonstrated how Populous can be used to develop an ontology describing cells of the kidney and urinary pathway system. This demonstration highlights how domain experts managed to generate a real application ontology without being exposed to an ontology language like OWL, or a tool like Protégé. Populous's main purpose is for knowledge gathering and not ontologising. By shielding users from the ontology, except for review later in the process, they are left to concentrate on the biology and not worry about the axioms needed to represent it. This separation is particularly useful should the ontologist wish to change the conceptualisation or experiment with different patterns for the representation.

Our experience in developing the KUPO with domain experts provided interesting insights into the benefits of developing an ontology in this way. Classical approaches to ontology development in the life science have tended to focus on building rich asserted hierarchies of concepts. The KUPO approach exploits the expressiveness of an ontology language like OWL to describe the cells in such a way that the class hierarchy is computed by the reasoner. This means we have a logical explanation for all the subsumptions in the hierarchy, that is useful for spotting erroneous or missing information. For example, there are cell types for the vasa recta descending limb and the vasa recta ascending limb, both of which have different functions. The imported anatomy ontology, however, only describes a vasa recta. The domain experts were able to spot this and can now submit a request for these two new concepts to be added to the anatomy ontology. Building normalised ontologies that facilitate the kinds of inferences we see in KUPO are generally considered to be harder and more time consuming than constructing class hierarchies manually, despite offering a clear benefit [17]. However, in cases where a repeating pattern can be abstracted from the ontology, as in the case of KUP cells, we see that domain experts can rapidly produce rich ontologies with considerably less investment using Populous.

The question now arises as to how far can you go with a tool like Populous? Populous is by no means a replacement for full blown ontology editors, nor is it intended to be. The scenarios where Populous is of benefit assume that the ontology being developed has repeating patterns in the modelling. Furthermore, specifying the patterns for new ontologies before they exist is particularly difficult and is often something that emerges later as the ontology matures. For example, it was assumed with the kidney cells that we could describe them all in terms of their anatomy, only to later find some exceptions to the pattern. Renal principal and renal intercalated cells are currently indistinguishable by anatomy and function alone. There are always going to be exceptions, especially when modelling a complex domain like biology. We hope that Populous can bring more domain experts into the ontology development process and engage them in the development process.

The template approach can be particularly advantageous in scenarios where the modelling needs to change. [16] showed how templates can be used to gen-

erate different ontological representations of the same data. The KUPO is also being used to link data in an RDF store to support the KUP KB, where only limited support for OWL inferences is possible. By developing a different set of patterns we could generate a simpler version of the KUP ontology from the same Populous data for use in such an application. This again highlights an added benefit of separating the pattern from its population.

OPPL provides an expressive language for generating OWL patterns. OPPL's support for variables make mapping single columns from tabular data to variables convenient. The built in macros means we can create abstract pattern where the axioms can be generated dynamically depending on the number of values stored in a variable. Populous is currently limited to working with uniform spreadsheets and assumes a row-per-entity paradigm, where single columns map to a particular variable. This structure keeps the template simple and should cover the majority of use cases for populating an ontology in this way. The extension to support multi-values per cell offers some additional flexibility over existing spreadsheet based approaches. In order to accommodate more complex spreadsheets we plan to extend Populous to support more complex mappings from columns to spreadsheets. We are also exploring integrating the M² language from MappingMaster directly into Populous, we note that templates created in Populous can already be exported as Excel document and loaded into MappingMaster for transformation should the user desire.

Populous offers a means of creating ontology content without the use of a standard ontology development tool. Just as data-entry tools exist for populating databases, so we also need such tools for populating ontologies. It is possible to separate knowledge gathering from conceptualisation and axiomatisation and Populous is one means of achieving this goal. Such a separation offers flexibility and simple form fill-in style of knowledge gathering that should make generation of axiomatically rich ontologies increasingly straight-forward.

Acknowledgements: We kindly acknowledge Mikel Egaña Aranguren for advice, requirements and testing Populous. This work is funded by the e-LICO project—EU/FP7/ICT-2007.4.4 and by SysMO-DB - BBSRC grant BBG0102181.

References

1. L.] Han, T.W. Finin, C.S. Parr, J. Sachs, and A. Joshi. RDF123: From Spreadsheets to RDF. In *ISWC 2008 LNCS 5318*, pages 451–466. Springer, 2008.
2. Mikel Egaña Aranguren, Erick Antezana, Martin Kuiper, and Robert Stevens. Ontology Design Patterns for bio-ontologies: a case study on the Cell Cycle Ontology. *BMC bioinformatics*, 9(Suppl 5):S1, 2008.
3. Michael Ashburner, Catherine A. Ball, Judith A. Blake, David Botstein, Heather Butler, J. Michael Cherry, Allan P. Davis, Kara Dolinski, Selina S. Dwight, Janan T. Eppig, Midori A. Harris, David P. Hill, Laurie Issel-Tarver, Andrew Kasarskis, Suzanna Lewis, John C. Matese, Joel E. Richardson, Martin Ringwald, Gerald M. Rubin, and Gavin Sherlock. Gene ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25–29, May 2000.
4. J. Bard, S. Y. Rhee, and M. Ashburner. An ontology for cell types. *Genome Biology*, 6(2), 2005.

