Semantic Evaluation at Large Scale Tutorial

Evaluating Ontology Engineering Tools

Raúl García-Castro

Ontology Engineering Group

Departamento de Lenguajes y Sistemas Informáticos e Ingeniería de Software, Facultad de Informática

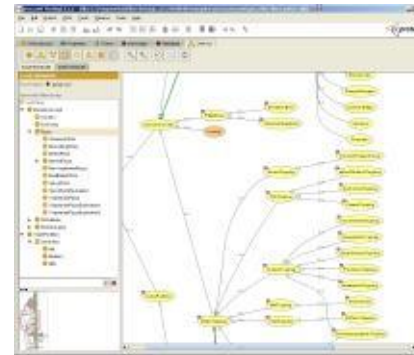Universidad Politécnica de Madrid, Spain

rgarcia@fi.upm.es

Index

- **Evaluation scenarios**
- Test data
- Evaluation descriptions
- Tools
- Results
- Conclusions

SEALS
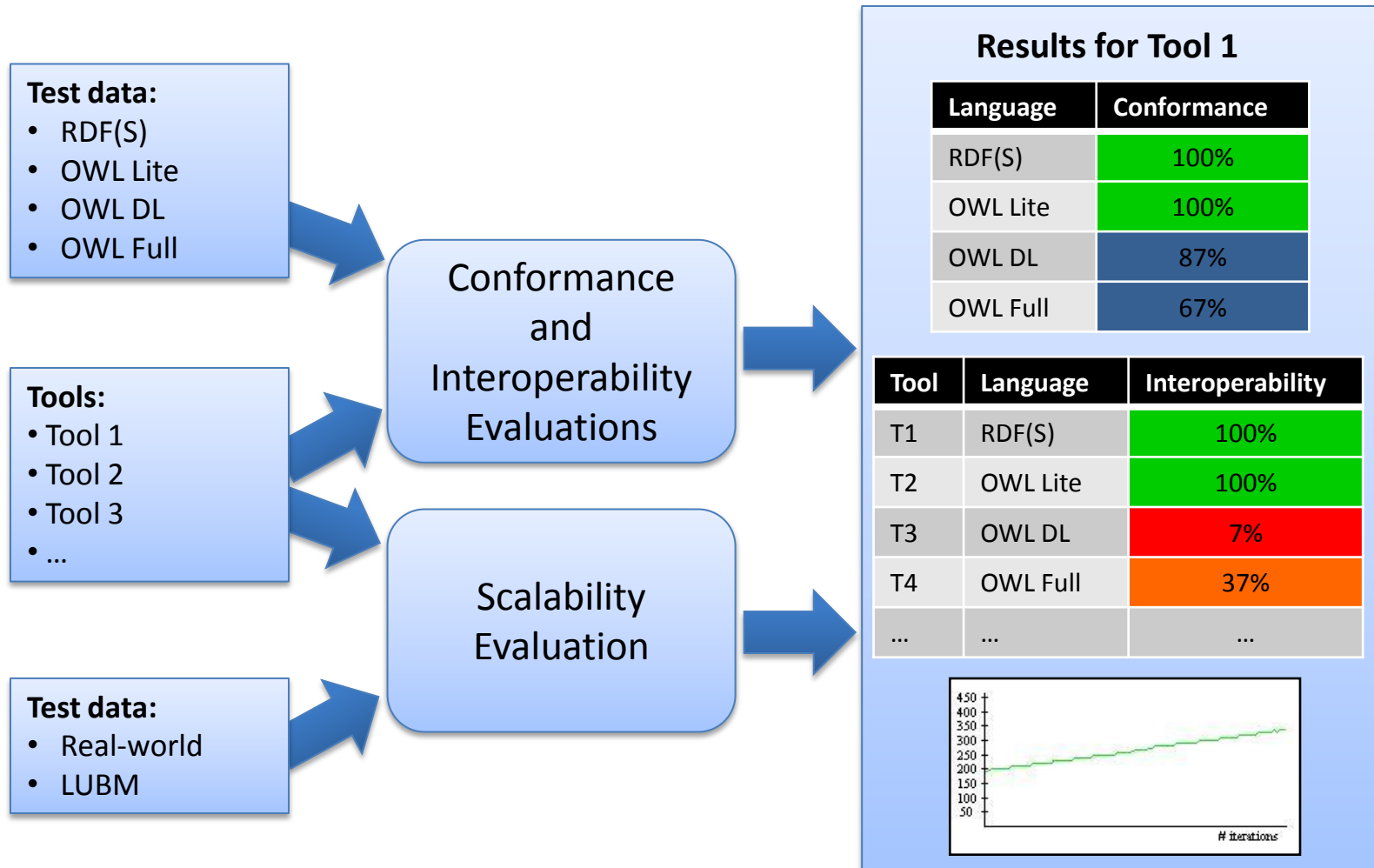Semantic Evaluation at Large Scale

Ontology Engineering Tools

Allow the creation and management of ontologies:

- Ontology editors
  - User oriented

- Ontology language APIs
  - Programming oriented

SEALS

Semantic Evaluation at Large Scale

Evaluation scenarios

Test data:
- RDF(S)
- OWL Lite
- OWL DL
- OWL Full

Tools:
- Tool 1
- Tool 2
- Tool 3
- …

Test data:
- Real-world
- LUBM

Conformance and Interoperability Evaluations

Scalability Evaluation

Results for Tool 1

Language	Conformance
RDF(S)	100%
OWL Lite	100%
OWL DL	87%
OWL Full	67%

Tool	Language	Interoperability
T1	RDF(S)	100%
T2	OWL Lite	100%
T3	OWL DL	7%
T4	OWL Full	37%
…	…	…

SEALS
Semantic Evaluation at Large Scale

Conformance evaluation

- **Goal**: to evaluate the conformance of semantic technologies with regards to ontology language specifications

Step 1: Import + Export

$$O1 = O1'' + \alpha - \alpha'$$
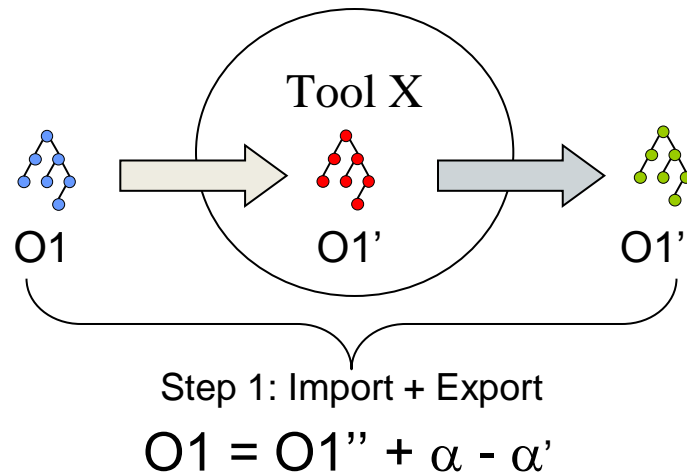
- **Metrics**: Information addition and loss when processing ontologies

