



Exploiting Semantics in Information Integration: a National Mapping Agency Perspective

Ian Holt

(ian.holt@ordnancesurvey.co.uk)

Second Workshop on Challenges and Promise of the Semantic Web,
27th September 2007

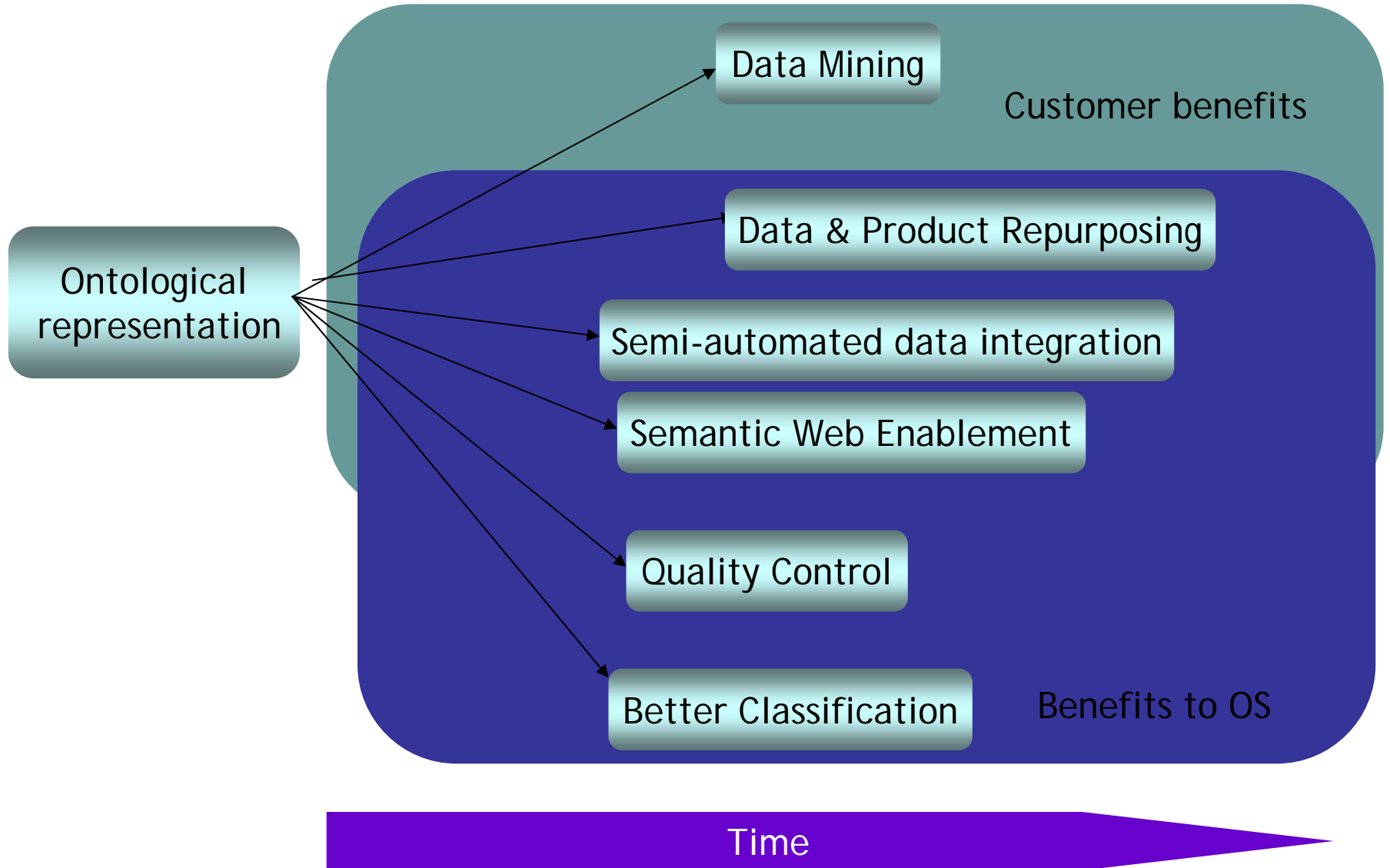
Cathy Dolbear, Paula Engelbrecht, John Goodwin,
Glen Hart

