Semantic Web Reasoning

Majid Sazvar
sazvar@stu-mail.um.ac.ir

Knowledge Engineering Research Group
Ferdowsi University of Mashhad

Overview

- Introduction
- What is the meaning?
- Where is meaning in the Semantic Web?
- Structured Web vs. Reasonable Web
- SWR Challenges and Solutions
- Conclusion

The WWW is penetrating our society

- Social contacts (social networking platforms, blogging, ...)
- Economics (buying, selling, advertising, ...)
- Administration (eGovernment)
- Education (eLearning, Web as information system, ...)
- Work life (information gathering and sharing)
- Recreation (games, role play, ...)

The current Web

Immensely successful.

- Huge amounts of data.
- Syntax standards for transfer of unstructured data.
- Machine-processable, human-readable documents.

BUT:

- Content/knowledge cannot be accessed by machines.
- Meaning (semantics) of transferred data is not accessible.

Limitations of the current Web

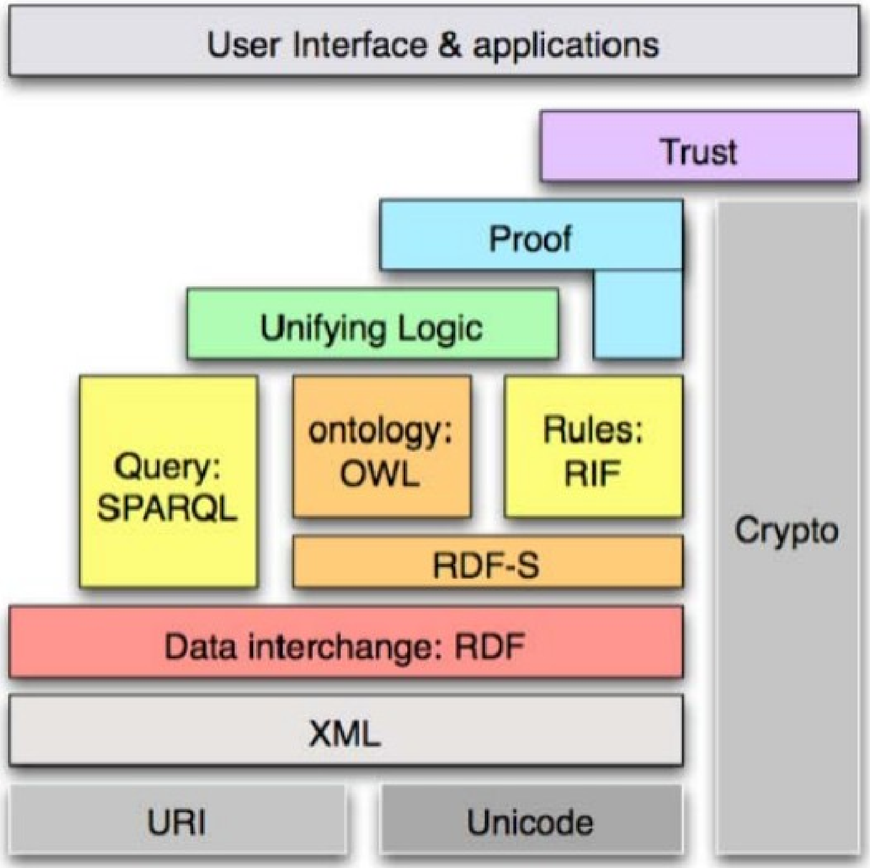
- Too much information with too little structure and made for human consumption
  - Content search is very simplistic
  - future requires better methods
- Web content is heterogeneous
  - in terms of content
  - in terms of structure
  - in terms of character encoding
    - future requires intelligent information integration
- Humans can derive new (implicit) information from given pieces of information

  but on the current Web we can only deal with syntax
  - requires automated reasoning techniques

