Basics

Semantic Modelling

Sponsored by:

CAPSICUM

Business Architects

IT Architectures and Information modelling

- IT Architecture
  - Helps manage complexity in modern software systems
  - Supports agile processes
  - Supports reuse of assets
  - Reduces costs
  - Helps focus on core assets
- Decomposition
  - Functions decomposed into services
  - Business processes
  - Information provided at abstract level

Information modelling

- Purpose of information modelling is to:
  - Create a representation of real-world concepts and meaningful relationships between them
  - Provide a high level understanding of data by abstracting it further away from physical aspect of data storage
  - Represent a user's perspective of the data
- Choosing a modelling language is trade-off between
  - Formality/Informality: is the meaning of the modelling language same regardless of audience ?
  - Commonality and variability: how to manage things that are common and things that are variable ?
  - Expressivity: how detailed in the model ?

How expressive should a model be ?

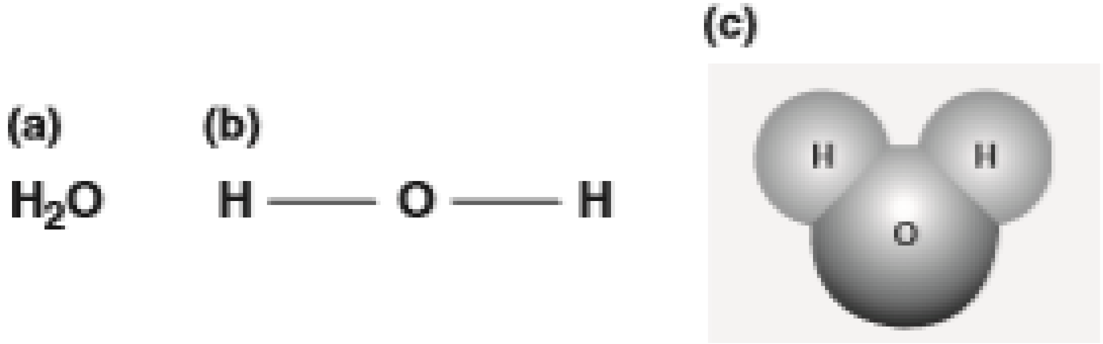
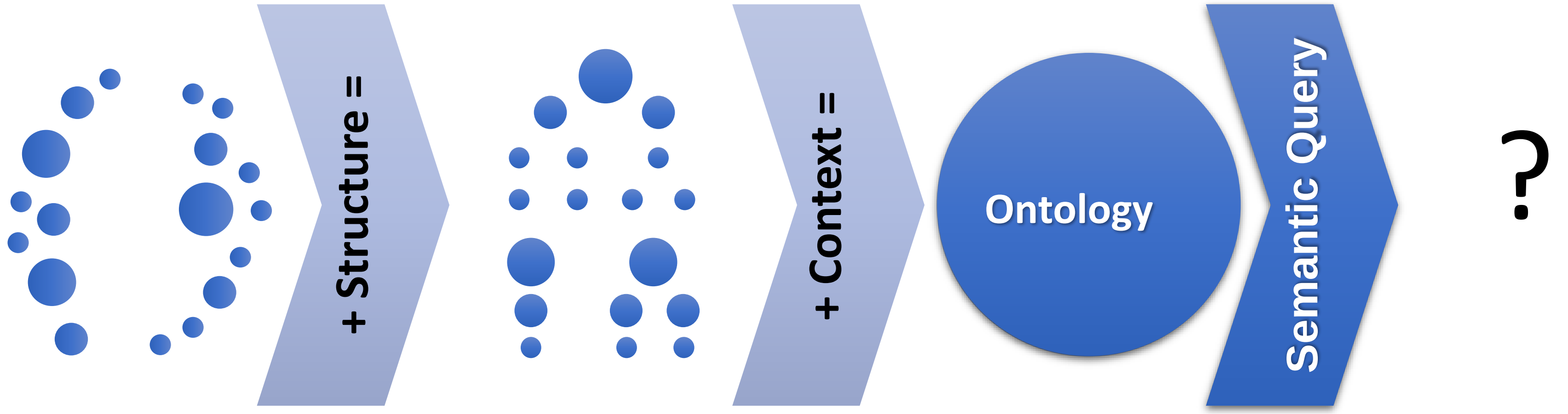


FIGURE 2.4

Different expressivity of models of a water molecule.