Ordnance Survey GeoSemantics research

- Motivation – why is a mapping agency interested in semantics?
- Our research
 - Ontology authoring
 - Linking ontologies to databases
 - Merging ontologies for data integration
 - Spatial and semantic reasoning & querying



Motivation



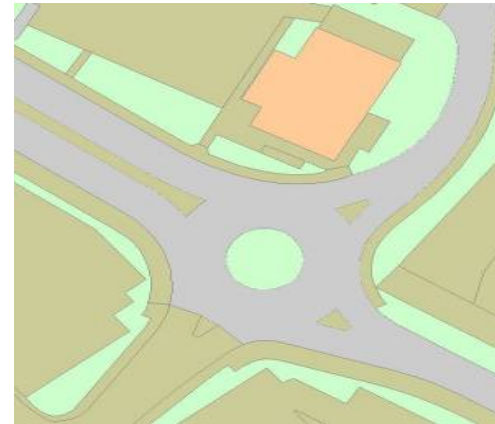


Ontology authoring



Topographic domain ontology

- Describes what we as an organisation know
- Provides a framework for specifying product content:
 - What our customers need
 - What can be captured and stored.
- Replaces legacy terminology e.g. Built Non Obstruction, Roofed Dolphin, Taxing Stone



Topographic domain ontology

- Developed Hydrology, Administrative Geography, Buildings and Places
- ~ 600 concepts, ALCOQ expressivity (OWL 1.1)
“An Allotment is owned by exactly 1 Local Authority”
- Working on Addresses, Settlements and Land forms
- Plus supporting modules: Mereology, Spatial Relations, Network topology.

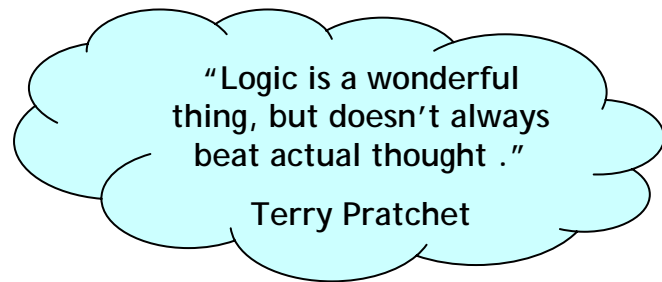
<http://www.ordnancesurvey.co.uk/ontology>

How to author ontologies - the literature

- Drawn on work from:
 - METHONTOLOGY [Fernandez Lopez et al, 1997]
 - Ontology Development 101 [Noy & McGuinness 2001]
 - Ontology authoring patterns [Semantic Web Best Practices Working Group, 2006]
- Assume a computer scientist writes the ontology

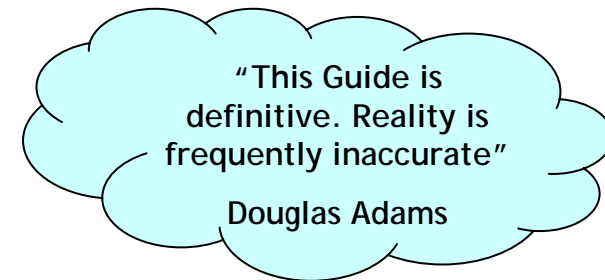
How to author ontologies - two views

Domain expert



Natural Language

Knowledge Engineer



Description
Logic
OWL

Our approach to building ontologies

Conceptual Aspect



Conceptual Ontology
Knowledge represented in a form understandable to people

Computational Aspect

Logical Ontology
Knowledge represented in a form manipulable by computers



A River Stretch
is part of a
River



River_Stretch \sqsubset
direct_part_of \exists River

```
<owl:ObjectProperty rdf:ID="directPartOf">  
  <rdf:type rdf:resource="#owl:FunctionalProp  
</owl:ObjectProperty>  
<owl:Class rdf:ID="River"/>  
<owl:Class rdf:ID="RiverStretch">  
  <rdfs:subClassOf>  
    <owl:Restriction>  
      <owl:onProperty rdf:resource="#directP  
      <owl:someValuesFrom rdf:resource="#R  
    </owl:Restriction>  
  </rdfs:subClassOf>
```