Basic ingredients for the Semantic Web

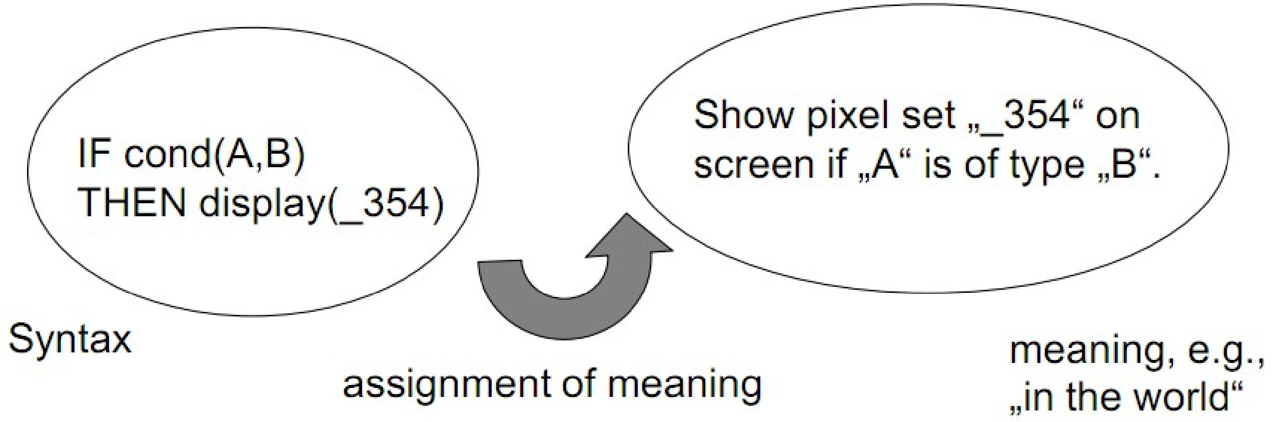
- Open Standards for describing information on the Web
- Methods for obtaining further information from such descriptions