SEALS

Semantic Evaluation at Large Scale

Interoperability evaluation

- **Goal**: to evaluate the interoperability of semantic technologies in terms of the ability that such technologies have to interchange ontologies and use them

Step 1: Import + Export
$$O1 = O1'' + \alpha - \alpha'$$

Step 2: Import + Export
$$O1''=O1'''' + \beta - \beta'$$
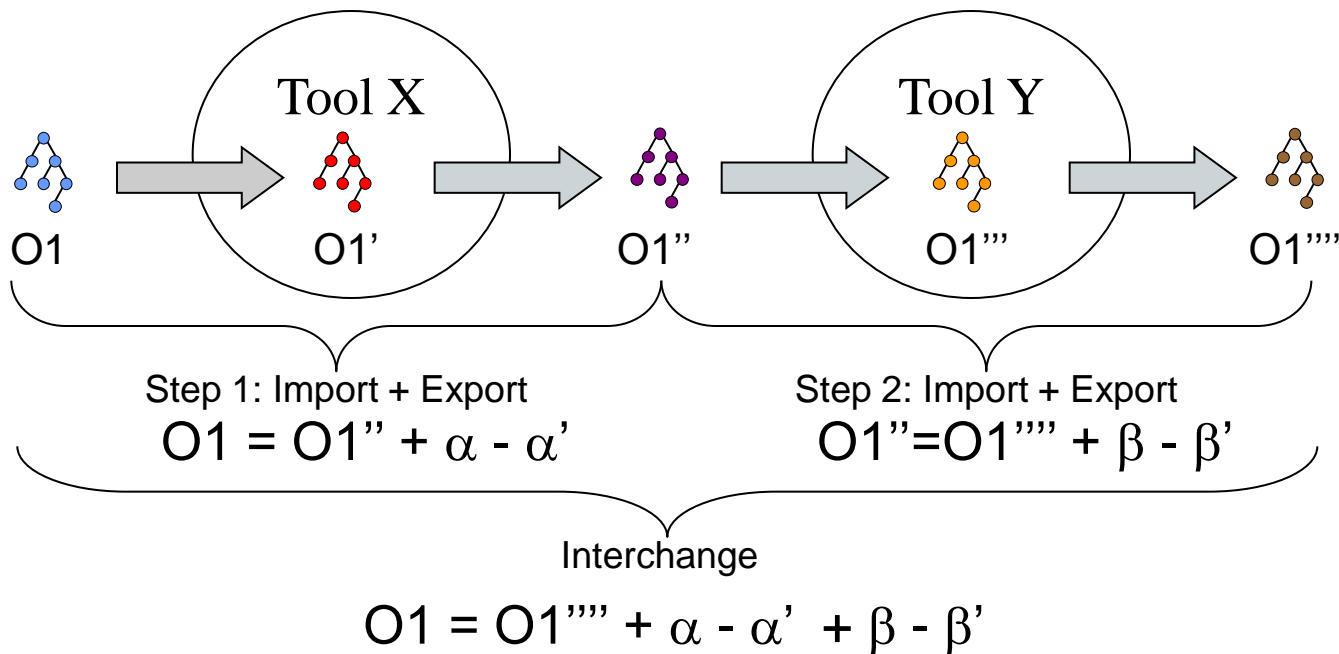
Interchange
$$O1 = O1'''' + \alpha - \alpha' + \beta - \beta'$$

- **Metrics**: Information addition and loss when interchanging ontologies

SEALS
Semantic Evaluation at Large Scale

Scalability evaluation

- **Goal**: to evaluate the ability to maintain performance under increasing loads
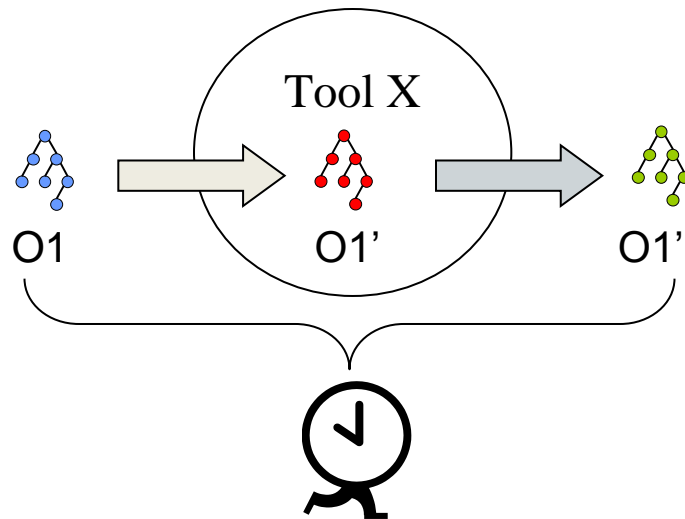
- **Metric**: Processing times for ontologies of increasing size

Index

- Evaluation scenarios
- **Test data**
- Evaluation descriptions
- Tools
- Results
- Conclusions

SEALS
Semantic Evaluation at Large Scale

Conformance and interoperability test suites

RDF(S)

Group	No.
Class	2
Metaclass	5
Subclass	5
Class and property	6
Property	2
Subproperty	5
Property with domain and range	24
Instance	4
Instance and property	14
Syntax and abbreviation	15
TOTAL	82