⏪ ⏩ <http://www.ordnancesurvey.co.uk/ontology/v1/BuildingsAndPlaces.owl>

Scope and Purpose Source Knowledge Capture Glossaries Formalisation Evaluation and Documentation

Class Annotations Matrix: Abbey

Class	label	NaturalLanguageTextDefinition	SentencesToUniquelyDefineConcept
TopographicObject			
FixedPermanentStructure			
Place			
Abbey	"Abbey"@en,"Abdi..."	"An abbey (from the Latin abbatia, which is deri..."	"An Abbey is a kind of Place.An Abbey has part an Abbey Church.An Abbey has purpos..."
Chapel			
Church			
Convent			
Gurdwara			
MeetingHouse			
Mission			
Monastery			
Mosque			

Object Property Features Matrix: appliesTo

Object Property	NaturalLanguageTextDefinition
appliesTo	"Use, utilise, employ in a particular situation"
hasName	"A name is a label for a thing, person, place, product (as in a brand name), and even an idea or concept, normally used to distinguish one from another. Names can i..."
hasPart	"something determined in relation to something that includes it; "he wanted to feel a part of something bigger than himself"; "I read a portion of the manuscript"; "the..."
hasPurpose	"an anticipated outcome that is intended or that guides your planned actions; "his intent was to provide a new translation"; "good intentions are not enough"; "it was ..."
isLedBy	
isLocatedIn	

Rabbit: Controlled natural language

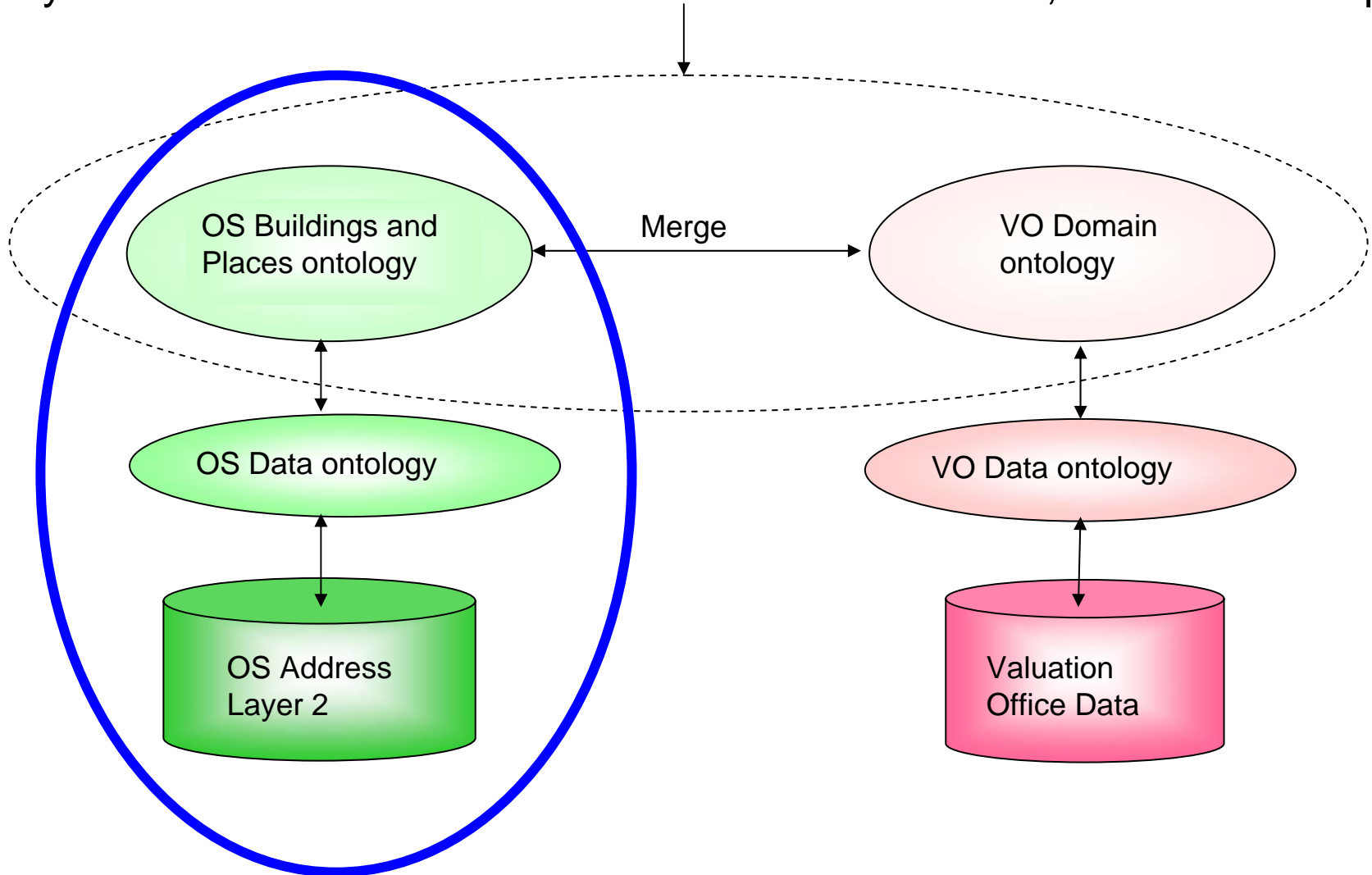
- Structured English, largely compilable to OWL
- Intelligible to domain expert
 - Can author conceptual ontology
 - Guides good modelling practice
 - Other domain experts are able to validate the work
 - Acts as documentation for the OWL
- Part of OWL 1.1 task force to develop a controlled natural language syntax



