Semantic Web Enabled Software Engineering

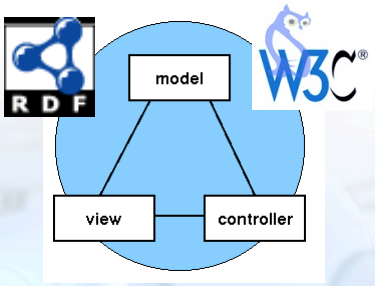
WebIST2012

Jeff Z. Pan

Department of Computing Science
University of Aberdeen, UK

How is it going, Semantic Web?

What's new in the Semantic Web

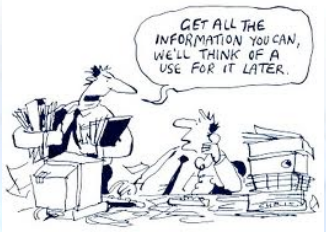


[Photo source: javaworld.com]

- **RDF**: a new **data model** for the Web
- **OWL**: standard **web ontology language** to define the meaning of vocabulary
- **SPARQL**: query language

4

SW became widely accepted



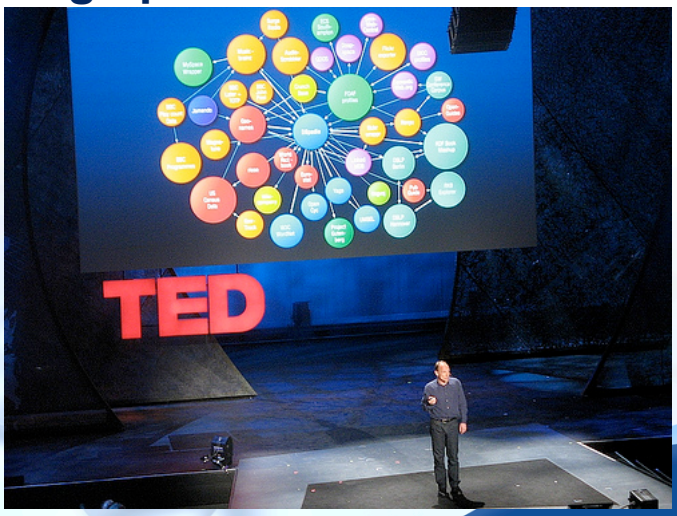
[Photo source: socialmeteor.com]

Two key steps:

1. **Building reusable ontologies**
2. **Using ontological vocabulary to annotate data**

5

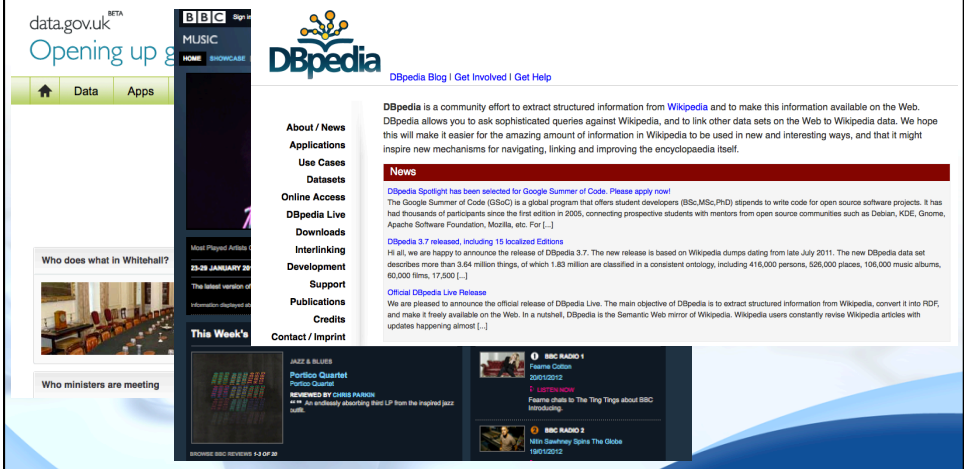
Linking open data



[Photo source: talis.com]

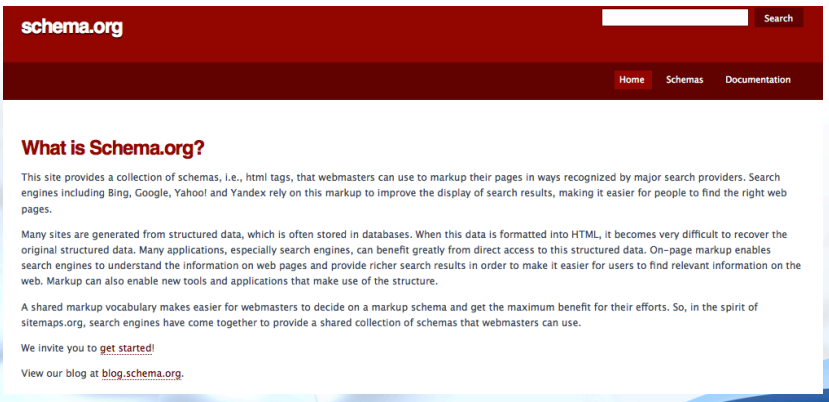
6

Linked open data



7

Schema.org



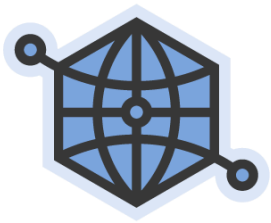
Official **OWL** ontology: **http://schema.org/docs/schemaorg.owl**

HTML microdata: **http://www.w3.org/TR/microdata/**

8

Data as rich objects

The Open Graph protocol



Introduction

The Open Graph protocol enables any web page to become a rich object in a social graph. For instance, this is used on Facebook to allow any web page to have the same functionality as any other object on Facebook.