OWL Lite

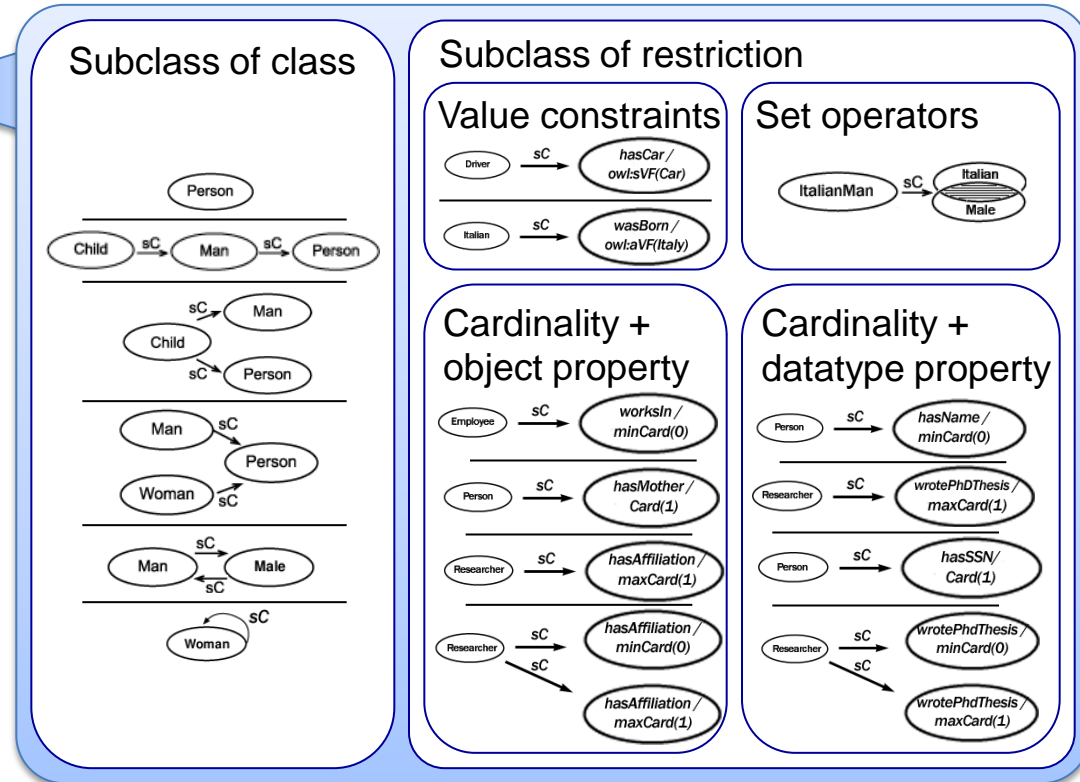| Group	No.
A - Class hierarchies	17
B - Class equivalences	12
C - Classes defined with set operators	2
D - Property hierarchies	4
E - Properties with domain and range	10
F - Relations between properties	3
G - Global cardinality constraints and logical property characteristics	5
H - Single individuals	3
I - Named individuals and properties	5
J - Anonymous individuals and properties	3
K - Individual identity	3
L - Syntax and abbreviation	15
TOTAL	82

OWL DL

Group	Number
Class descriptions	74
Class axioms	96
Combinations of class axioms	14
Property descriptions	8
Combinations of property axioms	24
Properties with domain and range	58
Individual descriptions	38
Combinations of individual axioms	9
Individuals and properties	11
Data ranges	76
Annotation properties	153
TOTAL	561

OWL Full

Group	Number
Ontology Header Relaxations	5
Construct Encoding Relaxations	3
Entity Type Declaration Relaxations	6
Entity Types as Classes	6
Non-Separated Entity Types	10
Metamodelling	12
Data Properties as Object Properties	5
Semantic Annotation Properties	6
Datatypes as Classes	7
Unrestricted Use of Literals	6
Unrestricted Use of Blank Nodes	8
Unrestricted Use of Complex Properties	1
RDF Lists Usage	4
Logical Vocabulary Reflection	11
TOTAL	90

SEALS
Semantic Evaluation at Large Scale

Example:
OWL Lite Import Test Suite

Group	No.
Class hierarchies	17
Class equivalences	12
Classes defined with set operators	2
Property hierarchies	4
Properties with domain and range	10
Relations between properties	3
Global cardinality constraints and logical property characteristics	5
Single individuals	3
Named individuals and properties	5
Anonymous individuals and properties	3
Individual identity	3
Syntax and abbreviation	15
TOTAL	82

Subclass of class

Subclass of restriction

Value constraints

Driver → sC → hasCar / owl:sVF(Car)

Italian → sC → wasBorn / owl:aVF(Italy)

Set operators

ItalianMan → sC → Italian / Male

Cardinality + object property

Employee → sC → worksIn / minCard(0)

Person → sC → hasMother / Card(1)

Researcher → sC → hasAffiliation / maxCard(1)

Researcher → sC → hasAffiliation / minCard(0), sC → hasAffiliation / maxCard(1)

Cardinality + datatype property

Person → sC → hasName / minCard(0)

Researcher → sC → wrotePhDThesis / maxCard(1)

Person → sC → hasSSN / Card(1)

Researcher → sC → wrotePhdThesis / minCard(0), sC → wrotePhDThesis / maxCard(1)

SEALS
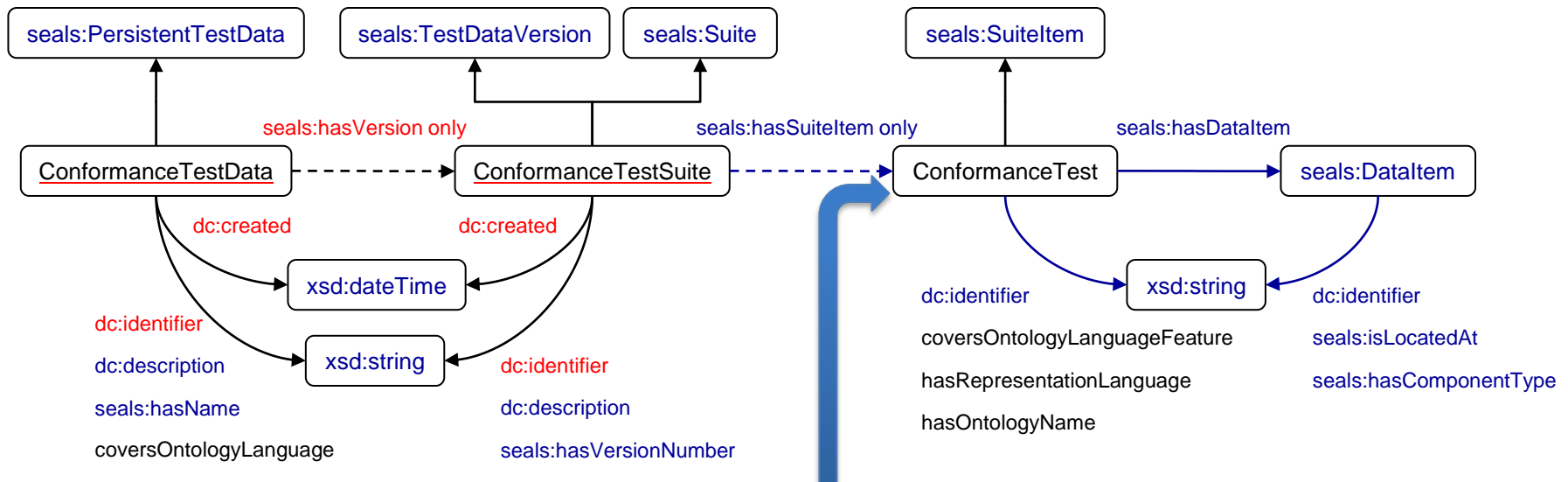**Semantic Evaluation at Large Scale**

