Semantic Technology

Leo Obrst, MITRE

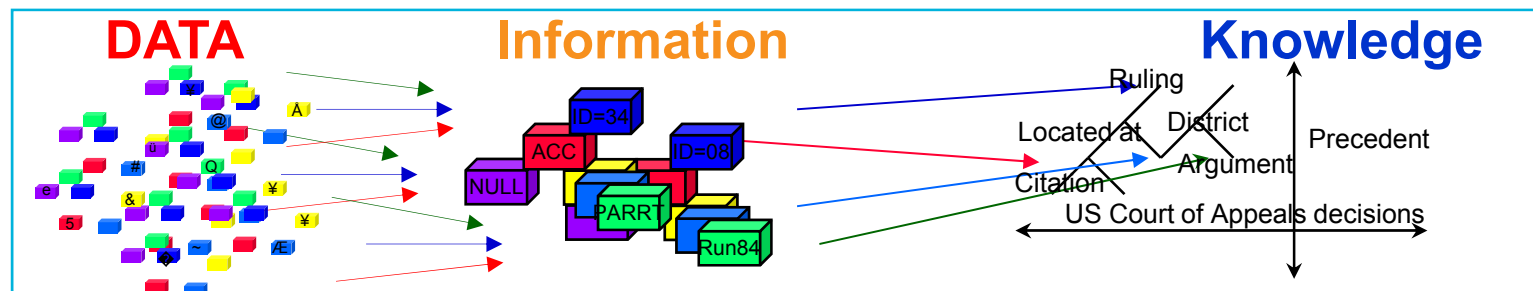**December 12, 2012**

MITRE

Towards Solving the Problem

- **With the increasing complexity of our systems and our IT needs, we need to go to human level interaction**

- **We need to maximize the amount of Semantics we can utilize**

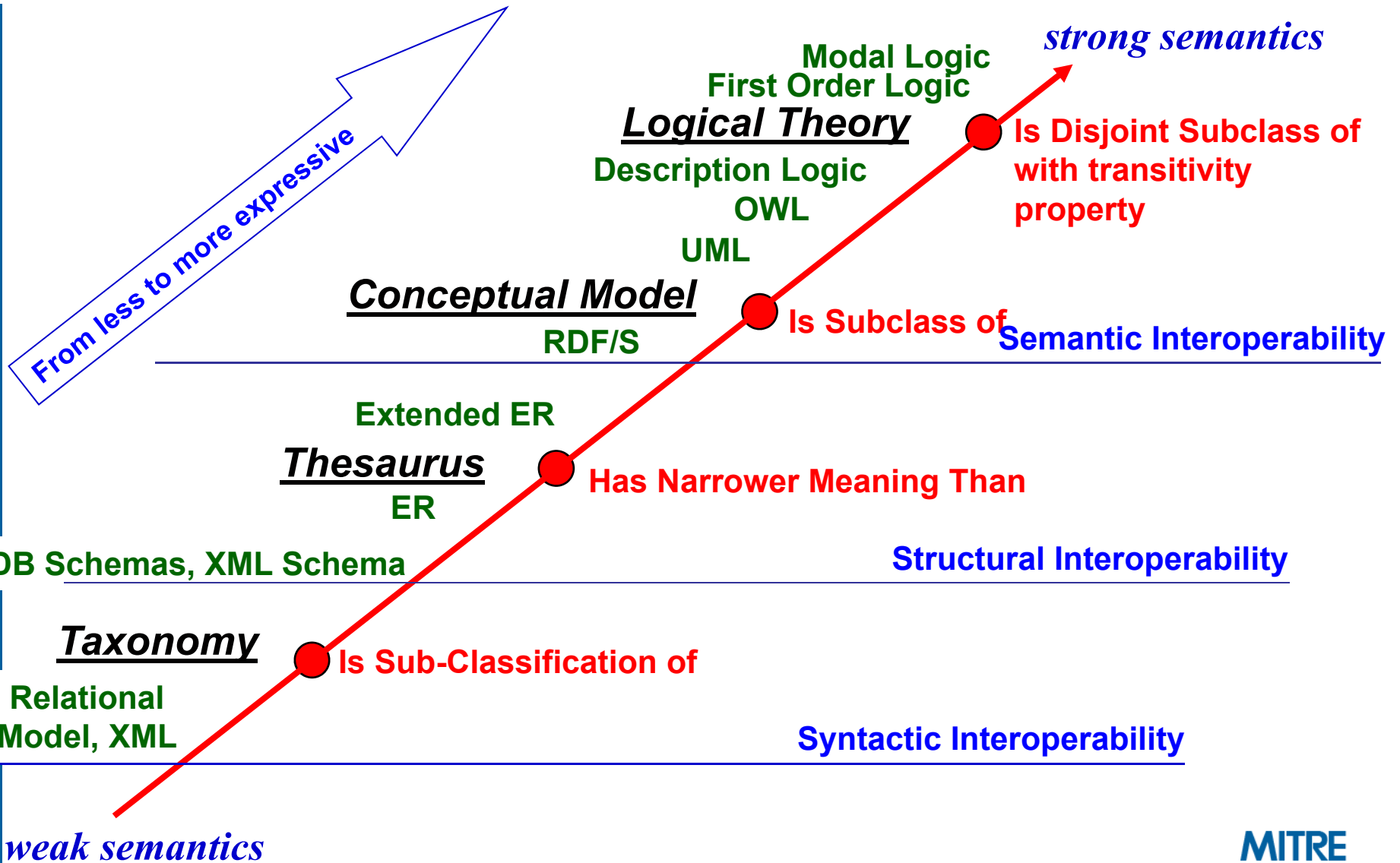- **From data and information level, we need to go to human *semantic* level interaction**

- **Semantic integration: Semantic technology is the key!**

MITRE

Premise & Promise

- **Semantic technologies can enrich search, discovery, navigation**
  - Applies to: data integration, social networking, intelligence analytics, situational awareness, etc.
  - Many emerging tools, existing and emerging semantic models, patterns, etc.