While many different technologies and schemas exist and could be combined together, there isn't a single technology which provides enough information to richly represent any web page within the social graph. The Open Graph protocol builds on these existing technologies and gives developers one thing to implement. Developer simplicity is a key goal of the Open Graph protocol which has informed many of the technical design decisions.

9

Creating semantic data with mobiles

Your mobile knows “everything” about you:

Where you are

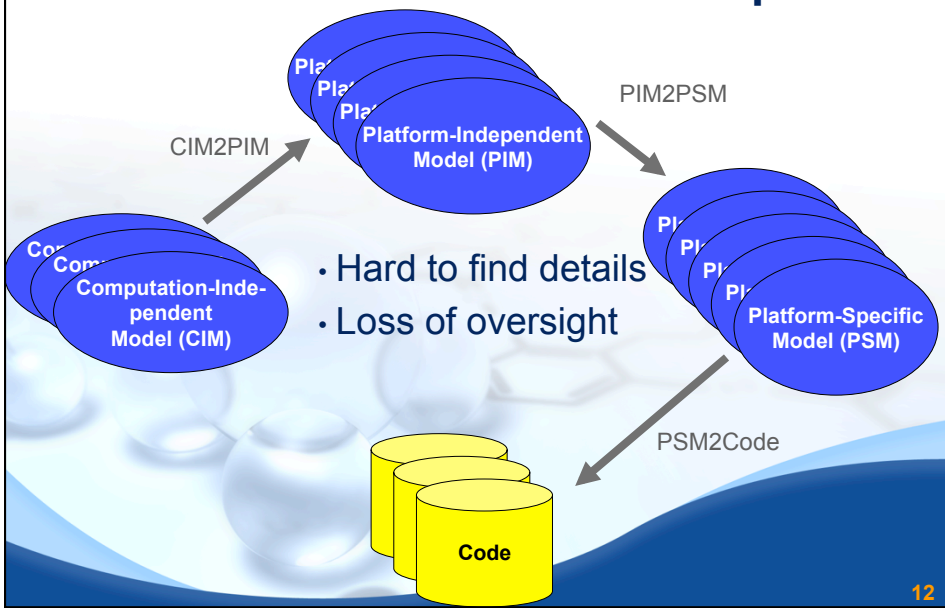
Who your social network contacts are

10

Semantic Web Enabled Software Engineering: Why and How

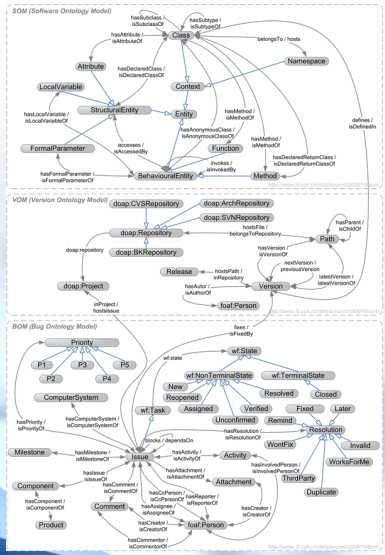
11

Model Driven Software Development