Scalability test suites

Real-world

Ontology	Name	Size	Collection
Wine ontology	ontology001	82KB	Wine
Athletics events ontology	ontology002	758KB	AOE
Geographic Information ontology	ontology003	119KB	AOE
Multimedia content ontology	ontology004	33KB	AOE
Multimedia descriptor ontology	ontology005	274KB	AOE
FMA	ontology006	37.7MB	FMA
Chemical Information ontology	ontology007	29KB	OBO Foundry
Infectious disease ontology	ontology008	233KB	OBO Foundry
Influenza ontology	ontology009	164KB	OBO Foundry
Information Artifact ontology	ontology010	4KB	OBO Foundry
Lipid ontology	ontology011	1.3MB	OBO Foundry
Gene Regulation ontology	ontology012	430KB	OBO Foundry
NIF-Cell	ontology013	476KB	OBO Foundry
Ontology for biomedical investigations ontology	ontology014	4.1MB	OBO Foundry
Subcellular anatomy ontology	ontology015	881KB	OBO Foundry
Vaccine ontology	ontology016	1.3MB	OBO Foundry
Digestive ontology	ontology017	14.6MB	GALEN
Paedriatic ontology	ontology018	12.6MB	GALEN
Sensory ontology	ontology019	13.9MB	GALEN
Skin plastic ontology	ontology020	13.8MB	GALEN

LUBM

Ontology	Name	Size
One generation unit ontology	university1	8212KB
Two generation units ontology	university2	18907KB
Three generation units ontology	university3	27745KB
Four generation units ontology	university4	39380KB

SEALS

Semantic Evaluation at Large Scale

Example:
Describing conformance test data


```
<rdf:Description rdf:about="&conf;OWLDLImportTest049">
    <rdf:type rdf:resource="&conf;ConformanceTest"/>
    <rdf:type rdf:resource="&seals;SuiteItem"/>
    <dc:identifier>DLTestOWLDLImportTest049</dc:identifier>
    <j.3:coversOntologyLanguageFeature>Class intersection</j.3:coversOntologyLanguageFeature>
    <j.3:coversOntologyLanguageFeature>Named class</j.3:coversOntologyLanguageFeature>
    <j.3:coversOntologyLanguageFeature>Cardinality restriction</j.3:coversOntologyLanguageFeature>
    <j.3:coversOntologyLanguageFeature>Object property</j.3:coversOntologyLanguageFeature>
    <j.3:hasRepresentationLanguage>OWL DL</j.3:hasRepresentationLanguage>
    <j.3:hasOntologyName>ISJ03.owl</j.3:hasOntologyName>
    <j.3:hasDataItem rdf:resource="&conf;OWLDLImportTestDocument049"/>
</rdf:Description>
```

Test data overview

Conformance + Interoperability Test Data

Name	Type	Nº Tests
RDF(S) Import Test Suite	*Synthetic, hand-crafted*	82
OWL Lite Import Test Suite	*Synthetic, hand-crafted*	82
OWL DL Import Test Suite	*Synthetic, generated*	561
OWL Full Import Test Suite	*Synthetic, hand-crafted*	90

Scalability Test Data

Name	Type	Nº Tests
Real-world ontologies	*Application-built*	20
Generated ontologies	*Synthetic, generated*	4

SEALS
Semantic Evaluation at Large Scale

Index

- Evaluation scenarios
- Test data
- **Evaluation descriptions**
- Tools
- Results
- Conclusions
- Links to resources

SEALS
Semantic Evaluation at Large Scale

Evaluation overview

Test Suite

Tool X

O1   O1'   O1''

