Ontology Engineering

Lecture 1: Introduction to Knowledge bases, ontologies, and the Semantic Web

Maria Keet

email: mkeet@cs.uct.ac.za

home: http://www.meteck.org

Department of Computer Science
University of Cape Town, South Africa

Semester 2, Block I, 2019

Outline

1. Introduction
2. Where is it used?
   - 'Ontology inside'
   - The Semantic Web
3. What is an Ontology?

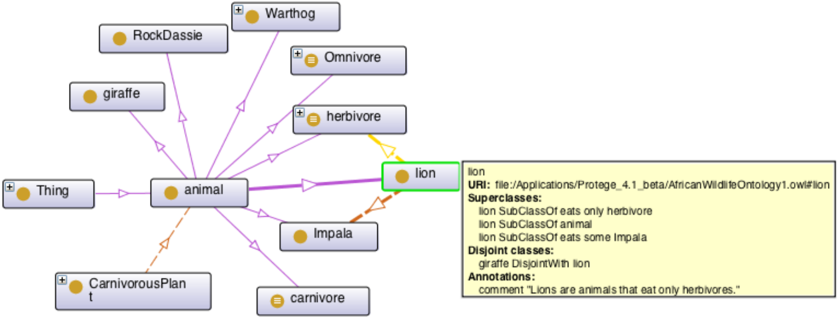
Outline

1. Introduction
2. Where is it used?
   - ‘Ontology inside’
   - The Semantic Web
3. What is an Ontology?

An ontology (very informally)

- classes, relationships between them, and constraints that hold between/for them, with possibly individuals and their relations
- as a representation of a particular subject domain

‘pretty’ picture of a section of the AWO



¡ there’s a lot going on behind the scenes !

Conceptual data models vs ontologies

- Main differences:
  - Information needs for one application vs. representing the knowledge of a subject domain (regardless the particular application)
  - Formalization in a logic language (though one could do that for conceptual models as well)

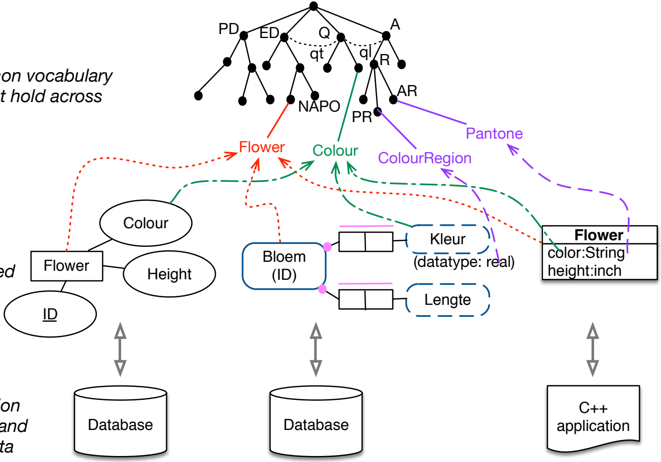
Conceptual data models vs ontologies

- Main differences:
  - Information needs for one application vs. representing the knowledge of a subject domain (regardless the particular application)
  - Formalization in a logic language (though one could do that for conceptual models as well)
- An ontology as a layer on top of conceptual data models
  - To improve the quality of a conceptual data model (hence, the software)
  - To facilitate system (database, application software) integration, or prevent the usual data integration problems

Ontology
provides the common vocabulary and constraints that hold across the applications

Conceptual model
shows what is stored in that particular application

Implementation
the actual information system that stores and manipulates the data



Databases vs. Knowledge bases

- Main differences:
  - Representation of the knowledge
  - Rules
  - Reasoning to infer new or implicit knowledge, detect inconsistencies of the knowledge base
  - Open World Assumption (vs. Closed World Assumption in databases)

What is the usefulness of an ontology?