- Hard to find details
- Loss of oversight

Ontology enabled software analysis

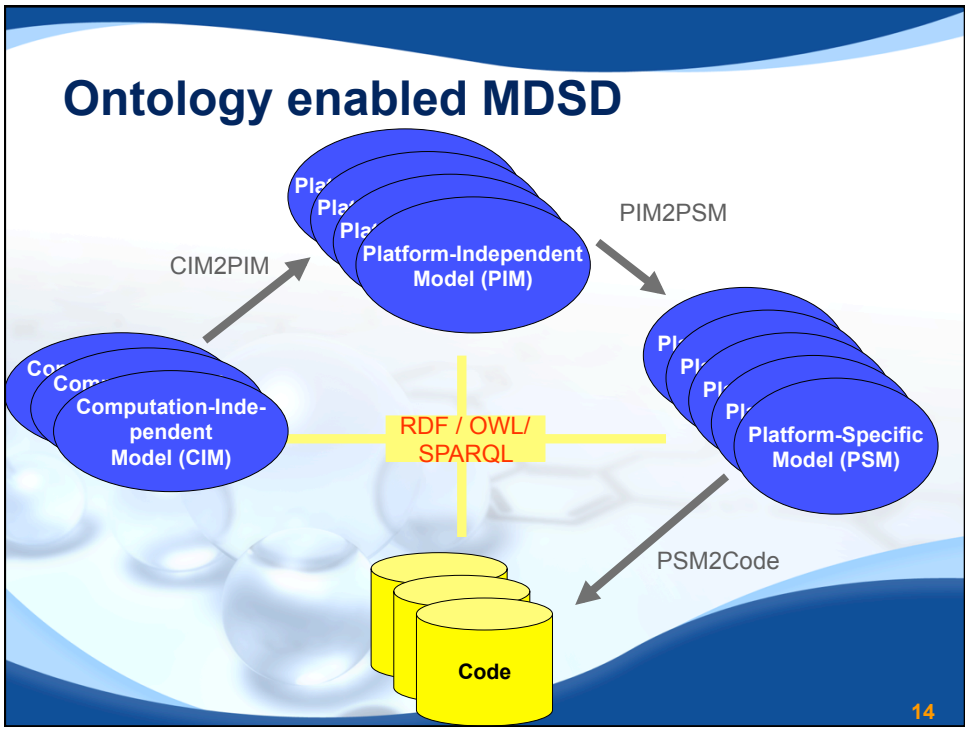


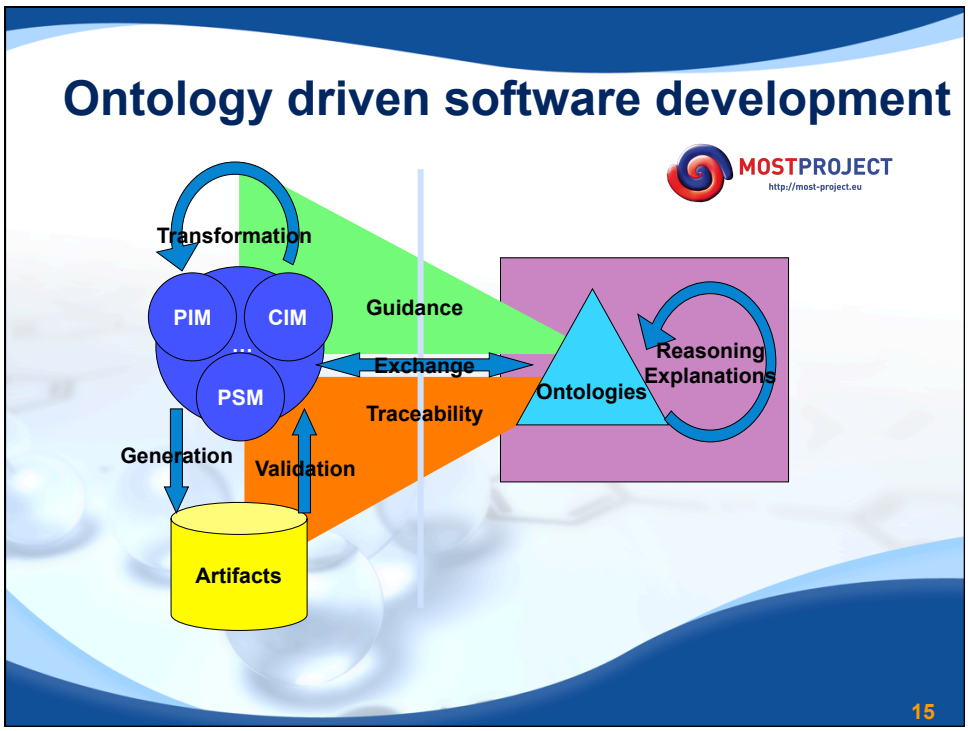
EvoOnt: A Software Evolution Ontology

- Software Analysis

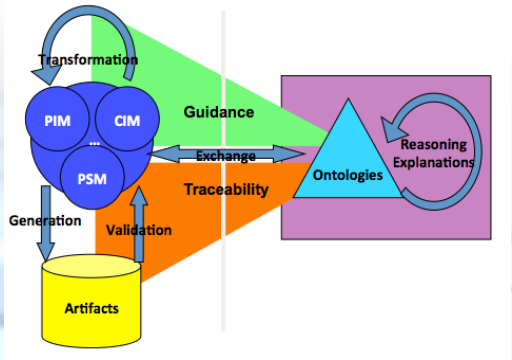
Many tasks in MSR (Mining Software Repository) Workshop 2004 - 2007 can be done by 1~2 SPARQL query

[Tappolet et al. 2010]


Ontology enabled MDSD
Platform-Independent Model (PIM)
CIM2PIM
PIM2PSM
Computation-Inde-pendent Model (CIM)
RDF / OWL/ SPARQL
Platform-Specific Model (PSM)
PSM2Code
Code
14


Ontology driven software development
MOSTPROJECT
http://most-project.eu
Transformation
PIM
CIM
PSM
Guidance
Exchange
Traceability
Ontologies
Reasoning
Explanations
Generation
Validation
Artifacts
15