Moving up the Knowledge Continuum

Data	Information	Knowledge	Wisdom



Facts	Descriptions	Meaning	Understanding
What	*Who, Where, When*	*How*	*Why*
(signals, symbols, measurements)	(statements, structure, state, content)	(routines, processes, truth, beliefs)	(ideas, judgments, decisions, rationalization)

Data modelling languages

- Many different types of data models
  - Entity-relationship (ER) model: traditional modelling technique associated with software analysis and design
  - Relational model: suitable for relational databases and SQL queries. Can be seen as a restricted ER-model
  - Object model: suitable for OO analysis and design.
  - Semantic data model

Semantic data model

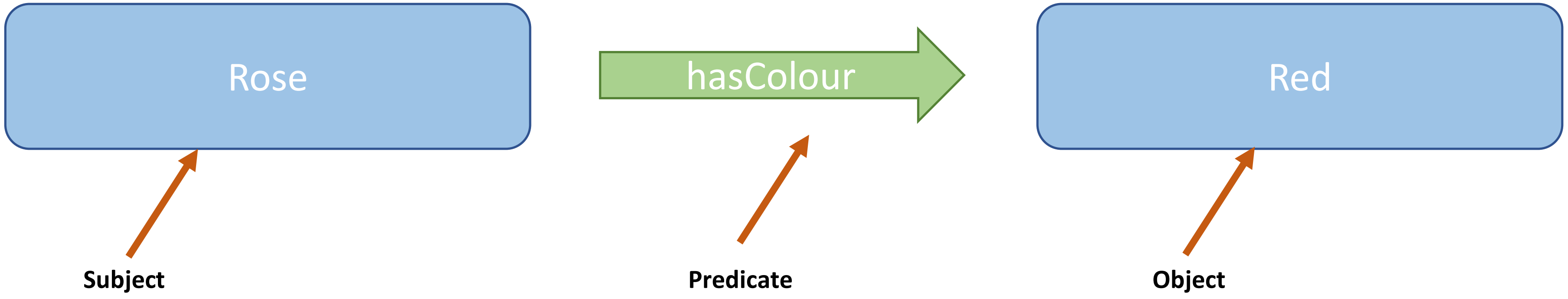
- Semantic data modelling is a method for representing data enriched with semantic information in the form of data values and relationships
- A semantic data model is more complex and expressive than the traditional data models
- Besides representing intensional structures like the traditional models, a SDM can also express implicit or derived knowledge from the explicit information, which is not possible with the traditional models

Why semantic models?

- **Reuse and interoperability**: Semantic models can be shared among applications and on the web
- **Flexibility**: Semantic models can operate in an open environment in which classes can be defined dynamically
- **Consistency and Quality Checking** across models
- **Reasoning**: Semantic models are supported by automated reasoning tools
- Semantic model creates ***ontologies***
- Semantic models are supported by a number of **technologies**

Subject–Predicate–Object expressions

- Semantics are expressed as triples of Subject-Predicate-Object
- RDF (Resource Description Framework) defines statements in this format



Comparison with relational models

- Traditional data modelling (Relational databases):
  - Concerned with structure of data
  - Inflexible (any changes would require changing entire tables and queries)
  - Well suited for large but simple data
  - No automated inferencing possible
  - Record-oriented modelling