- Making, more or less precisely, the (dis-)agreement among people explicit
- Enrich software applications with the additional semantics $\Rightarrow$ *ontology-driven information systems*
- Thus, practically, improving computer-computer, computer-human, and human-human communication

Outline

1. Introduction

2. Where is it used?
   - 'Ontology inside'
   - The Semantic Web

3. What is an Ontology?

Examples ontologies in information systems

- **e-learning** with *Inquire Biology* [Chaudhri et al., 2013]: textbook annotated with terms of the ontology, generates questions and answers.

Examples ontologies in information systems

- **e-learning** with *Inquire Biology* [Chaudhri et al., 2013]: textbook annotated with terms of the ontology, generates questions and answers.
- **data integration, cultural heritage**: combining resources of data and querying them, with a focus on the food system (in the Roman Empire) [Calvanese et al., 2016]

Examples ontologies in information systems

- **e-learning** with *Inquire Biology* [Chaudhri et al., 2013]: textbook annotated with terms of the ontology, generates questions and answers.
- **data integration, cultural heritage**: combining resources of data and querying them, with a focus on the food system (in the Roman Empire) [Calvanese et al., 2016]
- **publishing** of scientific papers, books: enable navigation and understanding of scholarly documents [Di Iorio et al., 2014]

Examples ontologies in information systems

- **e-learning** with *Inquire Biology* [Chaudhri et al., 2013]: textbook annotated with terms of the ontology, generates questions and answers.
- **data integration, cultural heritage**: combining resources of data and querying them, with a focus on the food system (in the Roman Empire) [Calvanese et al., 2016]
- **publishing** of scientific papers, books: enable navigation and understanding of scholarly documents [Di Iorio et al., 2014]
- **meta-mining of data mining experiments** (sections 1 and 5 of [Keet et al., 2015]): mine the (ontology-based) annotations of the data mining experiments, reason over that to have it propose the optimal data mining experiment

More Examples

- **For science** inside the scientific method: Outperforming humans (ontology+reasoner): classification of protein phosphatases [Wolstencroft et al., 2007]
- **Deep Question-Answering** with Watson beating human top-performers in 'Jeopardy!'; uses over 100 techniques, including ontologies for integration
- **Ontology-driven conceptual data modelling**: being more precise than just drawing diagrams, e.g., on those 'shared' and 'composite' aggregations in UML Class diagrams [Keet & Artale, 2008], finding contradictions.

Generalising from the examples:

- Data(base) integration
- Instance classification
- Matchmaking and services
- Querying, information retrieval
  - Ontology-Based Data Access
  - Ontologies to improve NLP
- Bringing more quality criteria into conceptual data modelling to develop a better model (hence, a better quality software system)
- Orchestrating the components in semantic scientific workflows, e-learning, etc.

The Semantic Web – Introduction (some motivations for ontologies and knowledge bases)

- AI put to the test in the (uncontrollable?) very large field
- Adding meaning to plain HTML pages and Web 2.0 by using theory and technologies of KBs and ontologies
  - *But there is more to ontologies and knowledge bases than their application in the Semantic Web!*
- See slides `semweb-intro.pdf` (bit outdated)
- Google's version of it: its "Knowledge graph" `https://www.youtube.com/watch?v=mmQl6VGvX-c`

Outline

1. Introduction

2. Where is it used?
   - 'Ontology inside'
   - The Semantic Web

3. What is an Ontology?

Background

– Aristotle and colleagues: **O**ntology
– Engineering: ontolog**ies** (count noun)

– Investigating reality, representing it
– Putting an engineering artefact to use