How to bridge SE models and ontologies

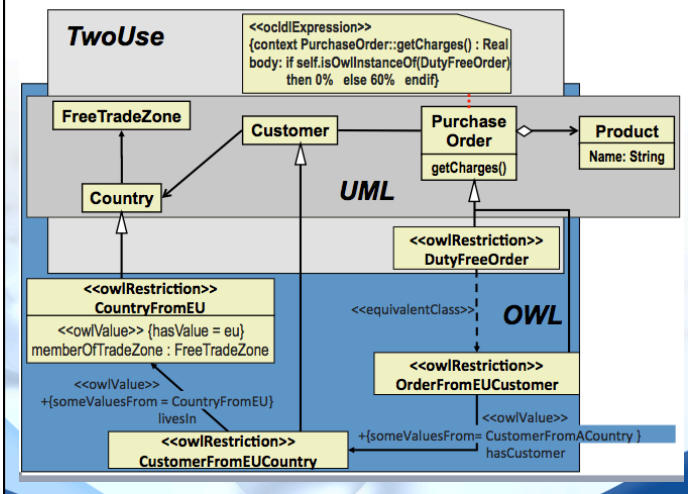


Approach 1: Integrated models

- Static SE models
- Ontologies

16

Integrated model: example 1



Two components

- UML class diagrams
- Ontologies

17

Integrated models: example 2

Two components

- Physical Device DSL (PDDSL) models
- Ontologies

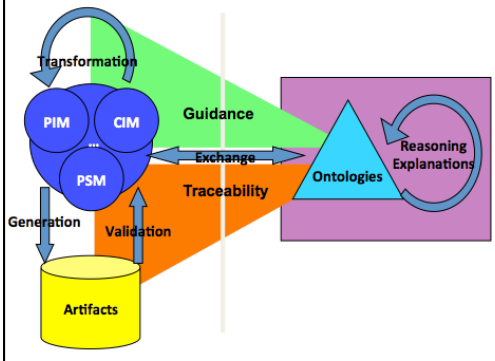
```
DeviceType "Cisco_7603"
  SubClassOf: pd_hasConfiguration some ( pd_hasSlot some ( pd_hasCard some Cisco_7600_SIP ) )
      longName : "CISCO 7603 CHASSIS" description : "The Cisco® 7603 Router is a high-performance ..." allowed : {
      PossibleConfiguration "Cisco_7603_Configuration" {
          Slot "1" allowed : "Supervisor_Engine_2" "Supervisor_Engine_720" required : true
          Slot "2" allowed : "Supervisor_Engine_2" "Supervisor_Engine_720" "Catalyst_6500_Module"
            required : false
          Slot "3" allowed : "Catalyst_6500_Module" required : false
```

```
Device serialNumber : "cisco_7603" hasType : "Cisco_7603"
configuration :
{
  Slot id : "1" : Card serialNumber : "supervisor_2_5" hasType : "WS-X6K-S2U-MSFC2"
  Slot id : "2" :
  Slot id : "3" :
}
```

[Miksa et al. 2009]

18

How to bridge SE models and ontologies

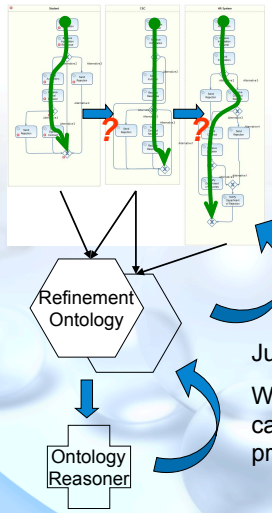


Approach 2: problem reduction to task ontologies

- To reduce an SE problem
- Into ontology reasoning problem

19

Bussiness process refinement



"Traceability":

Which activities cause the problem?

Justification:

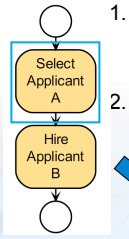
Which axioms cause the problem?

- To reduce business process refinement validation
- Into ontology consistency/coherency checking

[Ren et al. 2009b; Groener and Staab 2009]

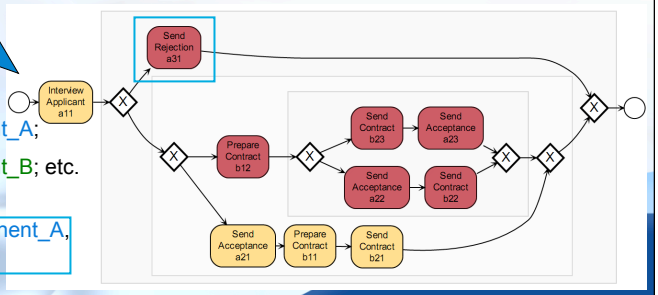
20

Bussiness process refinement: example



1. In pre-refinement process:
   Component_A subclassOf to only (Component_A or Component_B);
   Component_B subclassOf from only (Component_B or Component_A);
2. In post-refinement process:
   a31 subclassOf to some End
   b12 subclassOf (to some b23) and (to some a22), etc.

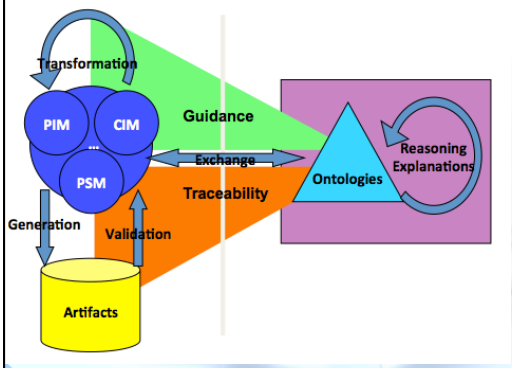
Invalid



3. Mapping:
   a31 subclassOf Component_A;
   b23 subclassOf Component_B; etc.
4. Uniqueness:
   Disjoint(Start, End, Component_A, Component_B), etc.

21

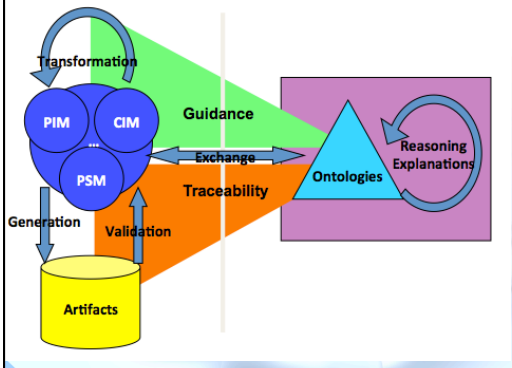
Traceability: exploiting the integrated models and task ontologies



- Help understand the associations and dependencies
- By exploiting reasoning and query answering

22

Dynamics and complete of knowledge



Two types of knowledge

- Some we have complete understanding (DB)
- Some we only know partially (SW)