The Semantic Web Layer Cake

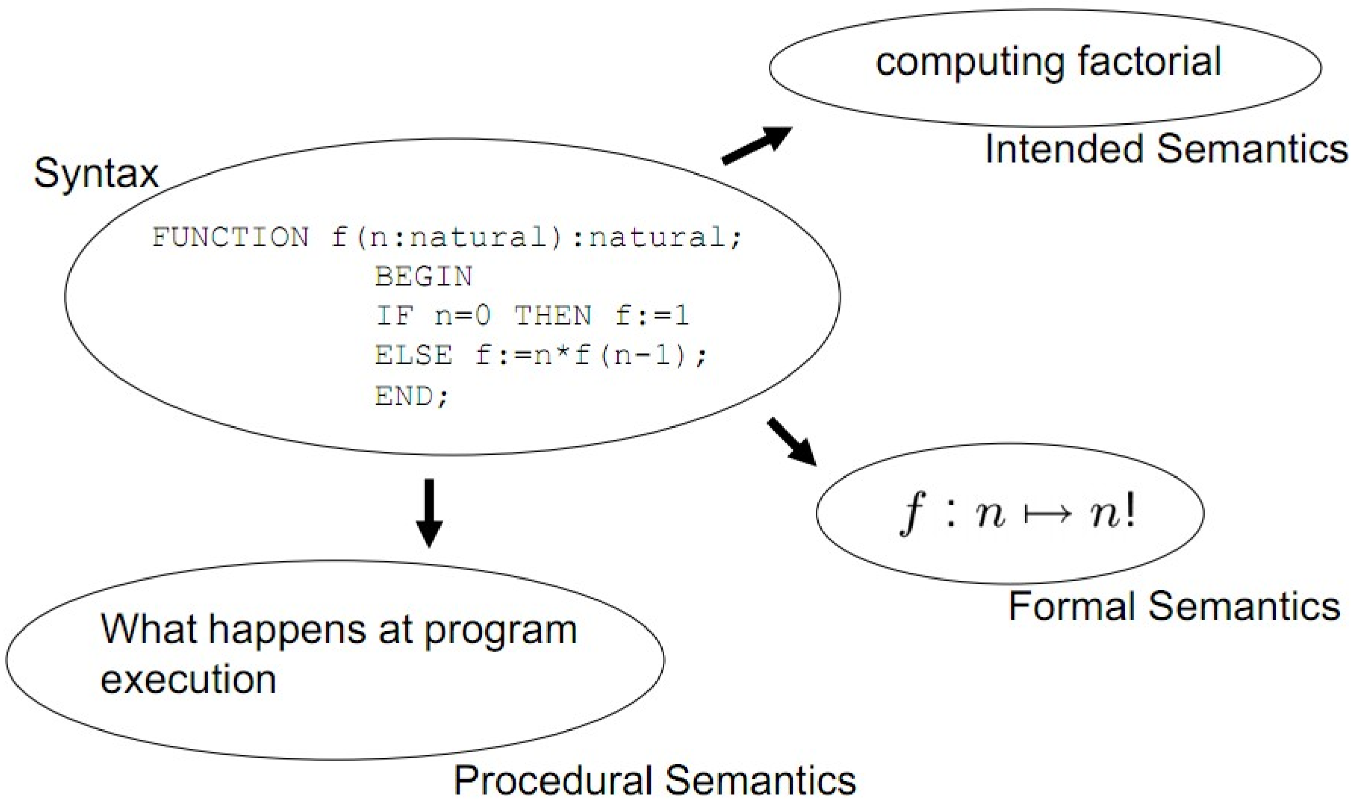


Syntax and Semantics

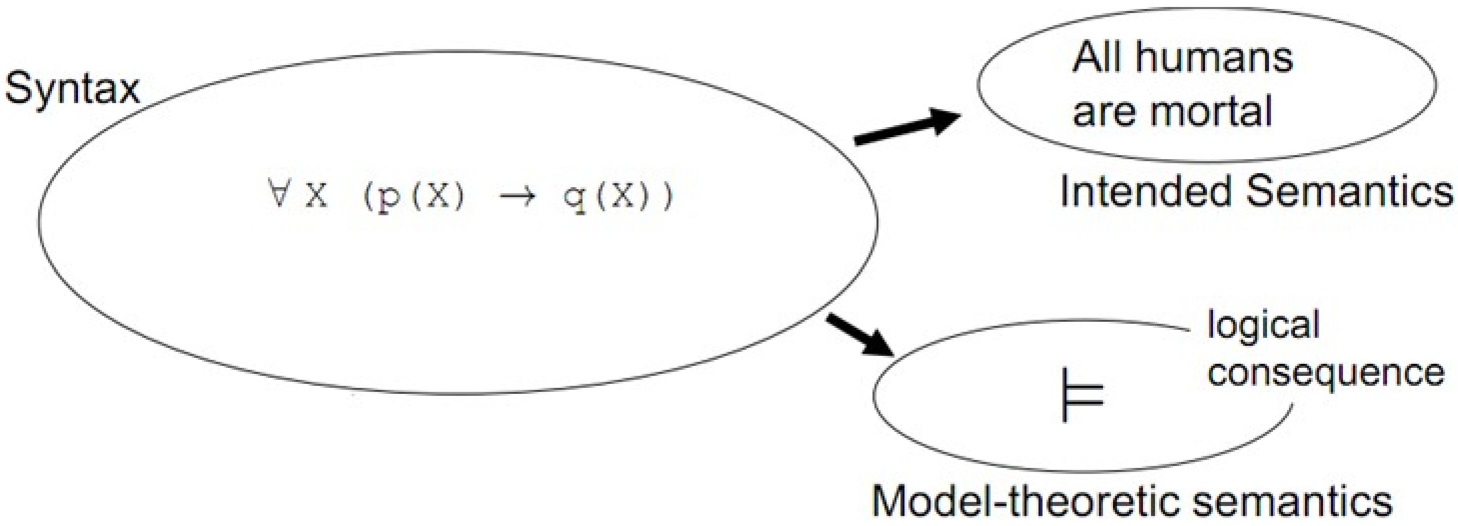
- Syntax: character strings without meaning
- Semantics: meaning of the character strings



Semantics of Programming Languages



Semantics of Logic



What Semantics Is Good For

- Semantic Web requires a shareable, declarative and computable semantics.
- I.e., the semantics must be a formal entity which is clearly defined and automatically computable.
- Ontology languages provide this by means of their formal semantics.
- Semantic Web Semantics is given by a relation – the logical consequence relation.

In other words

- We capture the meaning of information

  not by specifying its meaning (which is impossible)

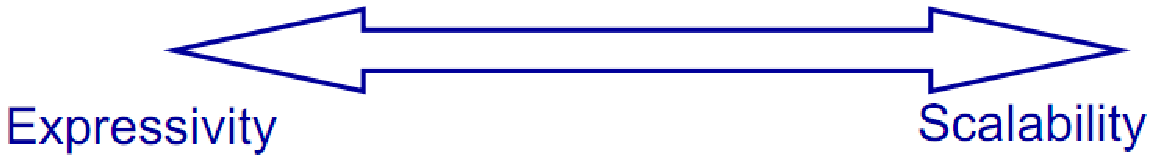
  but by specifying

  how information interacts with other information.