Linking ontologies to databases

Semantic data integration

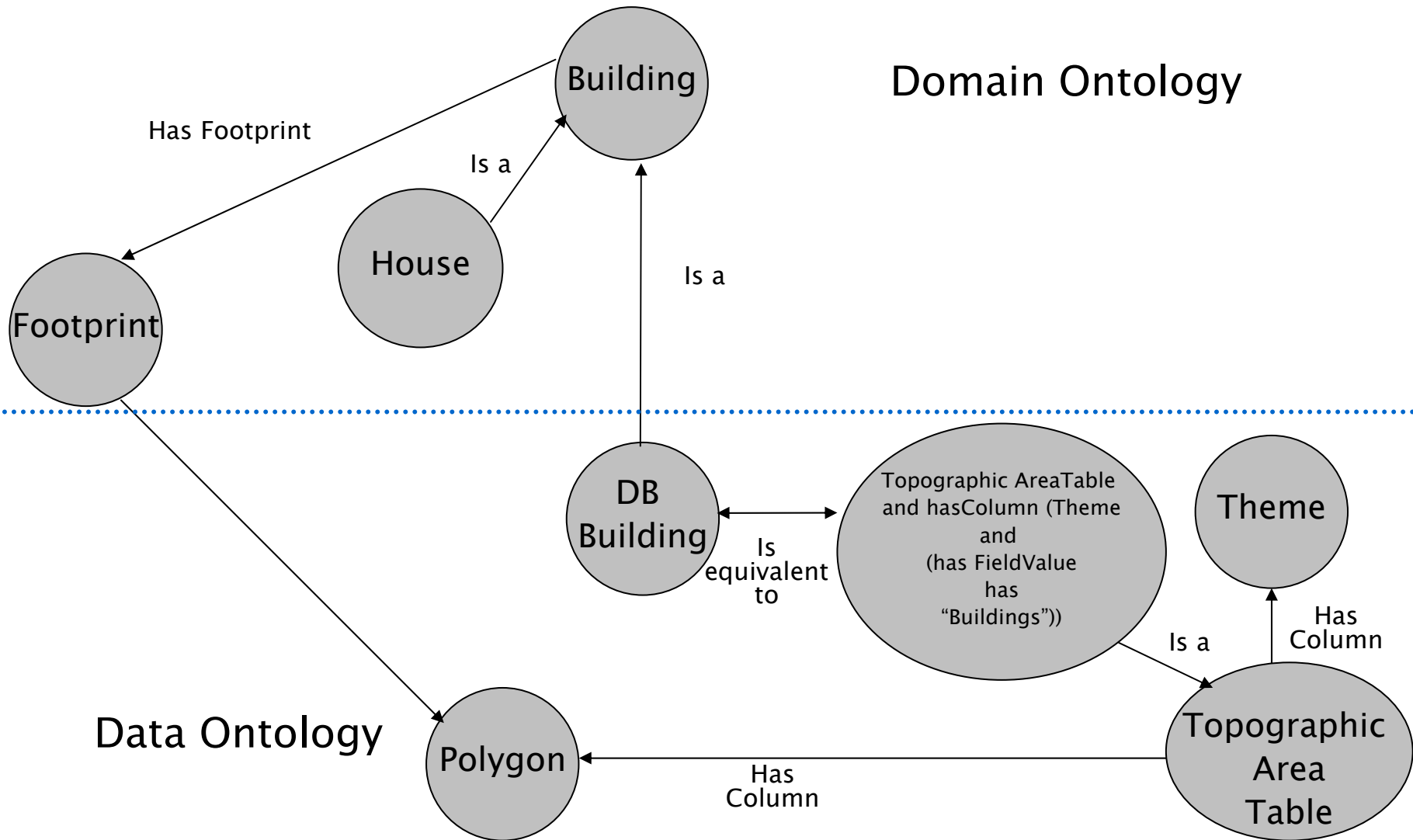
Query: Find all addresses with a taxable value over £500,000 in Southampton