23

Semantic Web supports BOTH

NBox (Negation As Failure Box) enabled ontologies **[JTS]**

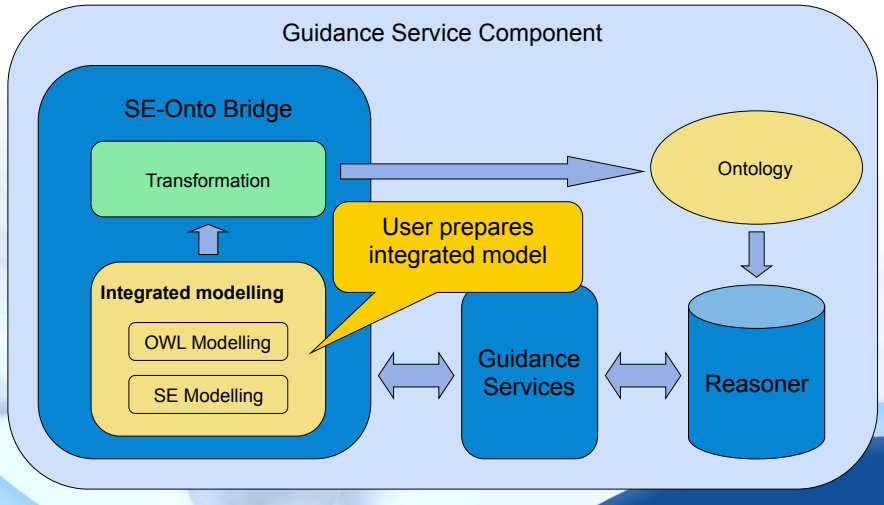
- TBox: **schema** axioms
- ABox: **data** axioms
- **NBox**: a set of closed vocabulary
  - NAF of NBox concepts can be simulated by full negation
    - Such as Spicy
  - Require **incremental reasoning**

Food	Note
Curry Chicken	Minor Spicy
Salmon Fillet	
Spicy Grilled Shrimp	Spicy
Pepper Salad	Vege

Name	Vegetarian
Jeff	No
Yuting	Yes
Jek	
Yuan	No

24

Architecture



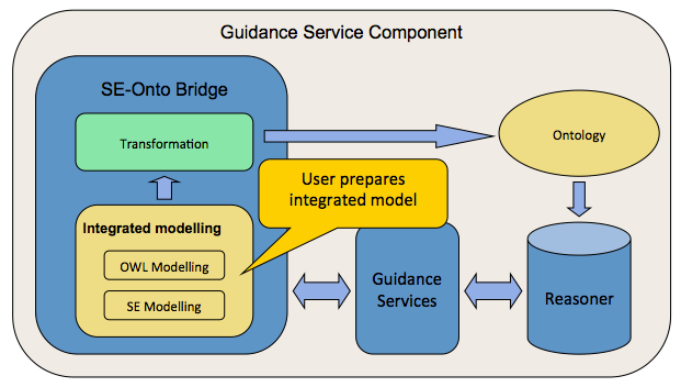
[Ren et al. 2009]

25

Guidance services

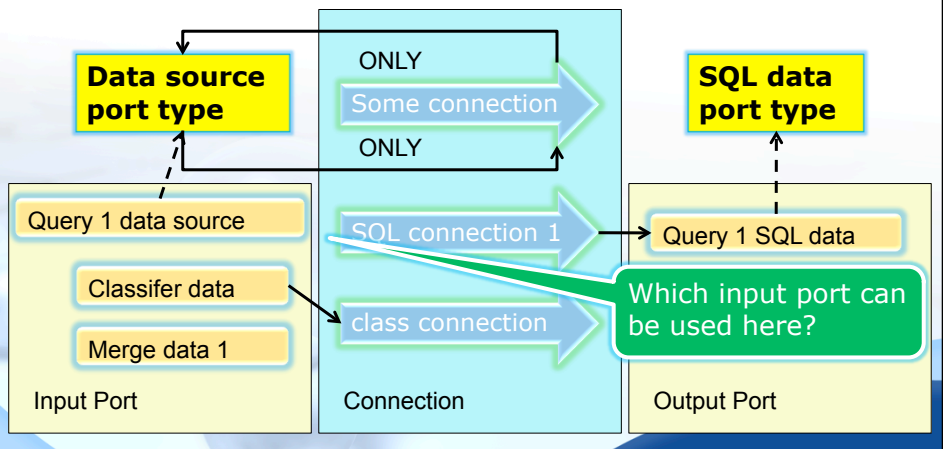
Todo list services for developers:

- To recommend the next available actions/options
- To validate the current action/option