5. Olivier Bodenreider. Ontologies and data integration in biomedicine: Success stories and challenging issues. In *DILS '08: Proceedings of the 5th international workshop on Data Integration in the Life Sciences*, pages 1–4, Berlin, Heidelberg, 2008. Springer-Verlag.
6. Mikel Egaña, Alan Rector, Robert Stevens, and Erick Antezana. Applying Ontology Design Patterns in Bio-ontologies. In A. Gangemi and J. Euzenat, editors, *EKAW 2008, LNCS 5268*, pages 7–16. Springer-Verlag, 2008.
7. Aldo Gangemi. Ontology design patterns for semantic web content. In *International Semantic Web Conference*, pages 262–276, 2005.
8. Georgios V. Gkoutos, Eain C. Green, Ann-Marie M. Mallon, John M. Hancock, and Duncan Davidson. Using ontologies to describe mouse phenotypes. *Genome biology*, 6(1), 2005.
9. Terry F Hayamizu, Mary Mangan, John P Corradi, James A Kadin, and Martin Ringwald. The adult mouse anatomical dictionary: a tool for annotating and integrating data. *Genome biology*, 6(3), 2005.
10. M Horridge, N Drummond, J Goodwin, A Rector, R Stevens, and H Wang. The Manchester OWL syntax. In *OWLed*, 2006.
11. Luigi Iannone, Alan L. Rector, and Robert Stevens. Embedding knowledge patterns into owl. In *ESWC*, pages 218–232, 2009.
12. Simon Jupp, Julie Klein, Joost Schanstra, and Robert Stevens. Developing a Kidney and Urinary Pathway Knowledge Base. In *Bio-ontologies SIG*, 2010.
13. Natalya F. Noy, Nigam H. Shah, Patricia L. Whetzel, Benjamin Dai, Michael Dorf, Nicholas Griffith, Clement Jonquet, Daniel L. Rubin, Margaret-Anne A. Storey, Christopher G. Chute, and Mark A. Musen. Biportal: ontologies and integrated data resources at the click of a mouse. *Nucleic acids research*, 37(Web Server issue):W170–173, July 2009.
14. M. J. O'Connor, C. Halaschek-Wiener, and M. A. Musen. M2: a Language for Mapping Spreadsheets to OWL. In *OWL: Experiences and Directions (OWLED), Sixth International Workshop*, 2010.
15. M. J. O'Connor, C. Halaschek-Wiener, and M. A. Musen. Mapping Master: a Spreadsheet to OWL Mapping Language. In *International Semantic Web Conference (ISWC)*, 2010.
16. Bjoern Peters, Alan Ruttenberg, Jason Greenbaum, Melanie Courtot, Ryan Brinkman, Patricia Whetzel, Daniel Schober, Susanna Assunta Sansone, Richard Scheuerman, and Philippe. Rocca-Serra. Overcoming the ontology enrichment bottleneck with quick term templates. *International Conference on Biomedical Ontology.*, 2009.
17. Alan L. Rector. Modularisation of domain ontologies implemented in description logics and related formalisms including owl. In *Proceedings of the 2nd International Conference on Knowledge Capture*, October, 2003, Sanibel Island, USA. 2003.
18. Barry Smith, Michael Ashburner, Cornelius Rosse, Jonathan Bard, William Bug, Werner Ceusters, Louis J. Goldberg, Karen Eilbeck, Amelia Ireland, Christopher J. Mungall, Neocles Leontis, Philippe Rocca-Serra, Alan Ruttenberg, Susanna-Assunta Sansone, Richard H. Scheuermann, Nigam Shah, Patricia L. Whetzel, and Suzanna Lewis. The obo foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology*, 25(11):1251–1255, November 2007.
19. Barry Smith, Werner Ceusters, Bert Klagges, Jacob Köhler, Anand Kumar, Jane Lomax, Chris Mungall, Fabian Neuhaus, Alan L. Rector, and Cornelius Rosse. Relations in biomedical ontologies. *Genome Biol*, 6:R46, Apr 2005.

20. The Gene Ontology Consortium. Gene Ontology: Tool for the Unification of Biology. *Nature Genetics*, 25:25–29, 2000.
21. Katy Wolstencroft, Stuart Owen, Matthew Horridge, Olga Krebs, Wolfgang Mueller, and Carole Goble. RightField: Rich Annotation of Experimental Biology Through Stealth Using Spreadsheets. In *Proceedings of the 7th Microsoft eScience Workshop*, pages 117–119., Berkeley, California, 2010.