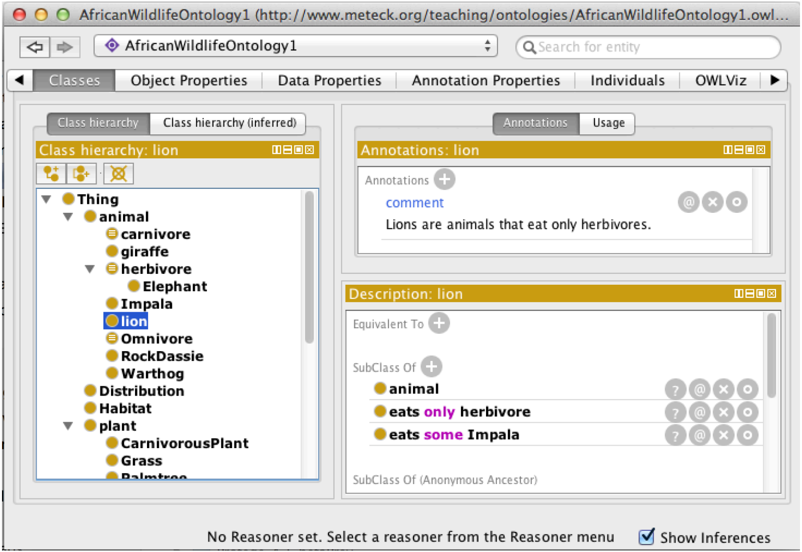
What then, is this engineering artefact?

First, let's look at an artefact: a text file....

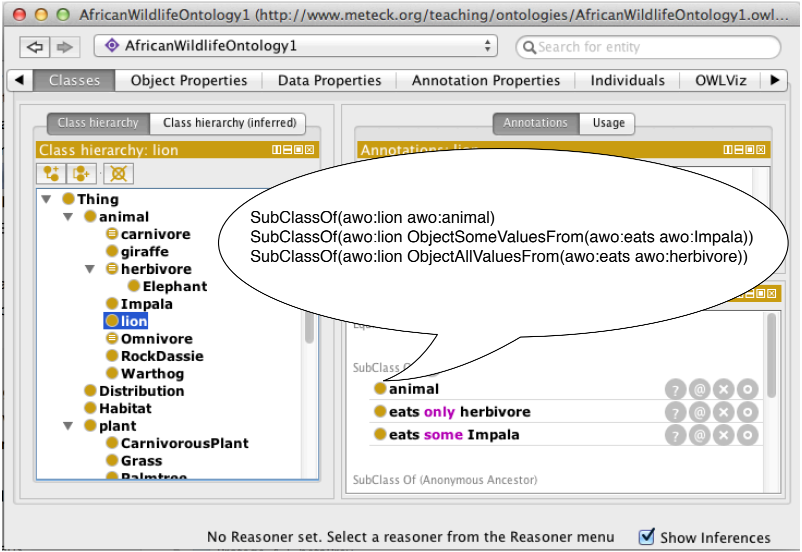
AfricanWildlifeOntology1.owl

```
<owl:Class rdf:about="&AfricanWildlifeOntology1;lion">
    <rdfs:subClassOf rdf:resource="&AfricanWildlifeOntology1;animal"/>
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:onProperty rdf:resource="&AfricanWildlifeOntology1;eats"/>
            <owl:allValuesFrom rdf:resource="&AfricanWildlifeOntology1;herbivore"/>
        </owl:Restriction>
    </rdfs:subClassOf>
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:onProperty rdf:resource="&AfricanWildlifeOntology1;eats"/>
            <owl:someValuesFrom rdf:resource="#Impala"/>
        </owl:Restriction>
    </rdfs:subClassOf>
    <rdfs:comment>Lions are animals that eat only herbivores.</rdfs:comment>
</owl:Class>



<!-- file:/Applications/Protege_4.1_beta/AfricanWildlifeOntology1.owl#plant -->

<owl:Class rdf:about="&AfricanWildlifeOntology1;plant">
    <rdfs:comment>Plants are disjoint from animals.</rdfs:comment>
</owl:Class>
```