Step 1: Import + Export

$$O1 = O1'' + \alpha - \alpha'$$

Raw results

Interpretations

SEALS
Semantic Evaluation at Large Scale
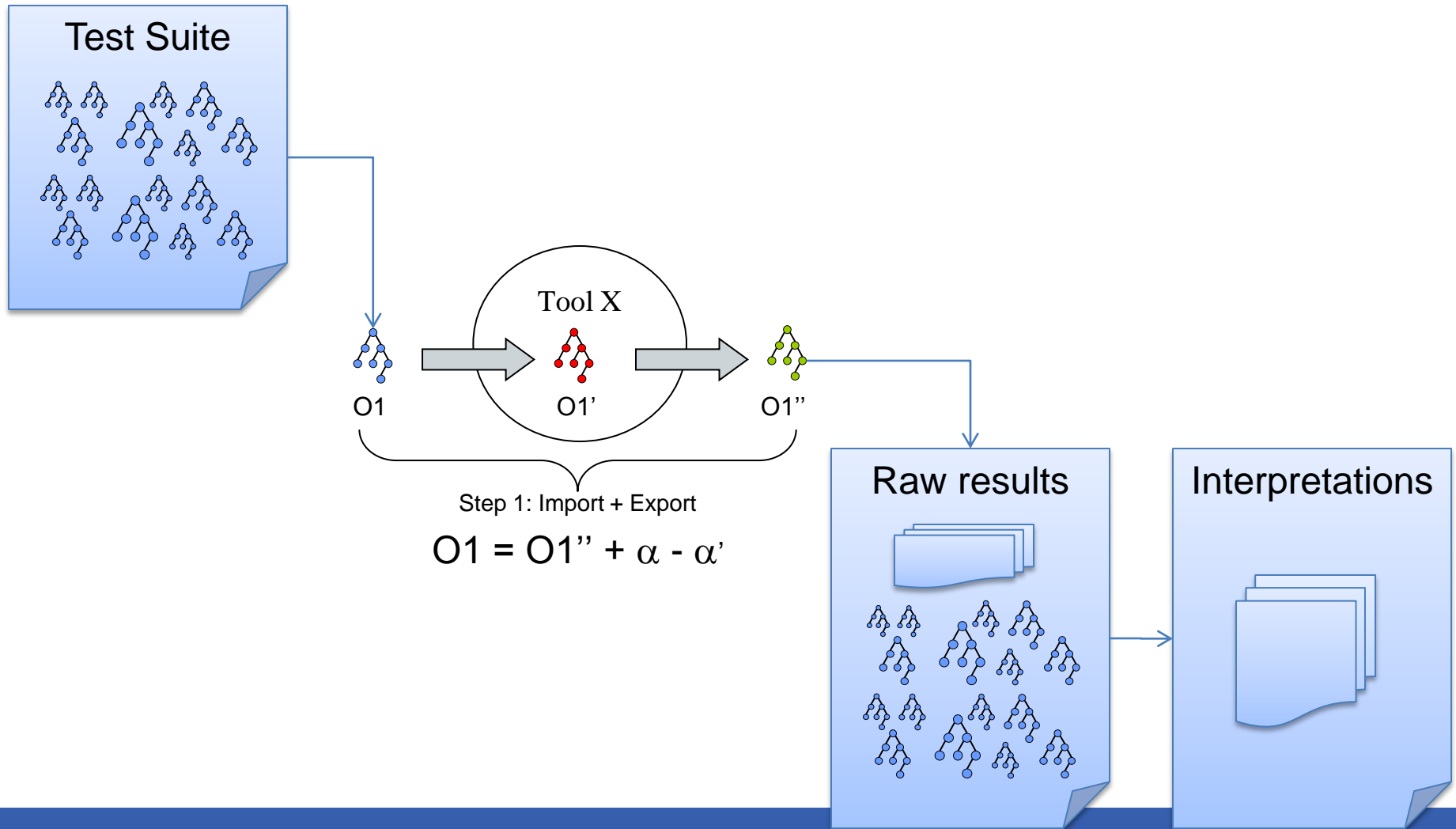
Implementing the evaluation
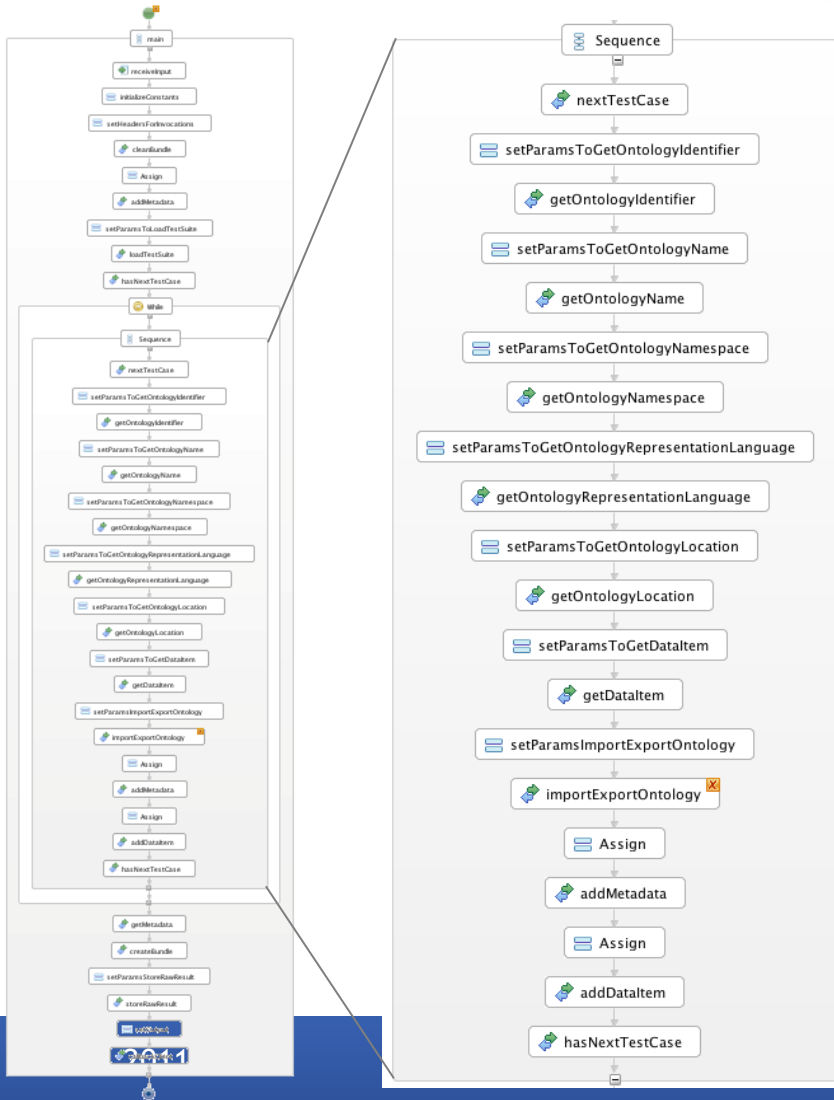
- BPEL workflow using
  - Platform services:
    - Test Data Repository
    - Result Repository
    - Result bundling
  - My services
    - Tool invocation
    - Interpretation
    - Any other custom service

SEALS

Semantic Evaluation at Large Scale

Example:
Describing conformance raw results

executionProblemInTool	false
executionProblemInPlatform	false

finalOntology	`<rdf:RDF`
	`  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"`
	`  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"`
	`  xmlns:owl="http://www.w3.org/2002/07/owl#>`
	
	`  <owl:Ontology rdf:about="#" />`
	`  <owl:Class rdf:about="http://www.example.org/ISA04#Man">`
	`    <rdfs:subClassOf rdf:resource="http://www.example.org/ISA04#Person"/>`
	`  </owl:Class>`
	`  <owl:Class rdf:about="http://www.example.org/ISA04#Woman">`
	`    <rdfs:subClassOf rdf:resource="http://www.example.org/ISA04#Person"/>`
	`  </owl:Class>`
	`  <owl:Class rdf:about="http://www.example.org/ISA04#Person" />`
	`</rdf:RDF>`