Linking ontologies to databases (1)

- D2RQ - maps SPARQL queries to SQL, creating “virtual” RDF [Bizer et al, 2006]
- No need to convert data to RDF explicitly
- Modifying the API to:
 - Use OWL-based description of mapping (data ontology)
 - Map queries via spatial relations to SQL spatial operators
 - Use views to reduce number of triples and improve efficiency

Simplified Example: "Building"



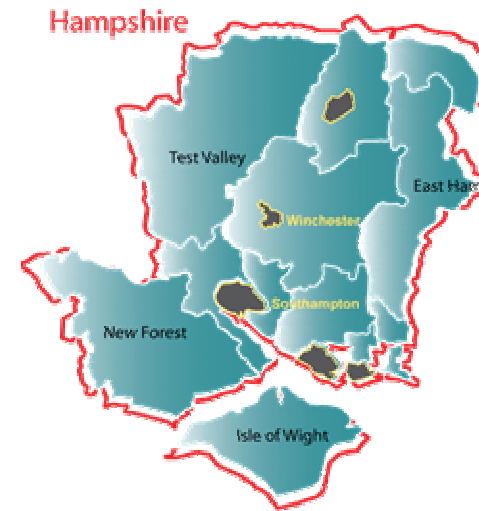
Linking ontologies to databases (2)

- Oracle RDF
 - Limited expressivity (RDFS in 10g, subset of OWL in 11g)
 - How much do we really need?
- Benefits likely to be using RDF as a data supply format rather than alternative to RDBMS
- RDF likely to degrade performance: we have > 10 billion triples
- Spatial component must be executed after RDF filtering
 - Would it be more efficient to perform the spatial query first to minimise the size of the RDF graph?

Semantically enabling data - an RDF Gazetteer

Experiment to semantically describe gazetteers

- OWL ontology to describe the concepts
- RDF version to represent the data
- Currently includes administrative regions
- Adding cities and other settlements, addresses etc
- Spatial boundary information embedded in the RDF using GML