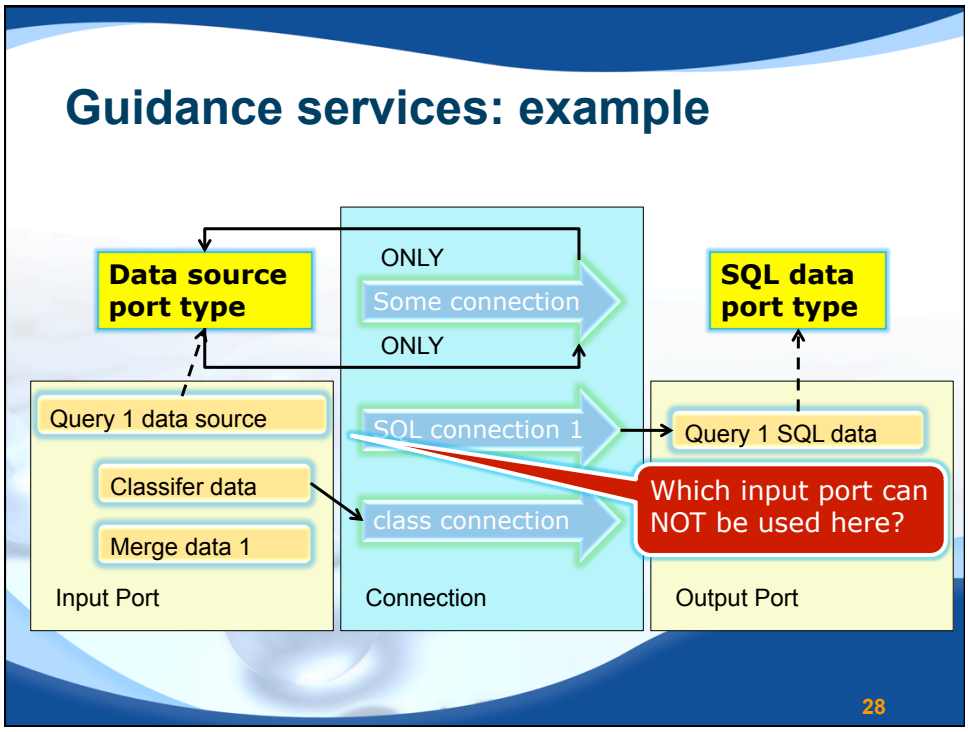


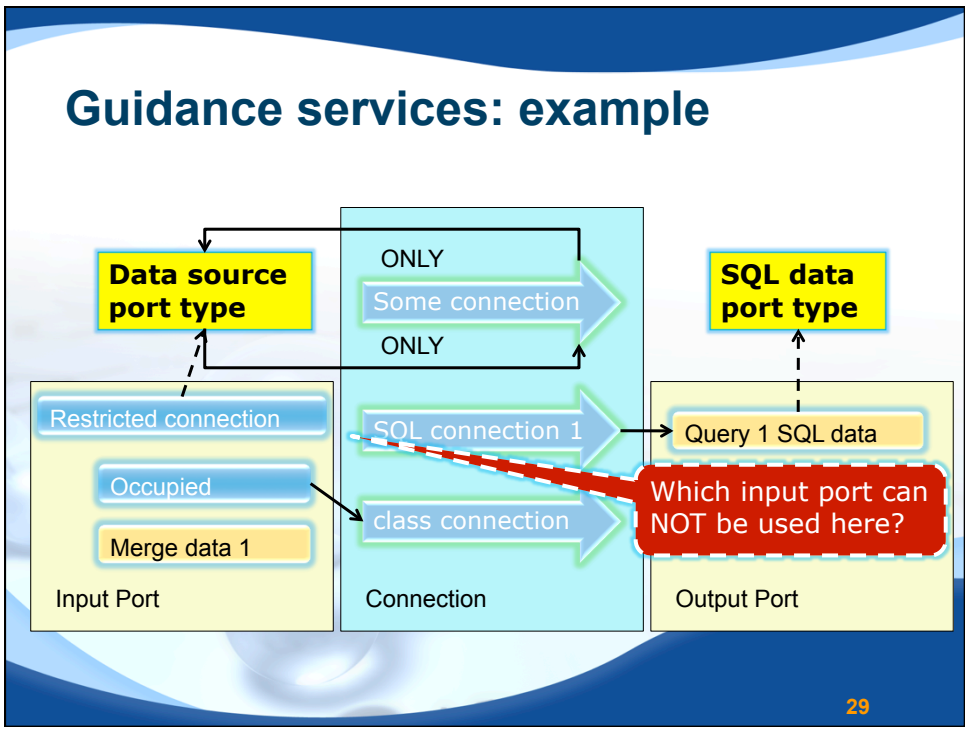
26

Guidance services: example

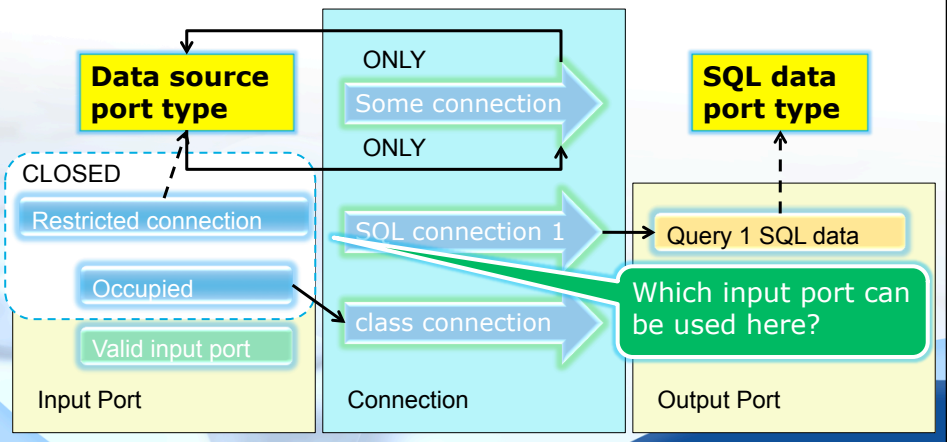


27


Guidance services: example
Data source port type
ONLY
Some connection
ONLY
SQL data port type
Query 1 data source
SQL connection 1
Query 1 SQL data
Classifer data
Which input port can NOT be used here?
class connection
Merge data 1
Input Port
Connection
Output Port
28