SEALS
Semantic Evaluation at Large Scale

Example:
Describing conformance interpretations

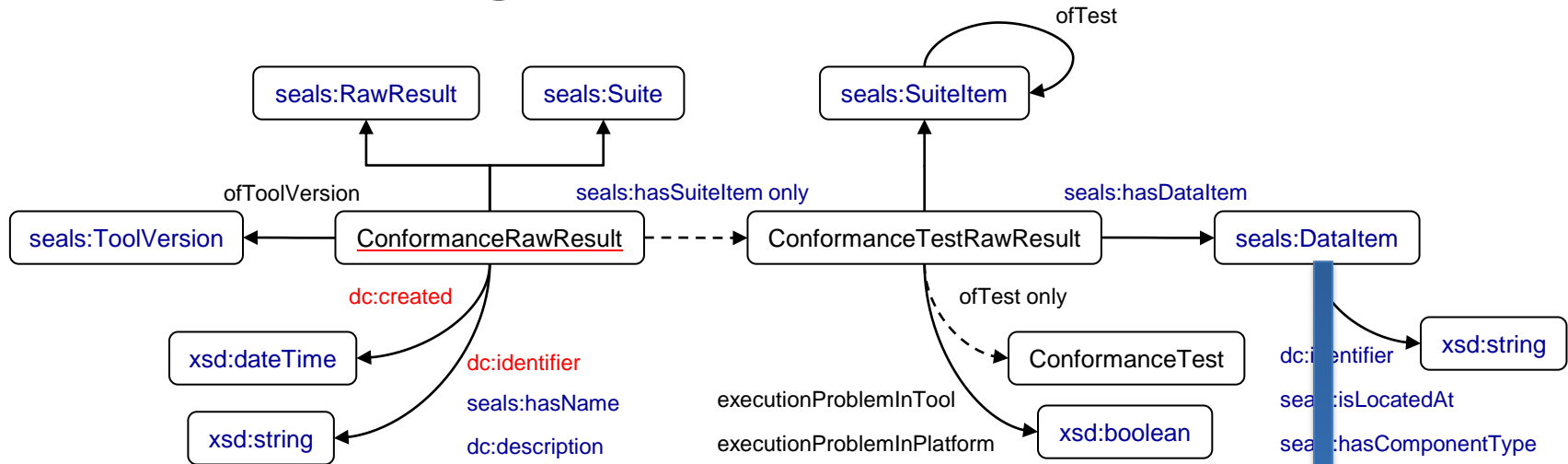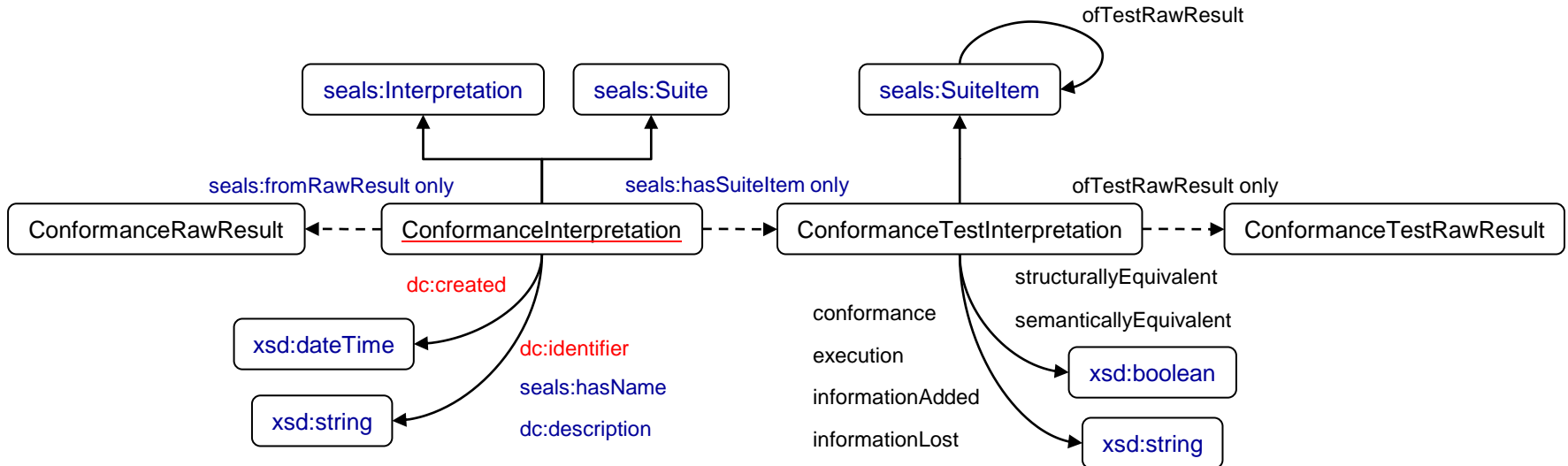| conformance	true
execution	true
informationAdded	`ont:A rdfs:subClassOf ont:C.`
informationLost	-
structurallyEquivalent	false
semanticallyEquivalent	true

SEALS

Semantic Evaluation at Large Scale

Index

- Evaluation scenarios
- Test data
- Evaluation descriptions
- **Tools**
- Results
- Conclusions

SEALS
Semantic Evaluation at Large Scale

Connecting an ontology engineering tool

- Implementation of a Java plugin with:
  - Tool Management API:
    - Deployment
    - Undeployment
    - Start (optional)
    - Stop (only if start)

    Java apps or shell scripts

  - Tool invocation API:
    - *importExportOntology*
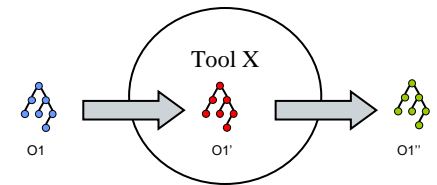

```
public interface OntologyEngineeringTool extends Tool{

    public URI importExportOntology(URI importFileURI, String ontologyLanguage)
              throws Exception;
}
```

SEALS
Semantic Evaluation at Large Scale

Packaging my tool

```
./descriptor.xml
./bin/<executable file>
./bin/...
./conf/<configuration file>
./conf/...
./lib/<tool distribution file>
./lib/...
```

Tool Bundle
(ZIP)

Maven archetype for tool bundle creation

Tool bundle description	
Description	• Id • Version • License • ...
Wrapper	• Tool management • Tool invocation • Dependencies
Requirements	• Modules • Configuration files