- We describe the meaning indirectly through its effects.

Ontology languages

- Of central importance for the realization of Semantic
- Meaning (semantics) provided via logic and deduction algorithms (automated reasoning).
- Scalability is a challenge.



Explicit vs. Implicit Knowledge

- if an RDFS document contains

```
u    rdf:type    ex:Textbook .
```

and

```
ex:Textbook    rdfs:subClassOf    ex:Book .
```

then

```
u    rdf:type    ex:Book .
```

- is *implicitly* also the case: it's a logical consequence. (We can also say it is deduced (deduction) or inferred (inference).
- We do not have to state this explicitly.
- Which statements are logical consequences is governed by the formal semantics (covered in the next session).

RDFS Semantic

Number	If...	then...
1	$s\ p\ o$ (where o is a literal)	_ : n rdf:type rdfs:Literal
2	s rdfs:domain x $u\ s\ y$	u rdf:type x
3	p rdfds:range o $s\ p\ v$	v rdf:type o
4a	$s\ p\ o$	s rdf:type rdfs:Resource
4b	$s\ p\ o$	o rdf:type rdfs:Resource
5	p rdfs:subPropertyOf $p1$ $p1$ rdfs:subPropertyOf $p2$	p rdfs:subPropertyOf $p2$
6	p rdf:type rdf:Property	p rdfs:subPropertyOf p
7	$s\ p\ o$ p rdfs:subPropertyOf $p1$	$s\ p1\ o$
8	s rdf:type rdfs:subClassOf	s rdfs:subClassOf rdfs:Resource
9	c rdfs:subClassOf $c1$ v rdf:type c	v rdf:type $c1$
10	u rdf:type rdfs:Class	u rdfs:subClassOf u
11	c rdfs:subClassOf $c1$ $c1$ rdfs:subClassOf $c2$	c rdfs:subClassof $c2$
12	s rdf:type rdfs:ContainerMembershipProperty	s rdfs:subPropertyOf rdfs:member
13	s rdf:type rdfs:Datatype	s rdfs:subClassOf rdfs:Literal