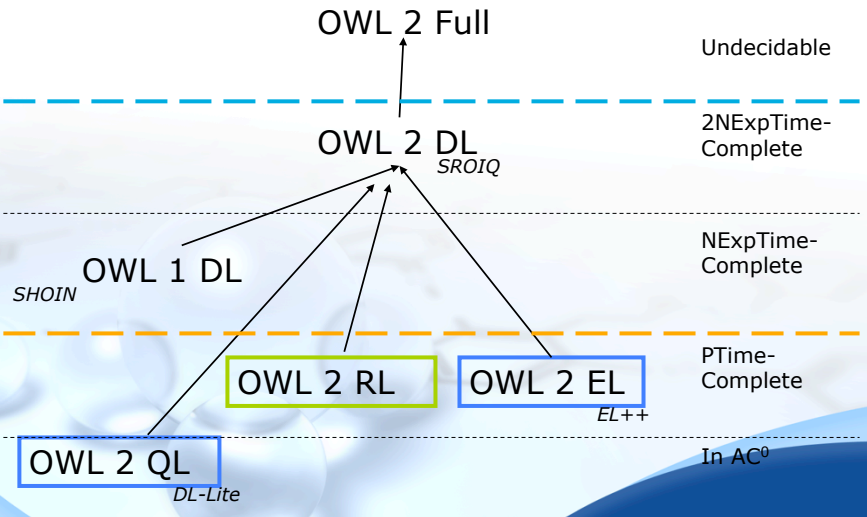

Guidance services: example
Data source port type
ONLY
Some connection
ONLY
SQL data port type
Restricted connection
SQL connection 1
Query 1 SQL data
Occupied
Which input port can NOT be used here?
class connection
Merge data 1
Input Port
Connection
Output Port
29

Guidance services: example



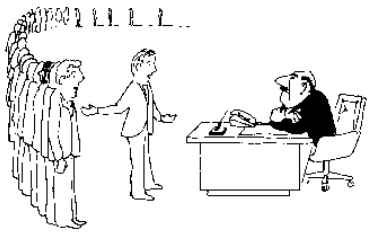
30

OWL: Standard Ontology Language



31

Approximate reasoning plays a key role



Inputs
- Ontology
- Queries

Outputs
- soundness preserving
- completeness preserving

I can't find any algorithms, but neither can all these famous people.

[Garey & Johnson. Computers and Intractability: A Guide to the Theory of NP-Completeness. Freeman, 1979.]

32

Faithful Syntactic Approximate Reasoning [AAAI2010]

Approximation
- Directly represent non-OWL2-EL concepts with fresh named concepts
  - E.g., ∀r.C subClassOf D ➔ $A_{\forall r.C}$ subClassOf D
- Maintain semantic relations for these named concepts
  - *complementary relations*
  - *cardinality relations*

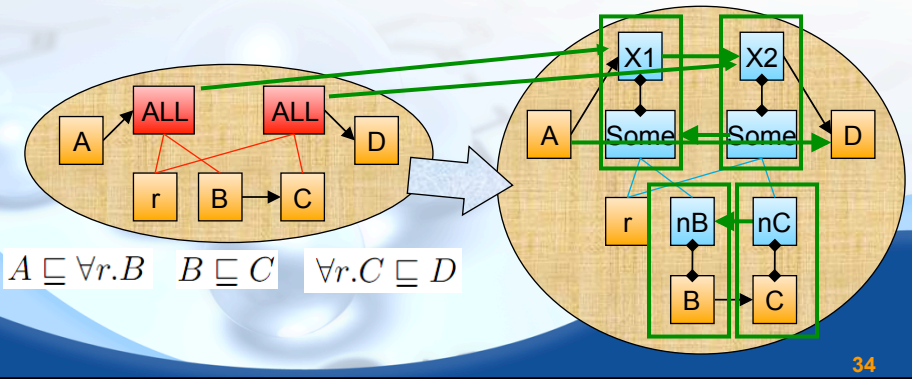
Reasoning
- Using additional **tractable** completion rules to recover the semantics

33

Example: How Does it Work

Additional completion Rules (on top of the EL ones), e.g.

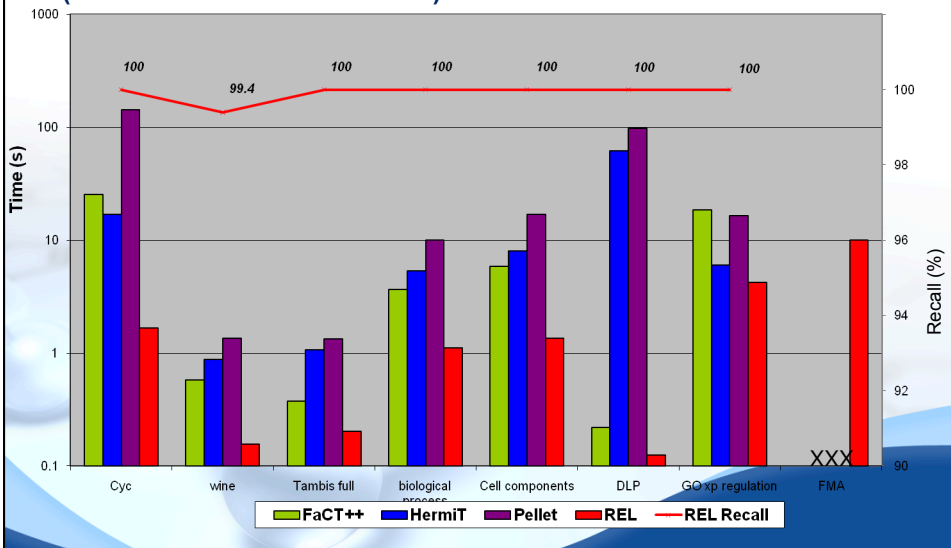
- Handling complement
  - E.g. B subClassOf C => ¬C subClassOf ¬B
- Handling cardinality
  - E.g. A subClassOf >= 3 r. B => A subClassOf >= 2 r. B
- **Soundness preserving and tractable**