SEALS
Semantic Evaluation at Large Scale
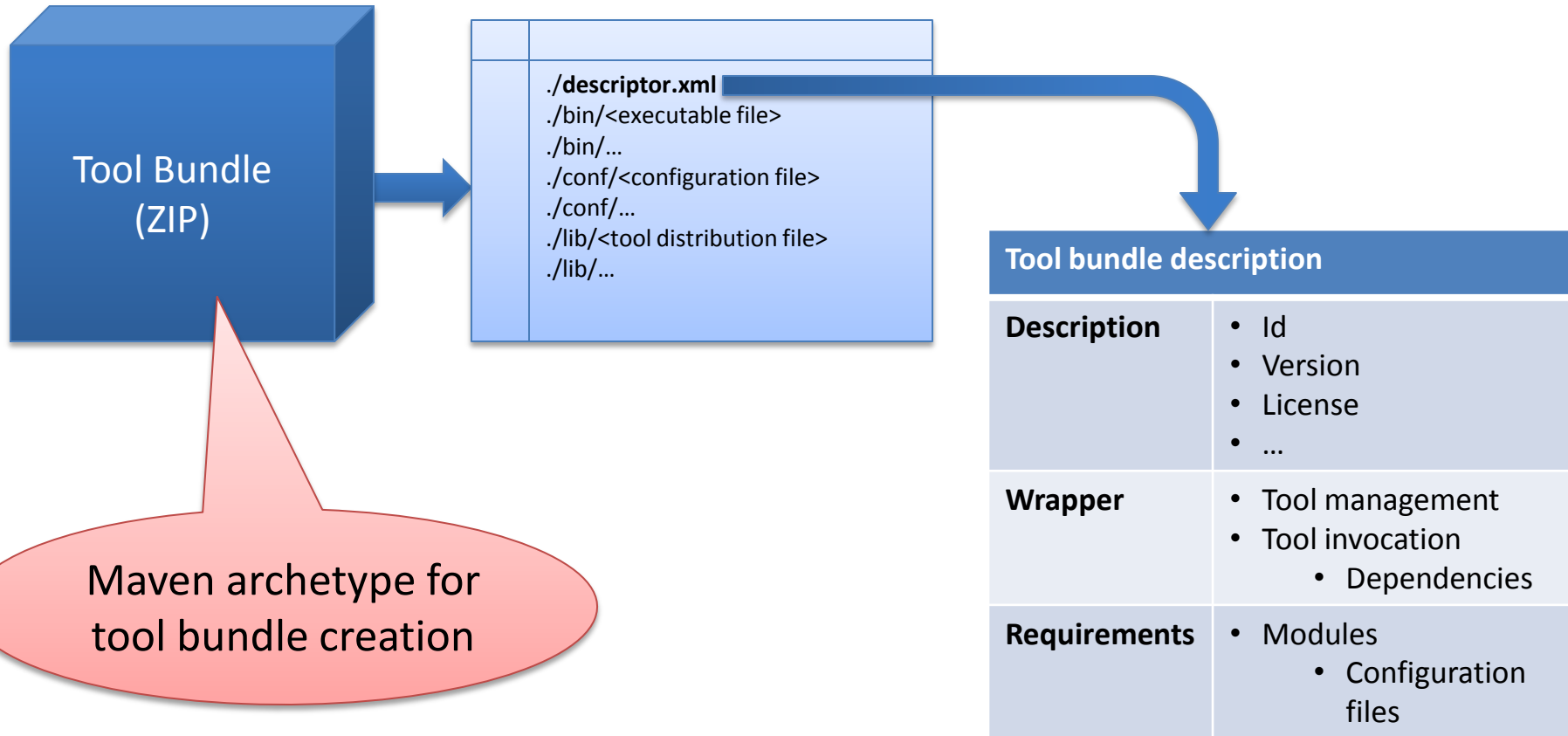
Tools evaluated

Ontology management frameworks			
Tool	**Version**		
Jena	2.6.3		
OWL API	3.1.0 1592		
Sesame	2.3.1		
Ontology editors			
Tool	**Version**	**API**	**API version**
NeOn Toolkit	2.3.2	OWL API	3.0.0 1310
Protégé 4	4.1 beta 209	OWL API	3.1.0 1602
Protégé OWL	3.4.4 build 579	Protégé OWL API	3.4.4 build 579

SEALS

Semantic Evaluation at Large Scale

Index

- Evaluation scenarios
- Test data
- Evaluation descriptions
- Tools
- **Results**
- Conclusions
- Links to resources

SEALS
Semantic Evaluation at Large Scale

Dynamic result visualization

Statistics

Index

Tool:OWL APIVersion1
Test: OWLLiteImportTestSuite V1
Date: 05.10.10 17:55:36

Category	No	%
Same ontologies	80	97
Different ontologies	2	2
Execution problem in tool	0	0
Execution problem in platform	0	0
TOTAL	82	100%

Information Added/Lost	No	%
Only information added	24	29
Only information lost	0	0
Information added and lost	5	6
No information added or lost	53	64
TOTAL	82	100%

Conformance results

Index ---> Conformance Test Suites ---> Conformance Results

Tool: OWL API Version: 1
OWLLiteImportTestSuite V 1
05.10.10 17:55:36

Show/Hide all data

[+] Annotation with literal

Test Id	Results ○	Equivalence ○	Information Added ○	Information Lost ○
OWLLiteImportTest82 ISL15.owl	Conformance=SAME Execution=OK	Semantic=true Structural=false	[+] Declaration(AnnotationProperty(rdfs:label))	

[+] Anonymous individual

Test Id	Results ○	Equivalence ○	Information Added ○
OWLLiteImportTest62 ISJ01.owl	Conformance=SAME Execution=OK	Semantic=true Structural=false	[+] Declaration(NamedIndividual(<http://www.example.org/ISJ01#John>)) ClassAssertion(<http://www.example.org/ISJ01#Person> _:http://www.example.org/ISJ
OWLLiteImportTest63 ISJ02.owl	Conformance=SAME Execution=OK	Semantic=true Structural=false	[+]
OWLLiteImportTest80 ISL13.owl	Conformance=DIFFERENT Execution=OK	Semantic=false Structural=false	[+]

SEALS
Semantic Evaluation at Large Scale

Conformance results

(a) RDF(S) conformance.

Category	JE	NT	OA	P4	PO^2	SE
SAME	82	0	0	0	68	82
DIFF	0	82	82	82	14	0
FAIL	0	0	0	0	0	0
TOTAL	82	82	82	82	82	82

(b) OWL Lite conformance.

Category	JE	NT^3	OA^3	P4^3	PO	SE
SAME	82	78	80	80	73	82
DIFF	0	2	2	2	9	0
FAIL	0	2	0	0	0	0
TOTAL	82	82	82	82	82	82

(c) OWL DL conformance.