OWL-Horst Semantic

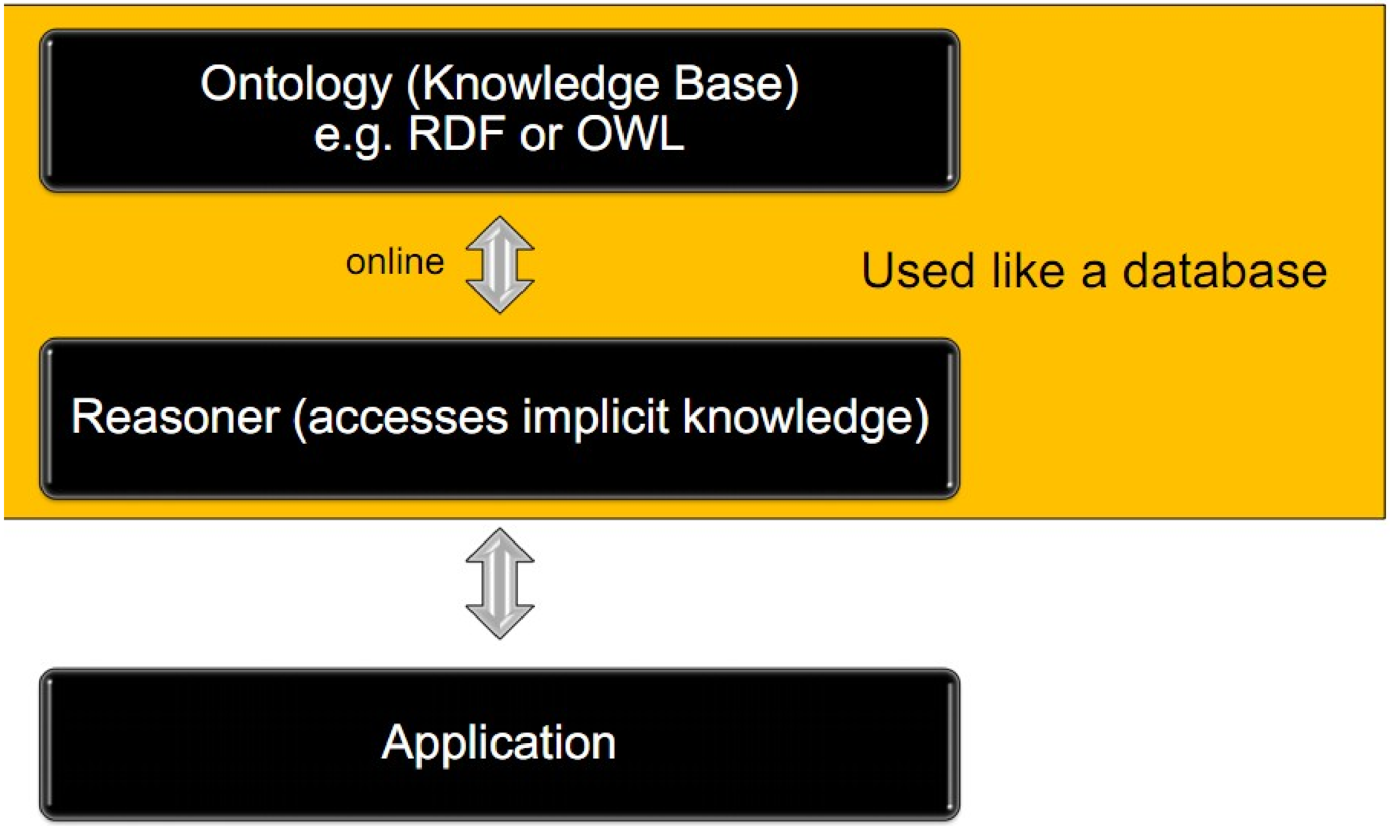
Number	If...	then...
1	p rdf:type owl:FunctionalProperty $u\ p\ v$ $u\ p\ w$	v owl:sameAs w
2	p rdf:type owl:InverseFunctionalProperty $v\ p\ u$ $w\ p\ u$	v owl:sameAs w
3	p rdf:type owl:SymmetricProperty $v\ p\ u$	$u\ p\ v$
4	p rdf:type owl:TransitiveProperty $u\ p\ w$ $w\ p\ v$	$u\ p\ v$
5a	$u\ p\ v$	u owl:sameAs u
5b	$u\ p\ v$	v owl:sameAs v
6	v owl:sameAs w	w owl:sameAs v
7	v owl:sameAs w w owl:sameAs u	v owl:sameAs u
8a	p owl:inverseOf q $v\ p\ w$	$w\ q\ v$
8b	p owl:inverseOf q $v\ q\ w$	$w\ p\ v$
9	v rdf:type owl:Class v owl:sameAs w	v rdfs:subClassOf w
10	p rdf:type owl:Property p owl:sameAs q	p rdfs:subPropertyOf q

OWL-Horst Semantic

11	$u\ p\ v$ u owl:sameAs x v owl:sameAs y	$x\ p\ y$
12a	v owl:equivalentClass w	v rdfs:subClassOf w
12b	v owl:equivalentClass w	w rdfs:subClassOf v
12c	v rdfs:subClassOf w w rdfs:subClassOf v	v rdfs:equivalentClass w
13a	v owl:equivalentProperty w	v rdfs:subPropertyOf w
13b	v owl:equivalentProperty w	w rdfs:subPropertyOf v
13c	v rdfs:subPropertyOf w w rdfs:subPropertyOf v	v rdfs:equivalentProperty w
14a	v owl:hasValue w v owl:onProperty p $u\ p\ v$	u rdf:type v
14b	v owl:hasValue w v owl:onProperty p u rdf:type v	$u\ p\ v$
15	v owl:someValuesFrom w v owl:onProperty p $u\ p\ x$ x rdf:type w	u rdf:type v
16	v owl:allValuesFrom w v owl:onProperty p u rdf:type v $u\ p\ x$	x rdf:type w