$A \sqsubseteq \forall r.B \quad B \sqsubseteq C \quad \forall r.C \sqsubseteq D$

34

Evaluations for the Oxford Benchmarks (REL – EL reasoner in TrOWL)



35

TrOWL: a tractable semantic reasoning infrastructure

Approximate reasoning [AAAI2007, AAAI2010]

Parallel reasoning [JIST2011]

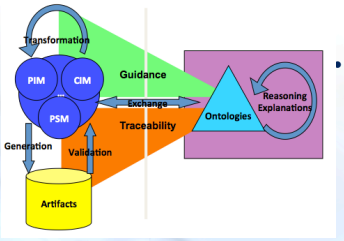
TrOWL
www.trowl.eu

Stream reasoning [CIKM2011]

Local closed world reasoning [JTS]

36

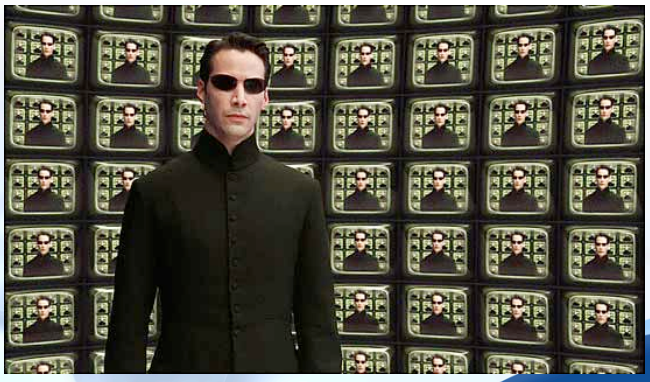
Benefits for Software Engineering



- Integrated modelling and querying
- Guidance services for developers:
  - To recommend the next available actions
  - To validate the current action
- Ontologies available at deployment time

37

Vision: Programs that can understand and improve themselves



[Photo source: northjersey.com]

API4KB

Possible impacts on the Semantic Web

- More data
  - Linked software data
- Easier to debug, reuse and extend semantic applications
- Software becomes semantics software
- Software/service as a first class citizen like data?

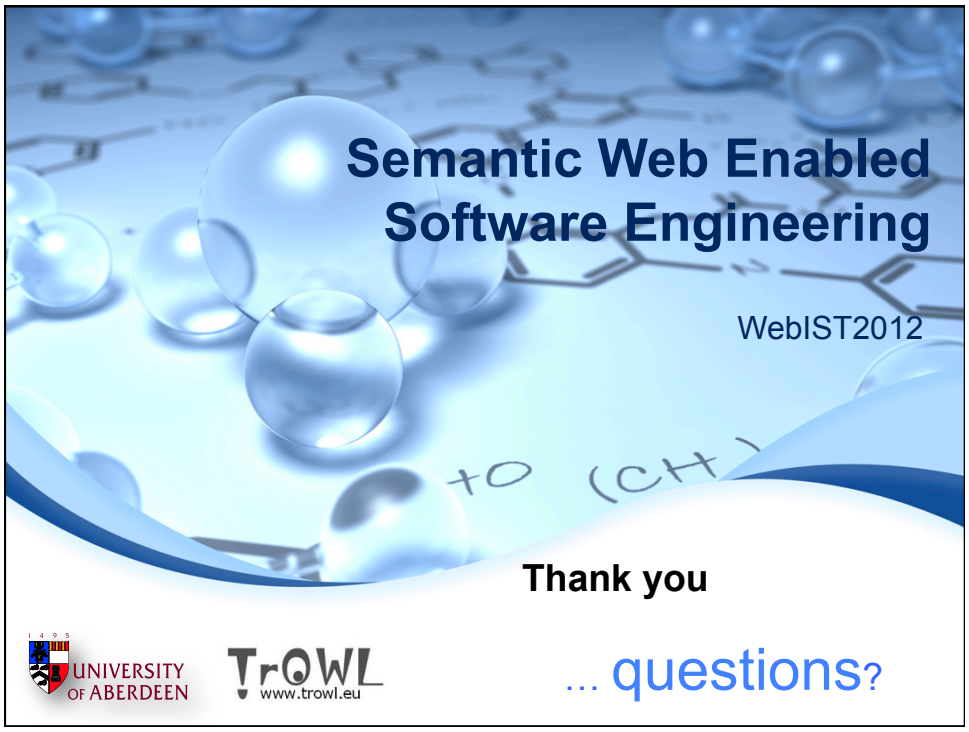
40

Where to find more information

International SWESE workshop series
http://www.abdn.ac.uk/~csc280/event/workshop/swese2011/

Two books coming:

- **J Z. Pan,** S. Staab, U. Assmann, J. Ebeert, Y. Zhao (Eds.), Ontology Driven Software Development. Springer. 2012
- **J. Z. Pan** and Y. Zhao (Eds.), Semantic Web Enabled Software Engineering. AKA Verlag Heidelberg and IOS Press, Amsterdam, The Netherlands. 2012


Semantic Web Enabled Software Engineering
WebIST2012
Thank you
UNIVERSITY OF ABERDEEN
TrOWL
www.trowl.eu
... questions?