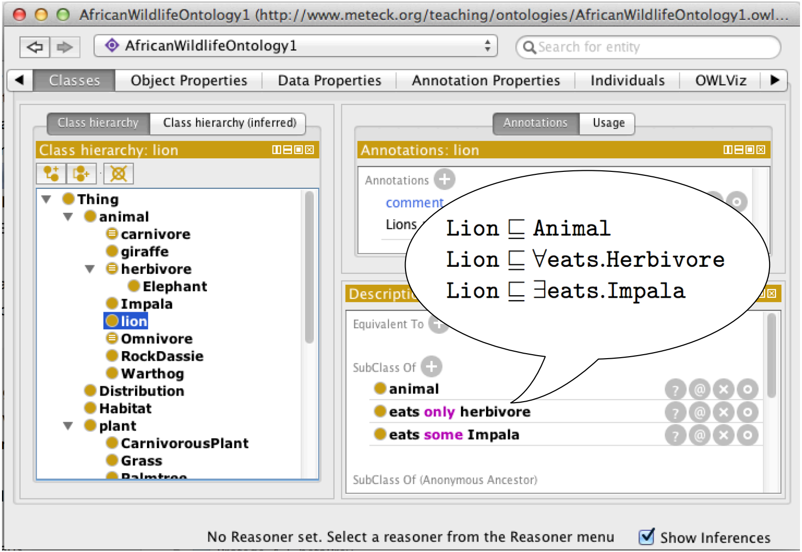
… or rendered in an ontology editor



Behind the facade



And behind that serialisation



A few definitions on what the text in the file is supposed to stand for

- Most cited (but very inadequate definition): "An ontology is a specification of a conceptualization" (by Tom Gruber, 1993)
- "a formal specification of a shared conceptualization" (by Borst, 1997)
- "An ontology is a formal, explicit specification of a shared conceptualization" (Studer et al., 1998)

A few definitions on what the text in the file is supposed to stand for

- Most cited (but very inadequate definition): "An ontology is a specification of a conceptualization" (by Tom Gruber, 1993)
- "a formal specification of a shared conceptualization" (by Borst, 1997)
- "An ontology is a formal, explicit specification of a shared conceptualization" (Studer et al., 1998)
- What is a *conceptualization*, and a *formal, explicit specification*? Why *shared*?

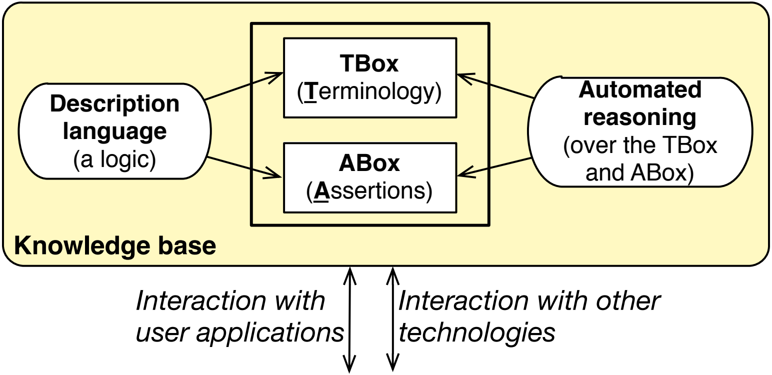
More definitions

- More detailed: “An ontology is a logical theory accounting for the *intended meaning* of a formal vocabulary, i.e. its *ontological commitment* to a particular *conceptualization* of the world. The intended models of a logical language using such a vocabulary are constrained by its ontological commitment. An ontology indirectly reflects this commitment (and the underlying conceptualization) by approximating these intended models.” (Guarino, 1998)

More definitions

- More detailed: "An ontology is a logical theory accounting for the *intended meaning* of a formal vocabulary, i.e. its *ontological commitment* to a particular *conceptualization* of the world. The intended models of a logical language using such a vocabulary are constrained by its ontological commitment. An ontology indirectly reflects this commitment (and the underlying conceptualization) by approximating these intended models." (Guarino, 1998)
- And back to a simpler definition: "with an ontology being equivalent to a Description Logic knowledge base" (Horrocks et al, 2003)