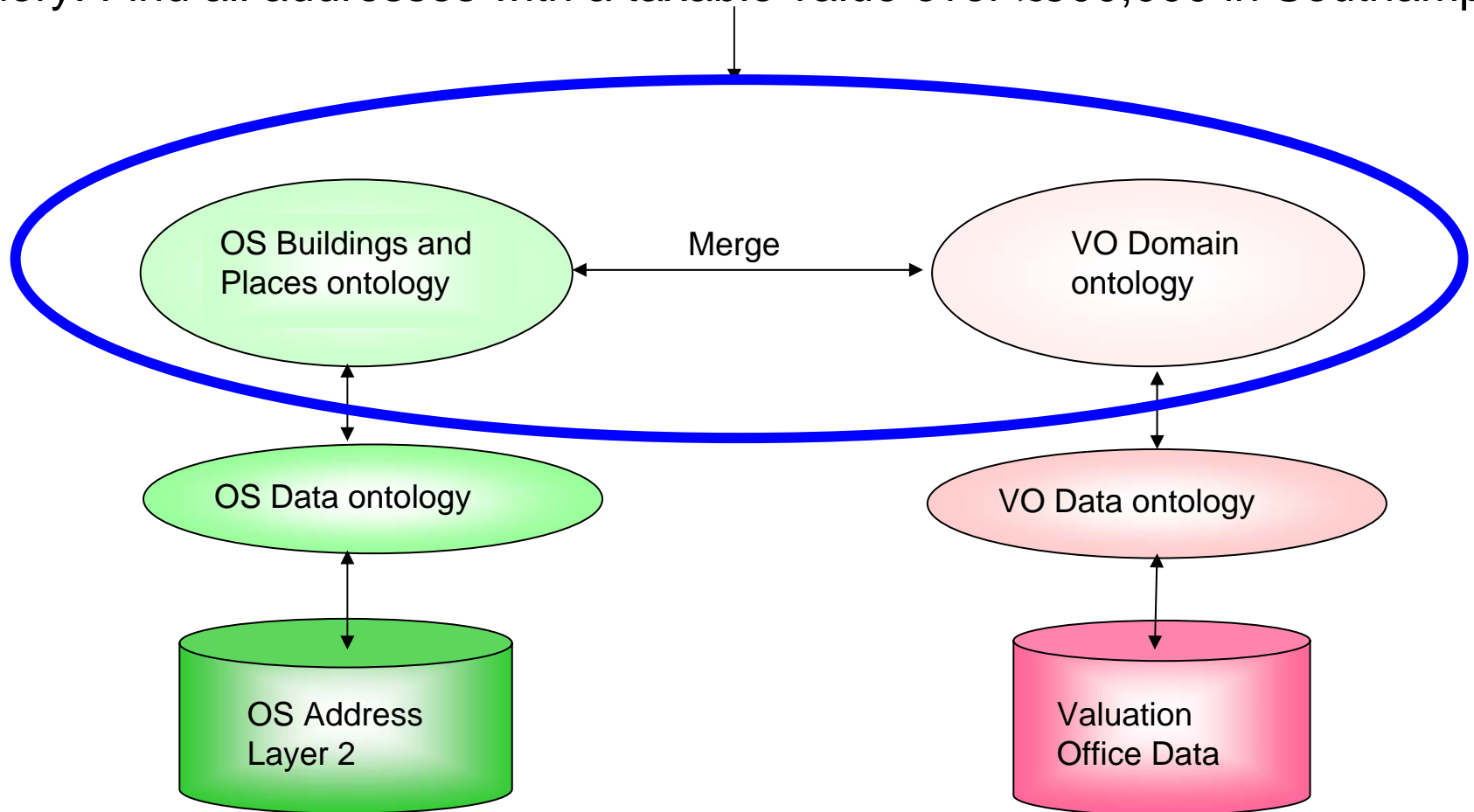


Ontology merging for data integration



Semantic data integration

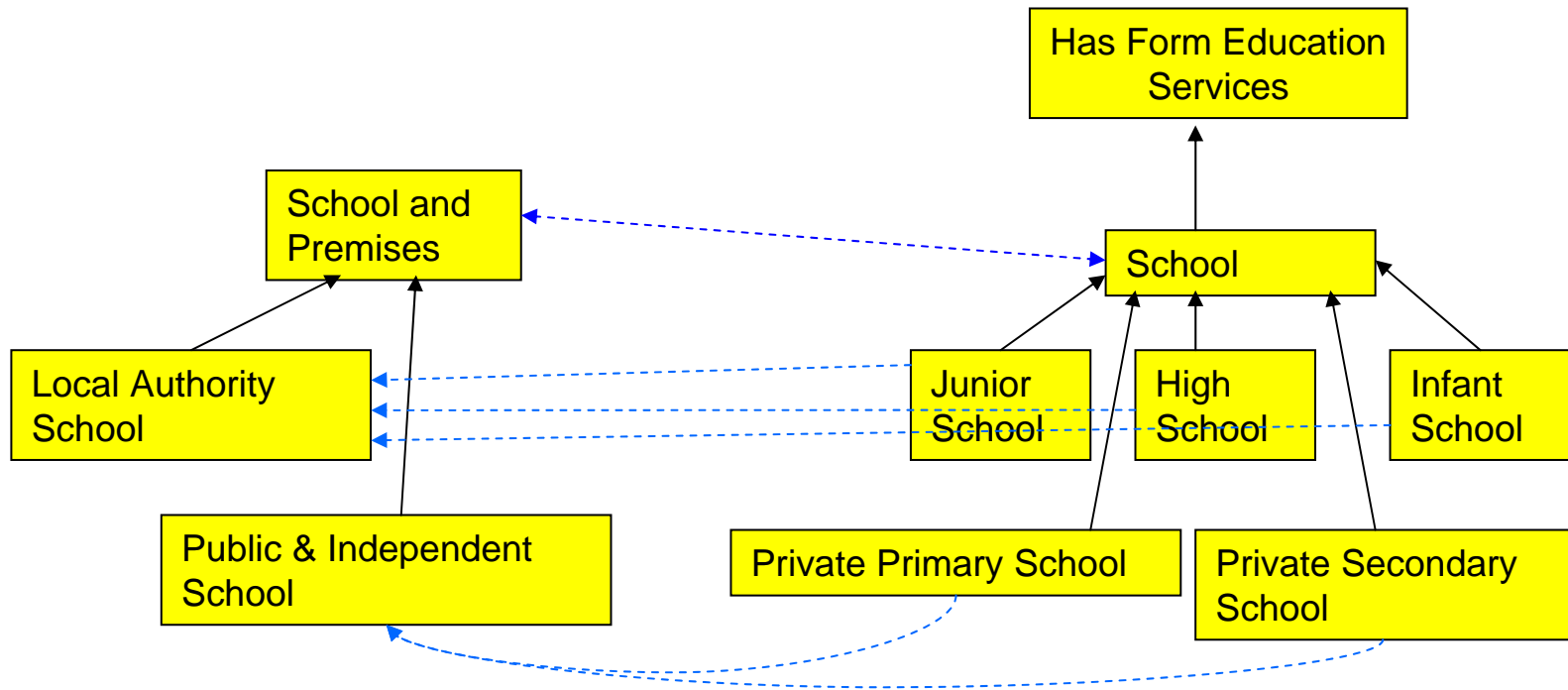
Query: Find all addresses with a taxable value over £500,000 in Southampton



Alignment: currently at data level

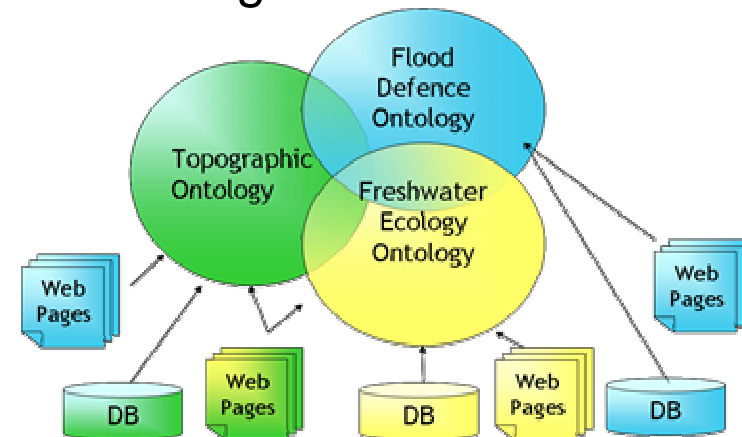
Valuation Office

Ordnance Survey



Merging ontologies

- Evaluation
 - Assess which concepts from one ontology should be included in another
- Modularisation
 - Isolate these concepts from the rest of the donor ontology
- Integration
 - Map and/or adapt the donor module for integration into the receiving ontology.