- Semantic data modelling:
  - Concerned with the meaning of data (relationships)
  - Very flexible (new data can be added without affecting the existing data and queries)
  - Well suited for fewer but complex data
  - Automated inferencing possible using semantic reasoners
  - User-oriented modelling (user’s view of the real-world concepts)

Relational Database tables

Person table

ID	First Name	Last Name
P1	Sandra	Ferreira
P2	Steve	Barrett
P3	Mia	Shaw

Country table

ID	Country
C1	Australia
C2	France
C3	Greece

Capital city table

ID	Capital City
CC1	Athens
CC2	Canberra
CC3	Paris

isBornIn table

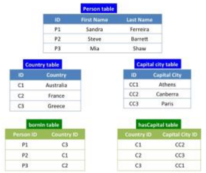
Person ID	Country ID
P1	C3
P2	C1
P3	C2

hasCapital table

Country ID	Capital City ID
C1	CC2
C2	CC3
C3	CC1

Example of Triples

Relationships



Subj. → *Obj.*

"Sandra Ferreira" —isBornIn→ Greece

"Steve Barrett" —isBornIn→ Australia

"Mia Shaw" —isBornIn→ France

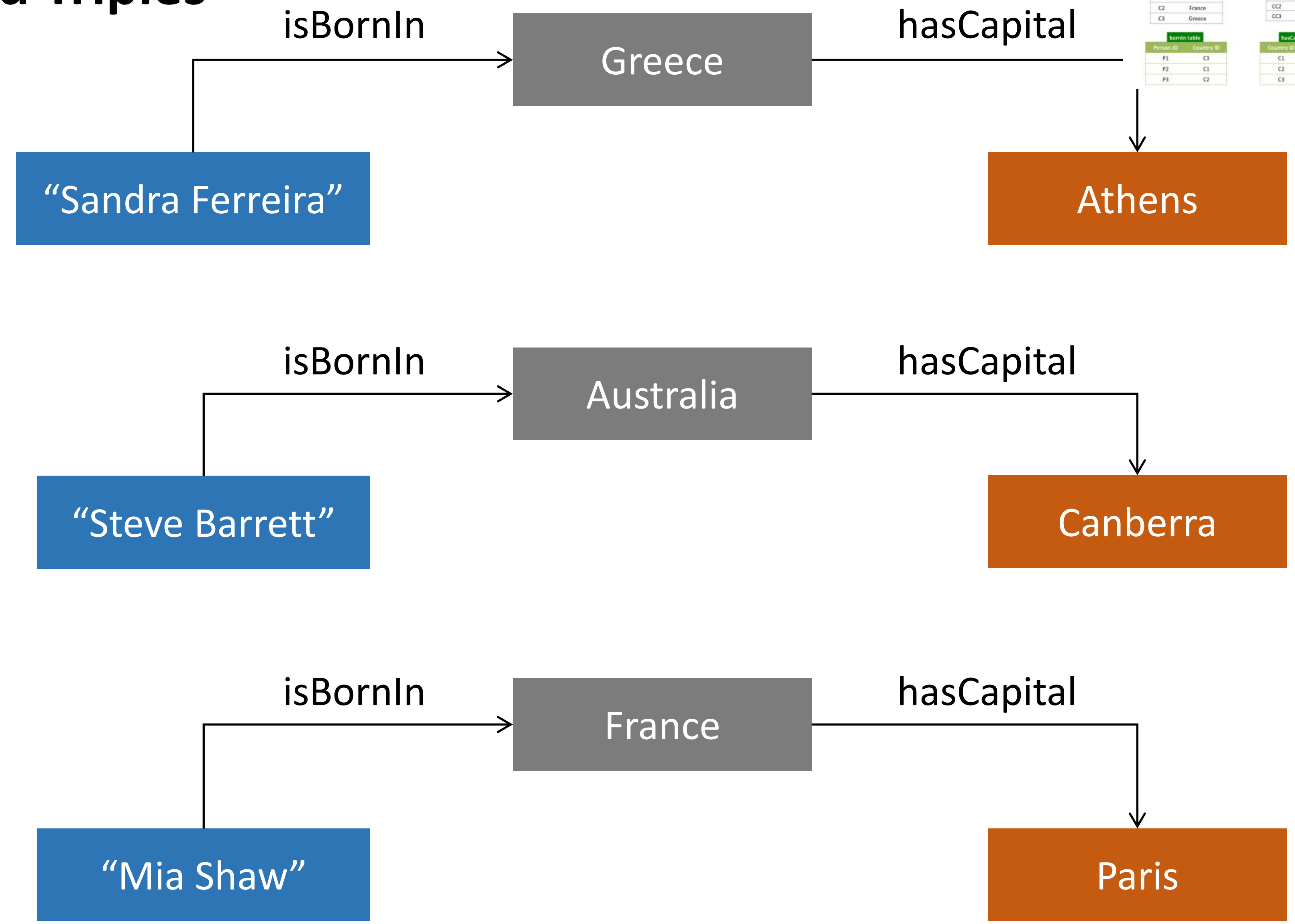
Subj. → *Obj.*

Australia —hasCapital→ Canberra

France —hasCapital→ Paris

Greece —hasCapital→ Athens

Linked Triples



Person table

ID	First Name	Last Name
P1	Sandra	Ferreira
P2	Steve	Barrett
P3	Mia	Shaw

Country table

ID	Country
C1	Australia
C2	France
C3	Greece

Capital city table

ID	Capital City
CC1	Athens
CC2	Canberra
CC3	Paris

bornIn table