Description Logic knowledge base



From logical to ontological level (1/2)

- Logical level (no structure, no constrained meaning[1]):
  - $\exists x(Apple(x) \wedge Green(x))$
  - "there exists an object that is an apple and it is green"

[1] meaning in the sense of subject domain semantics, not formal semantics

[2] DL has a model-theoretic semantics, so the axioms have a meaning in that sense of 'meaning/semantics'

From logical to ontological level (1/2)

- Logical level (no structure, no constrained meaning[1]):
  - $\exists x(Apple(x) \land Green(x))$
  - "there exists an object that is an apple and it is green"
- Epistemological level (structure, no constrained meaning):
  - $\exists x : apple\ Green(x)$ (many-sorted logics)
  - "there exists an apple-object that is green"

[1] meaning in the sense of subject domain semantics, not formal semantics

[2] DL has a model-theoretic semantics, so the axioms have a meaning in that sense of 'meaning/semantics'

From logical to ontological level (1/2)

- Logical level (no structure, no constrained meaning[1]):
  - $\exists x(Apple(x) \wedge Green(x))$
  - "there exists an object that is an apple and it is green"
- Epistemological level (structure, no constrained meaning):
  - $\exists x : apple\ Green(x)$ (many-sorted logics)
  - "there exists an apple-object that is green"
  - ~~$\exists x : green\ Apple(x)$~~
  - "there exists a green-object that is an apple"

[1] meaning in the sense of subject domain semantics, not formal semantics

[2] DL has a model-theoretic semantics, so the axioms have a meaning in that sense of 'meaning/semantics'

From logical to ontological level (1/2)

- Logical level (no structure, no constrained meaning[1]):
  - $\exists x(Apple(x) \land Green(x))$
  - "there exists an object that is an apple and it is green"
- Epistemological level (structure, no constrained meaning):
  - $\exists x : apple\ Green(x)$ (many-sorted logics)
  - "there exists an apple-object that is green"
  - ~~$\exists x : green\ Apple(x)$~~
  - "there exists a green-object that is an apple"
  - $Apple(a)$ and $hasColor(a, green)$ (description logics[2])
  - "object a is an apple and that object a has the colour green"

---

1 meaning in the sense of subject domain semantics, not formal semantics

2 DL has a model-theoretic semantics, so the axioms have a meaning in that sense of 'meaning/semantics'

From logical to ontological level (1/2)

- Logical level (no structure, no constrained meaning[1]):
  - $\exists x(Apple(x) \land Green(x))$
  - "there exists an object that is an apple and it is green"
- Epistemological level (structure, no constrained meaning):
  - $\exists x : apple\ Green(x)$ (many-sorted logics)
  - "there exists an apple-object that is green"
  - ~~$\exists x : green\ Apple(x)$~~
  - "there exists a green-object that is an apple"
  - $Apple(a)$ and $hasColor(a, green)$ (description logics[2])
  - "object *a* is an apple and that object *a* has the colour green"
  - ~~$Green(a)$ and $hasShape(a, apple)$~~
  - "object *a* is a green and that object *a* has the shape of an apple"

[1] meaning in the sense of subject domain semantics, not formal semantics

[2] DL has a model-theoretic semantics, so the axioms have a meaning in that sense of 'meaning/semantics'

From logical to ontological level (2/2)

- Ontological level (structure, constrained meaning):
  - Some structuring choices are excluded because of ontological constraints
  - ‘apple objects’ seems better than ‘green objects’
  - objects having the colour green seems more sensible than having an ‘apple-shape’