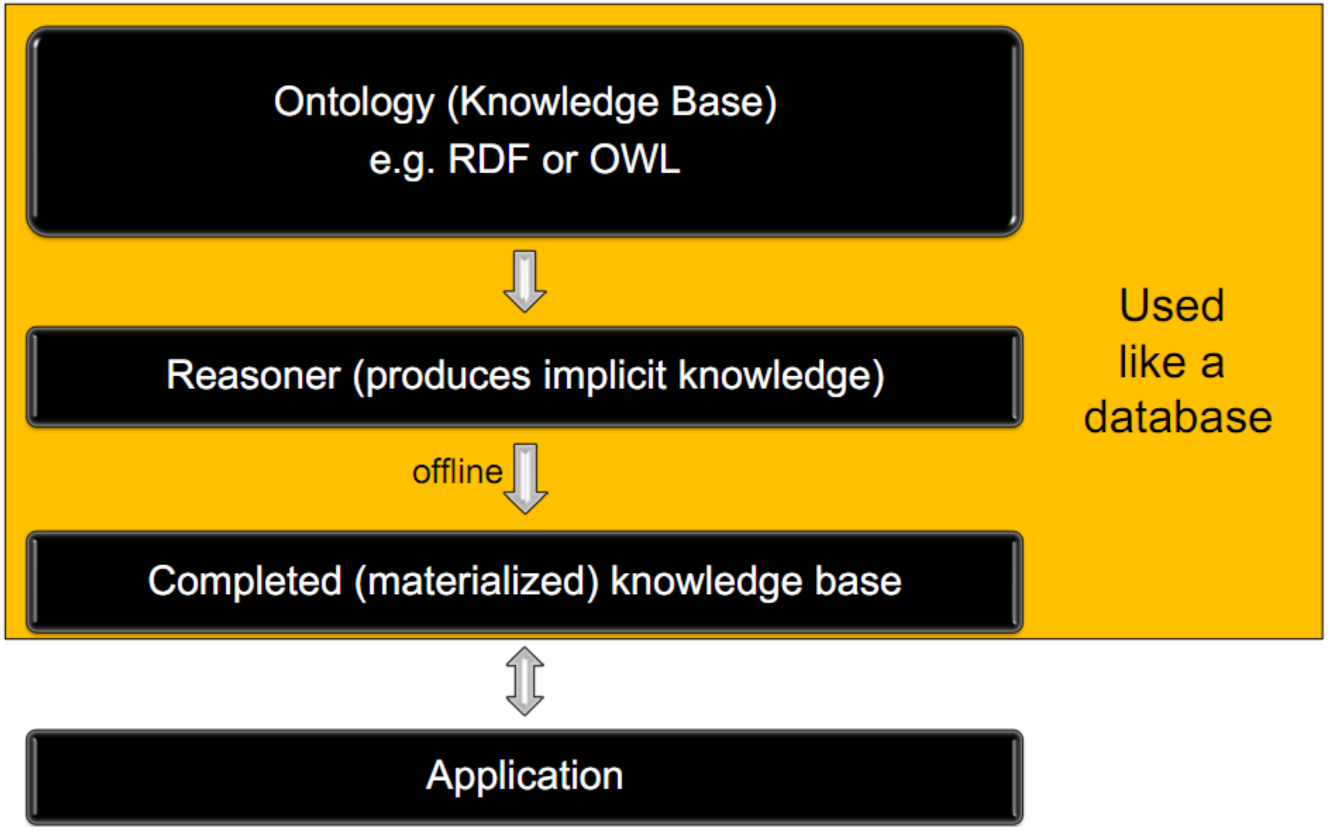
Using implicit knowledge

Backward Chaining Method (Goal-Driven)



Using implicit knowledge

Forward Chaining Method (Data-Driven)