Person ID	Country ID
P1	C3
P2	C1
P3	C2

hasCapital table

Country ID	Capital City ID
C1	CC2
C2	CC3
C3	CC1

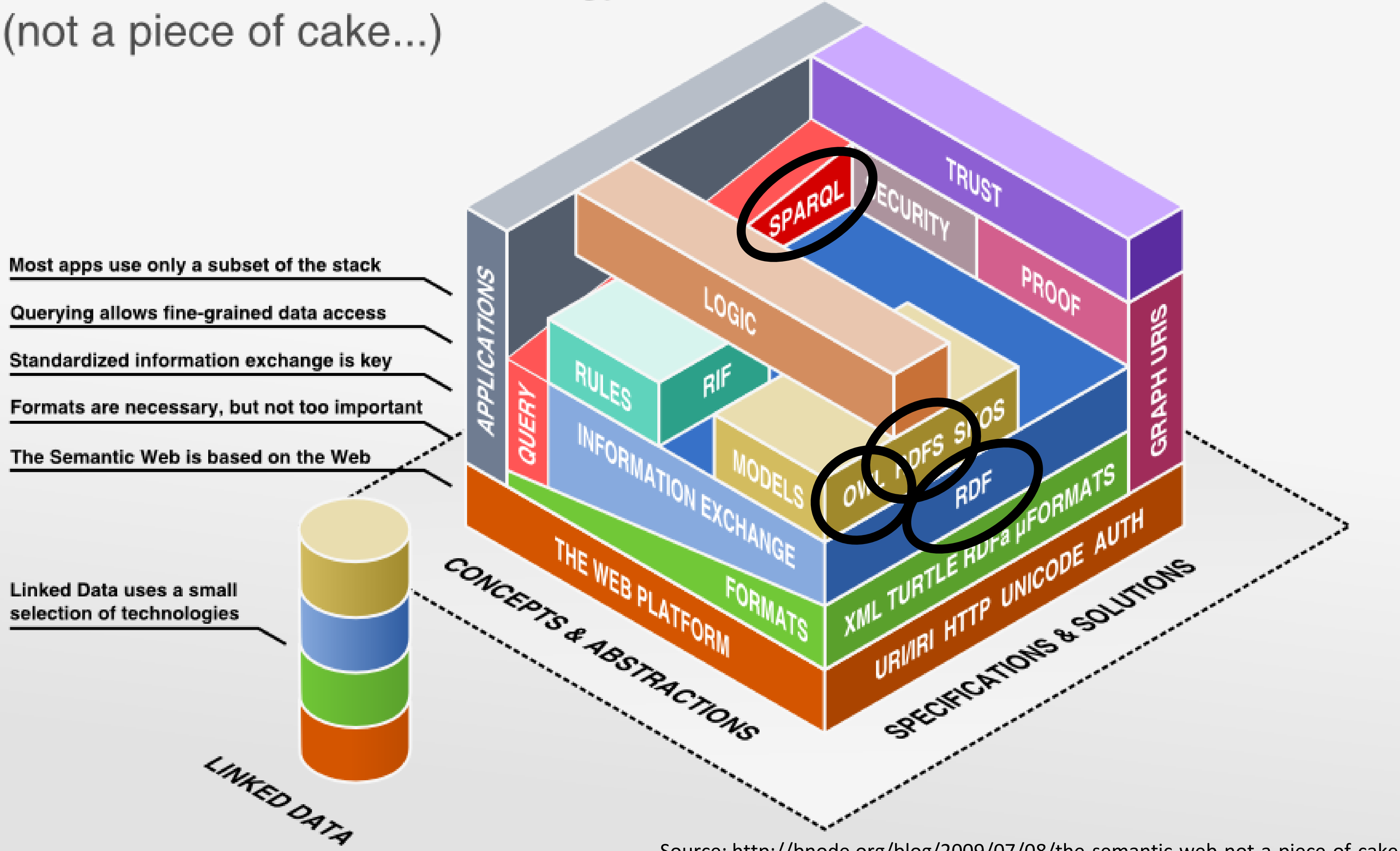
Triples representation

- A triple is a (3 tuple) an abstract representation in the form of <subject> <object> <predicate>
- The format of such representation is called RDF (**R**esource **D**escription **F**ramework)
- Triples can be encoded using text (e.g. XML, Turtle) and exchanged between different parties
- Triples from different files can be easily merged together

Technology stack overview

- Semantic Web provides a number of modelling languages that differ in their level of expressivity
  - RDF—The Resource Description Framework: provides a mechanism for allowing anyone to make a basic statement about anything and layering these statements into a single model.
  - RDFS—The RDF Schema language: is a language with the expressivity to describe the basic notions of commonality and variability familiar from object languages and other class systems—namely classes, subclasses, and properties.
  - OWL – Ontology Web Language: brings the expressivity of logic to the Semantic Web. It allows modelers to express detailed constraints between classes, entities, and properties
- SPARQL is the query language associated with Semantic Web

The Semantic Web Technology Stack
(not a piece of cake...)



Source: http://bnode.org/blog/2009/07/08/the-semantic-web-not-a-piece-of-cake

Semantic Modelling Tools

- There are a number of editors that assist users in defining semantic models
- Protégé
  - A free, open-source ontology editor and framework for building semantic models
  - http://protege.stanford.edu/
- Top Braid Composer
  - comes in multiple editions with free trials
  - http://www.topquadrant.com/tools/modeling-topbraid-composer-standard-edition/
- Jalapeno
  - part of the Capsicum Methodology for assisting enterprises in conducting business analysis
  - http://www.capsi.com.au/

Conclusion

- Semantic data modelling is an approach for representing data models using richer types of relationships
- At its most basic level, a semantic data model represents information in the form of triples
- More complex representations can be built on top of this representation
- There are a number of technologies and standards that support semantic models: RDF, RDFS, OWL, SPARQL
- Semantic models underpin the Semantic Web
- There are a number of editors that assist users in defining semantic models: Protégé, Top Braid, Jalapeno

References

- Dean Allemang and James Hendler, Semantic Web for the Working Ontologist, 2nd Edition, Morgan Kaufmann, 2011.