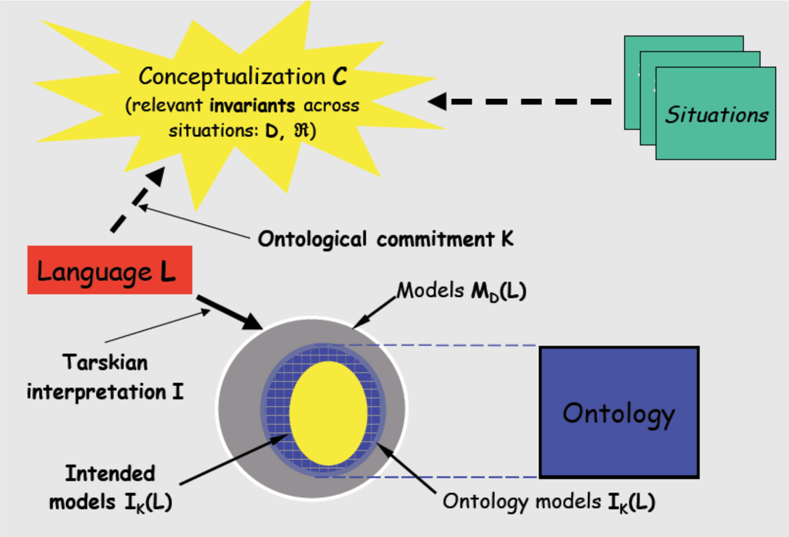
From logical to ontological level (2/2)

- Ontological level (structure, constrained meaning):
  - Some structuring choices are excluded because of ontological constraints
  - ‘apple objects’ seems better than ‘green objects’
  - objects having the colour green seems more sensible than having an ‘apple-shape’
  - There are reasons for that:
    - *Apple* carries an identity condition, so one can identify the object somehow (it is a ‘sortal’),
    - *Green* does not (is a value [‘qualia’] of the attribute [‘quality’] *hasColor* that a thing has)

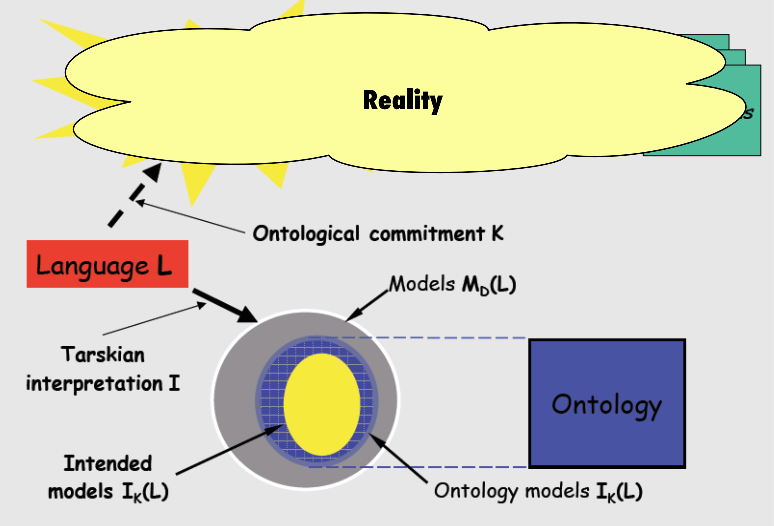
From logical to ontological level (2/2)

- Ontological level (structure, constrained meaning):
  - Some structuring choices are excluded because of ontological constraints
  - ‘apple objects’ seems better than ‘green objects’
  - objects having the colour green seems more sensible than having an ‘apple-shape’
  - There are reasons for that:
    - *Apple* carries an identity condition, so one can identify the object somehow (it is a ‘sortal’),
    - *Green* does not (is a value [‘qualia’] of the attribute [‘quality’] *hasColor* that a thing has)
- Put differently: one way of representing things turn out to be *better* than others.