Importance of meaning in the Semantic Web

Structured Web

+

Semantic

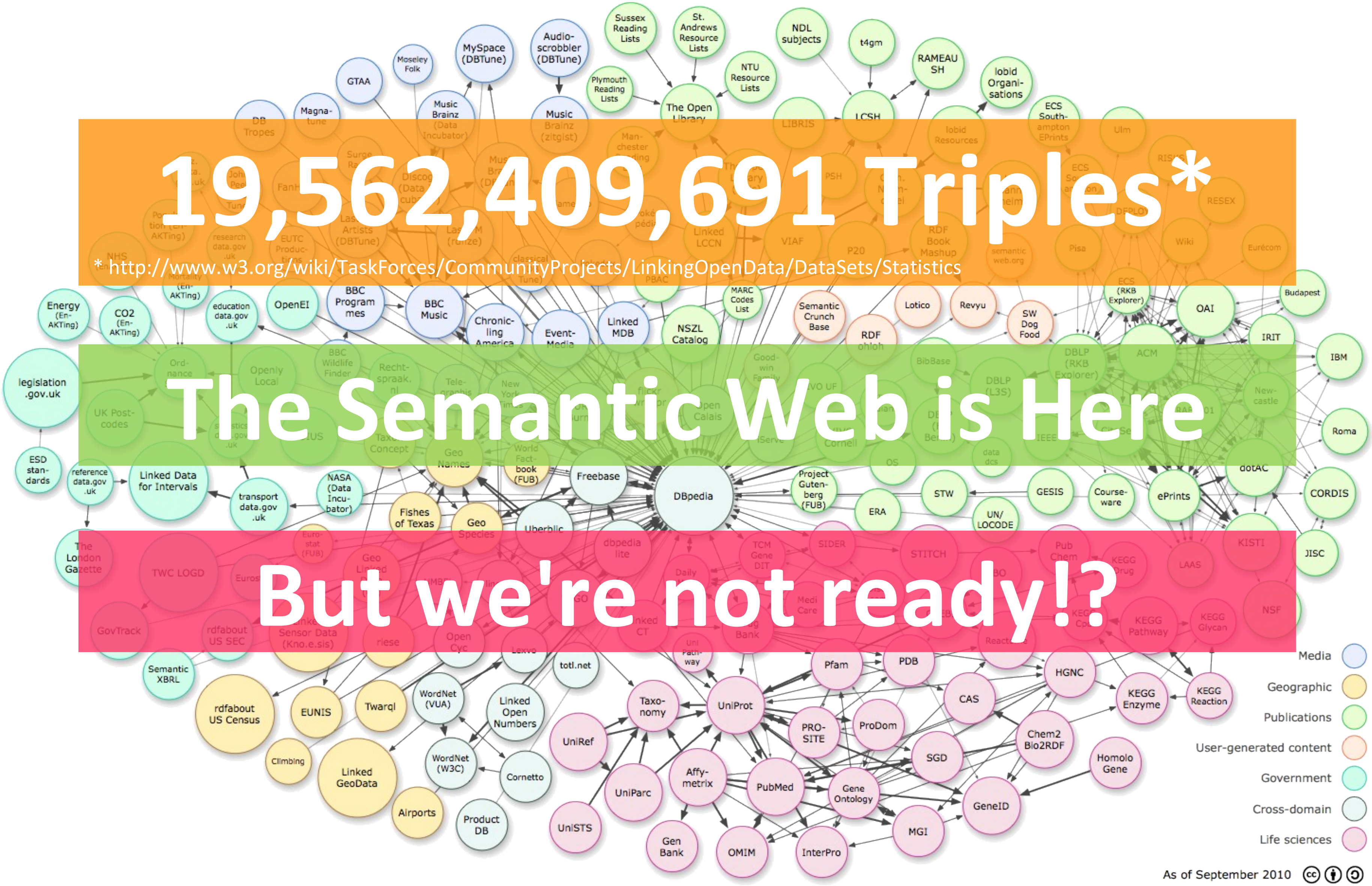
=

Reasonable Web

=

Semantic Web

LOD Cloud



19,562,409,691 Triples*

* http://www.w3.org/wiki/TaskForces/CommunityProjects/LinkingOpenData/DataSets/Statistics

The Semantic Web is Here

But we're not ready!?

Vastness

- The World Wide Web contains at least 30 billion pages.
- Any automated reasoning system will have to deal with truly huge inputs.
- Distributed Reasoning is a promising technique.

Vagueness

- These are imprecise concepts like "young" or "tall".
- This arises from the vagueness of user queries, of concepts represented by content providers, of matching query terms to provider terms and of trying to combine different knowledge bases with overlapping but subtly different concepts.
- Fuzzy logic is the most common technique for dealing with vagueness.

Uncertainty

- These are precise concepts with uncertain values. For example, a patient might present a set of symptoms which correspond to a number of different distinct diagnoses each with a different probability.

- Probabilistic reasoning techniques are generally employed to address uncertainty.

Inconsistency

- These are logical contradictions which will inevitably arise during the development of large ontologies, and when ontologies from separate sources are combined.

- Deductive reasoning fails catastrophically when faced with inconsistency, because "anything follows from a contradiction".

- Inductive reasoning, defeasible reasoning and paraconsistent reasoning are techniques which can be employed to deal with inconsistency.

Privacy

- Q: Can a reasoner answer queries using hidden knowledge without exposing hidden knowledge?

- A: Privacy-Preserving Reasoning

Deceit

- This is when the producer of the information is intentionally misleading the consumer of the information.

- Trust and Cryptography techniques are currently utilized to alleviate this threat.

Conclusion

We need to move towards the Reasonable Web.

?