Category	JE	NT^3	OA^3	P4^3	PO	SE
SAME	561	549	549	549	429	561
DIFF	0	8	11	11	132	0
FAIL	0	4	1	1	0	0
TOTAL	561	561	561	561	561	561

- **Jena** and **Sesame** have no problems
- The **OWL API** (and **Protégé 4**)
  - Transforms ontologies to OWL 2
    - E.g., Individual → Named Individual[3]
    - Sometimes with unexpected effects
  - Some problems
    - Anonymous individuals and object properties
    - Datatype property with range an enumerated datatype
    - Execution fails with *owl:imports*
- The **NeOn Toolkit** (old OWL API version)
  - More problems with anonymous individuals
- **Protégé OWL**
  - RDF(S): Creates and OWL ontology with a random name[2]
  - Minor issue with literals
  - Class descriptions subject or object of a *rdfs:subClassOf* property

SEALS

Semantic Evaluation at Large Scale

Interoperability results

(a) RDF(S) interoperability.

	JE	SE	PO	NT	OA	P4
JE	100	100	83	0	0	0
SE	100	100	83	0	0	0
PO	83	83	83	0	0	0
NT	0	0	0	0	0	0
OA	0	0	0	0	0	0
P4	0	0	0	0	0	0

(b) OWL Lite interoperability.

	JE	SE	OA	P4	NT	PO
JE	100	100	98	98	95	89
SE	100	100	98	98	95	89
OA	98	98	98	98	95	89
P4	98	98	98	98	95	89
NT	95	95	95	95	95	87
PO	89	89	89	89	87	89

(c) OWL DL interoperability.

	JE	SE	OA	P4	NT	PO
JE	100	100	98	98	98	76
SE	100	100	98	98	98	76
OA	98	98	98	98	98	75
P4	98	98	98	98	98	75
NT	98	98	98	98	98	75
PO	76	76	75	75	75	76

- Same conclusions as in conformance
- New fact:
  - From the **OWL API** (or tools that use it) to **Protégé OWL**: Protégé OWL has execution problems with anonymous individuals related through a datatype properties with literal values
- **Conclusions**:
  - While Jena and Sesame have no interoperability problems, the other tools have some issues
  - Tools based on the OWL API cannot interoperate using RDF(S) (they convert ontologies into OWL 2)

SEALS

Semantic Evaluation at Large Scale

Scalability results

Table 4.1: Scalability results with real-world ontologies.

ScalTest	OntSize	Jena	NeOn	OWL API	Protégé 4	ProtégéOWL	Sesame
ScalTest001	82KB	1 sec	2 sec	3 sec	3 sec	5 sec	1 sec
ScalTest002	758KB	2 sec	3 sec	3 sec	2 sec	4 sec	2 sec
ScalTest003	119KB	less 1	1 sec	less 1	less 1	1 sec	less 1
ScalTest004	33KB	less 1	less 1	1 sec	less 1	less 1	less 1
ScalTest005	274KB	less 1	1 sec	1 sec	1 sec	1 sec	1 sec
ScalTest006	37.7MB	-	-	-	-	-	-
ScalTest007	29KB	less 1	8 sec	6 sec	6 sec	19 sec	less 1
ScalTest008	233KB	less 1	10 sec	11 sec	9 sec	14 sec	1 sec
ScalTest009	164KB	less 1	-	-	-	81 sec	less 1
ScalTest010	4KB	less 1	-	-	-	27 sec	less 1
ScalTest011	1.3MB	1 sec	3 sec	2 sec	1 sec	4 sec	2 sec
ScalTest012	430KB	1 sec	2 sec	less 1	1 sec	1 sec	1 sec
ScalTest013	476KB	1 sec	2 sec	1 sec	less 1	1 sec	less 1
ScalTest014	4.1MB	4 sec	-	-	-	38 sec	2 sec
ScalTest015	881KB	less 1	3 sec	3 sec	3 sec	8 sec	less 1
ScalTest016	1.3MB	2 sec	-	-	-	99 sec	1 sec
ScalTest017	14.6MB	17 sec	-	-	-	30 sec	6 sec
ScalTest018	12.6MB	12 sec	-	-	-	23 sec	5 sec
ScalTest019	13.9MB	14 sec	-	-	-	27 sec	5 sec
ScalTest020	13.8MB	12 sec	-	-	-	26 sec	6 sec

Table 4.2: Scalability results with LUBM data generator ontologies

ScalTest	OntSize	Jena	NeOn	OWLAPI	Protégé4	ProtégéOWL	Sesame
ScalTest001	8212KB	10 sec	-	-	-	23 sec	4 sec
ScalTest002	18907KB	16 sec	-	-	-	56 sec	5 sec
ScalTest003	27745KB	22 sec	-	-	-	105 sec	6 sec
ScalTest004	39380KB	32 sec	-	-	-	failed	9 sec

- Dependence between ontology size and execution time
- Results vary between the different tools

SEALS

Semantic Evaluation at Large Scale

Index

- Evaluation scenarios
- Test data
- Evaluation descriptions
- Tools
- Results
- **Conclusions**

SEALS

Semantic Evaluation at Large Scale

Conclusions

- Knowledge representation language
  - RDF-based tools have no problems
  - OWL / OWL 2 tools deal better with OWL ontologies
- Development decisions
  - Results highly influenced by development decisions
    - E.g., OWL API developers' decision of converting all ontologies into OWL 2
- Dependency between OWL Lite and OWL DL
  - The OWL DL evaluation unveiled more problems (not only related OWL DL but also to OWL Lite)
- Use of ontology APIs
  - Does not isolate a tool from having problems
  - May help increasing conformance and interoperability
    - Be aware of defects and regularly update!

SEALS

Semantic Evaluation at Large Scale

Evaluation Campaign 2011

- Conformance & interoperability
  - RDF(S)
  - OWL Lite
  - OWL DL
  - OWL Full
  - **OWL 2 Full**
  - **Ontology design pattern test data**
- Scalability
  - **Real-world**
  - **LUBM**

SEALS
Semantic Evaluation at Large Scale

Links to resources

- **Evaluation campaign:**
  - http://www.seals-project.eu/seals-evaluation-campaigns/2nd-evaluation-campaigns/ontology-engineering-tools
- **Test suites:**
  - RDF(S):
    - http://seals.sti2.at/tdrs-web/testdata/persistent/RDF(S)+Import+Test+Suite/1.0/
  - OWL Lite:
    - http://seals.sti2.at/tdrs-web/testdata/persistent/OWL+Lite+Import+Test+Suite/1.0/
  - OWL DL:
    - http://seals.sti2.at/tdrs-web/testdata/persistent/OWL+DL+Import+Test+Suite/1.3/
  - OWL Full:
    - http://seals.sti2.at/tdrs-web/testdata/persistent/OWL+Full+Import+Test+Suite/1.0/
- **Result visualization:**
  - http://seals-ui.sti2.org/

SEALS

Semantic Evaluation at Large Scale

Join the SEALS Evaluation Campaigns!