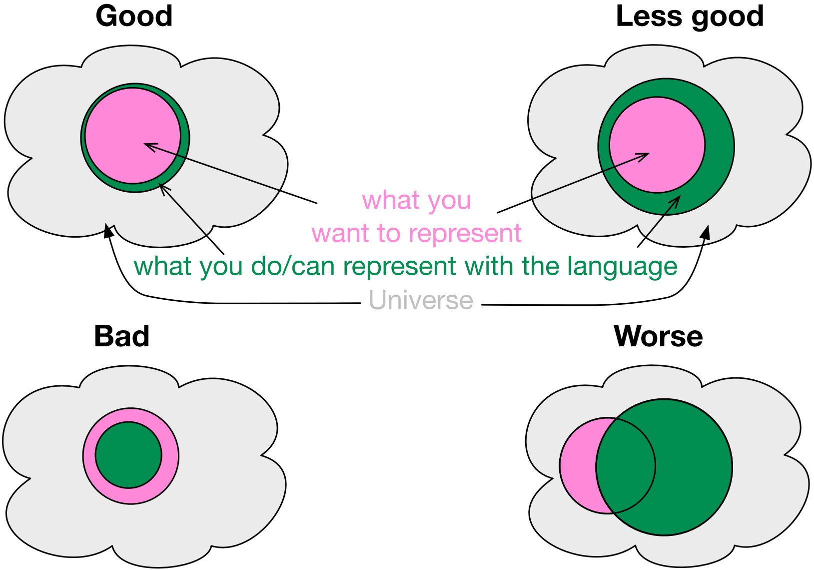
Ontologies and meaning



Ontologies and reality



Quality of the ontology



Initial Ontology Dimensions that have Evolved (Ontology Summit 2007)

- Semantic
  - Degree of Formality and Structure
  - Expressiveness of the Knowledge Representation Language
  - Representational Granularity

Initial Ontology Dimensions that have Evolved (Ontology Summit 2007)

- Semantic
  - Degree of Formality and Structure
  - Expressiveness of the Knowledge Representation Language
  - Representational Granularity
- Pragmatic
  - Intended Use
  - Role of Automated Reasoning
  - Descriptive vs. Prescriptive
  - Design Methodology
  - Governance

Summary

1. Introduction
2. Where is it used?
   - 'Ontology inside'
   - The Semantic Web
3. What is an Ontology?

Additional references

Calvanese, D., Liuzzo, P., Mosca, A., Remesal, J, Rezk, M., Rull, G. Ontology-Based Data Integration in EPNet: Production and Distribution of Food During the Roman Empire. *Engineering Applications of Artificial Intelligence*, 2016, 51:212-229.

Chaudhri, V.K., Cheng, B., Overholtzer, A, Roschelle, J., Spaulding, A., Clark, P., Greaves, M., Gunning, D.. Inquire Biology: A Textbook that Answers Questions. *AI Magazine*, 2013, 34(3): 55-72.

Di Iorio, A., Peroni, S., Vitali, F., Zingoni, J. (2014). Semantic lenses to bring digital and semantic publishing together. In Zhao, J., van Erp, M., Kessler, C., Kauppinen, T., van Ossenbruggen, J., van Hage, W. R. (Eds.), *Proceedings of the 4th Workshop on Linked Science (LISC 2014)*, CEUR Workshop Proceedings 1282: 12-23. Aachen, Germany: CEUR-WS.org.

Keet, C.M., Ławrynowicz, A., d'Amato, C., Kalousis, A., Nguyen, P., Palma, R., Stevens, R., Hilario, M. The Data Mining OPtimization ontology. *Web Semantics: Science, Services and Agents on the World Wide Web*, 2015, 32:43-53.

Keet, C.M., Artale, A. Representing and Reasoning over a Taxonomy of Part-Whole Relations. *Applied Ontology*, 2008, 3(1-2):91-110.

Note: where pictures and figures were taken from elsewhere, a note of the source is made in the LATEX source file as a comment. If there is no note about the source in that frame, then I made the figure.