Merging ontologies – some thoughts

- Ontologies are never perfect: not a complete descriptions of any domain
- Automatic tools [PROMPT, FOAM, SWOOP modularisation etc] can help with navigation / suggestion *but*
 - Domain experts needed to add in their knowledge/ judgement
 - Some concepts are more important than others [noted by Sowa, 2005]
- Early stages of research

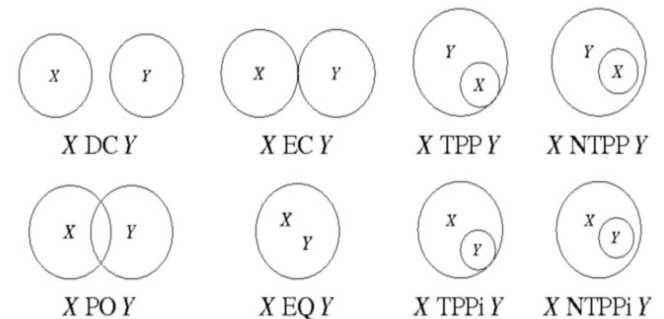


Spatial and semantic querying and reasoning

Spatial and semantic reasoning

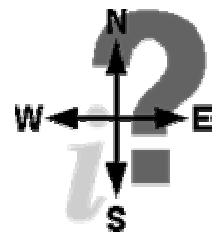
- Several options in the literature:
 - Convert OWL reasoner to RCC8 reasoner [Katz & Grau, 2005]
 - ε Connections to link the two logics
 - Concrete domains
- Trying something simpler – data ontology spatial relationships mapped to Oracle SDO_RELATE operators

spatially_contains maps to `SDO_RELATE(<geometry1>, <geometry2>, MASK='CONTAINS') = 'TRUE'`



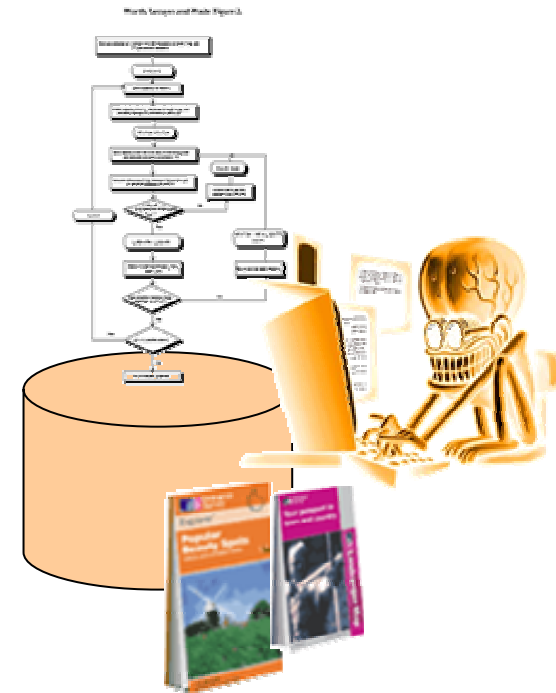
Spatial and semantic querying

- Graph matching using SPARQL
 - Can do SPARQL custom functions for spatial queries
 - Or split the query into spatial (SQL) and semantic (SPARQL) parts?
- Waiting for conjunctive query language (SPARQL DL)
- Should knowledge be included in the query or the ontology?
- How do we make it easier for domain experts to make queries?



Take home message

- “Semantics” research often doesn’t consider the semantics!
- Domain experts need to be at the centre of the process
- Technology transfer is difficult:
 - Benefits of semantics in products and applications must be clarified





Challenges

- Semantic Web Services
- Semantic Web Services



Thank you for your attention

<http://www.ordnancesurvey.co.uk/ontology>