- **Application of semantic technologies to search, discovery, & navigation, by adding semantics to raw data, will provide:**
  - Finer grained search and navigation
  - Richer results in discovery, finding people and communities
  - Traversal of complex social and information relationships
  - Entity disambiguation and correlation; identity management

- **Humans have complex semantics, in both language, mental models**
  - We can't duplicate that fidelity, granularity, complexity, intent, context of usage

- **But semantic search can discover, develop, deploy, disseminate**
  - More complex patterns of relationships
  - Semantic engines can better use those patterns

MITRE

Ontology Spectrum: The Range of Semantic Models

strong semantics

Modal Logic
First Order Logic

Logical Theory

Description Logic

OWL

UML

● Is Disjoint Subclass of with transitivity property

Conceptual Model

RDF/S

● Is Subclass of   Semantic Interoperability

From less to more expressive

Extended ER

Thesaurus

ER

● Has Narrower Meaning Than

DB Schemas, XML Schema   Structural Interoperability

Taxonomy

● Is Sub-Classification of

Relational Model, XML

Syntactic Interoperability

weak semantics

MITRE

Ontology & Ontologies

- **An ontology defines the terms used to describe and represent an area of knowledge (subject matter)**
  - An ontology also is the model (set of concepts) for the meaning of those terms
  - An ontology thus defines the vocabulary and the meaning of that vocabulary
- **Ontologies are used by people, databases, and applications that need to share domain information**
  - Domain: a specific subject area or area of knowledge, like medicine, tool manufacturing, real estate, automobile repair, financial management, etc.
- **Ontologies include computer-usable definitions of basic concepts in the domain and the relationships among them**
  - They encode domain knowledge (modular)
  - Knowledge that spans domains (composable)
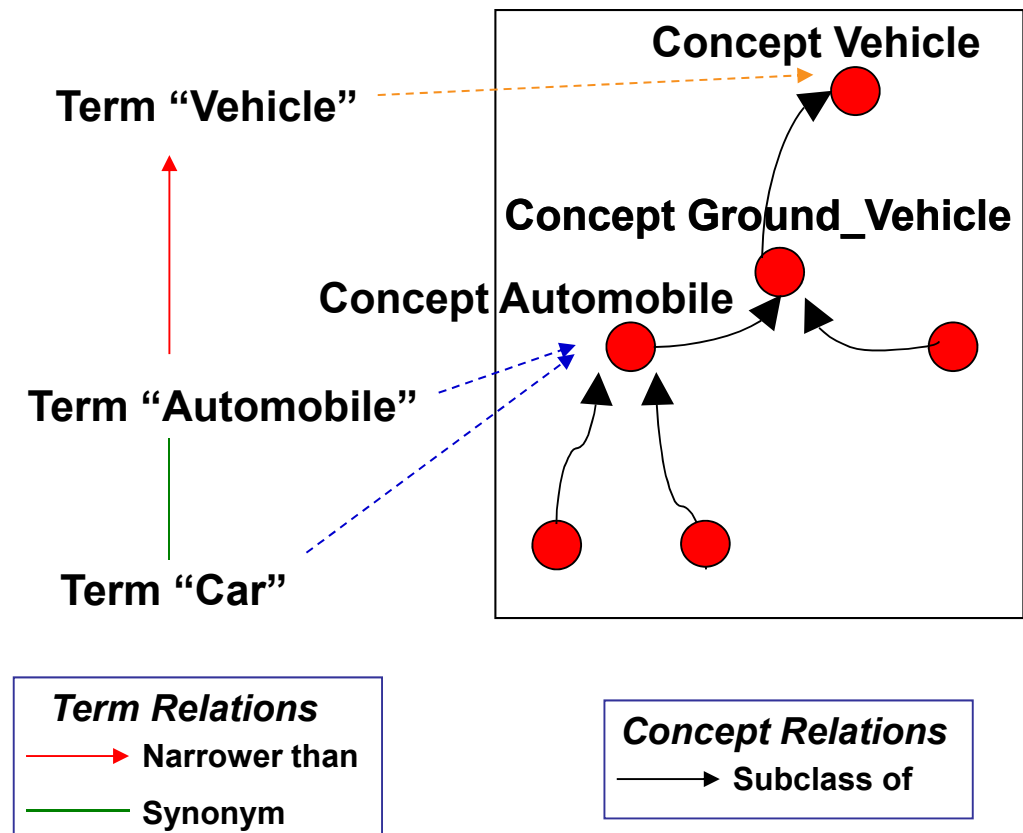 - Make knowledge available (reusable)

MITRE
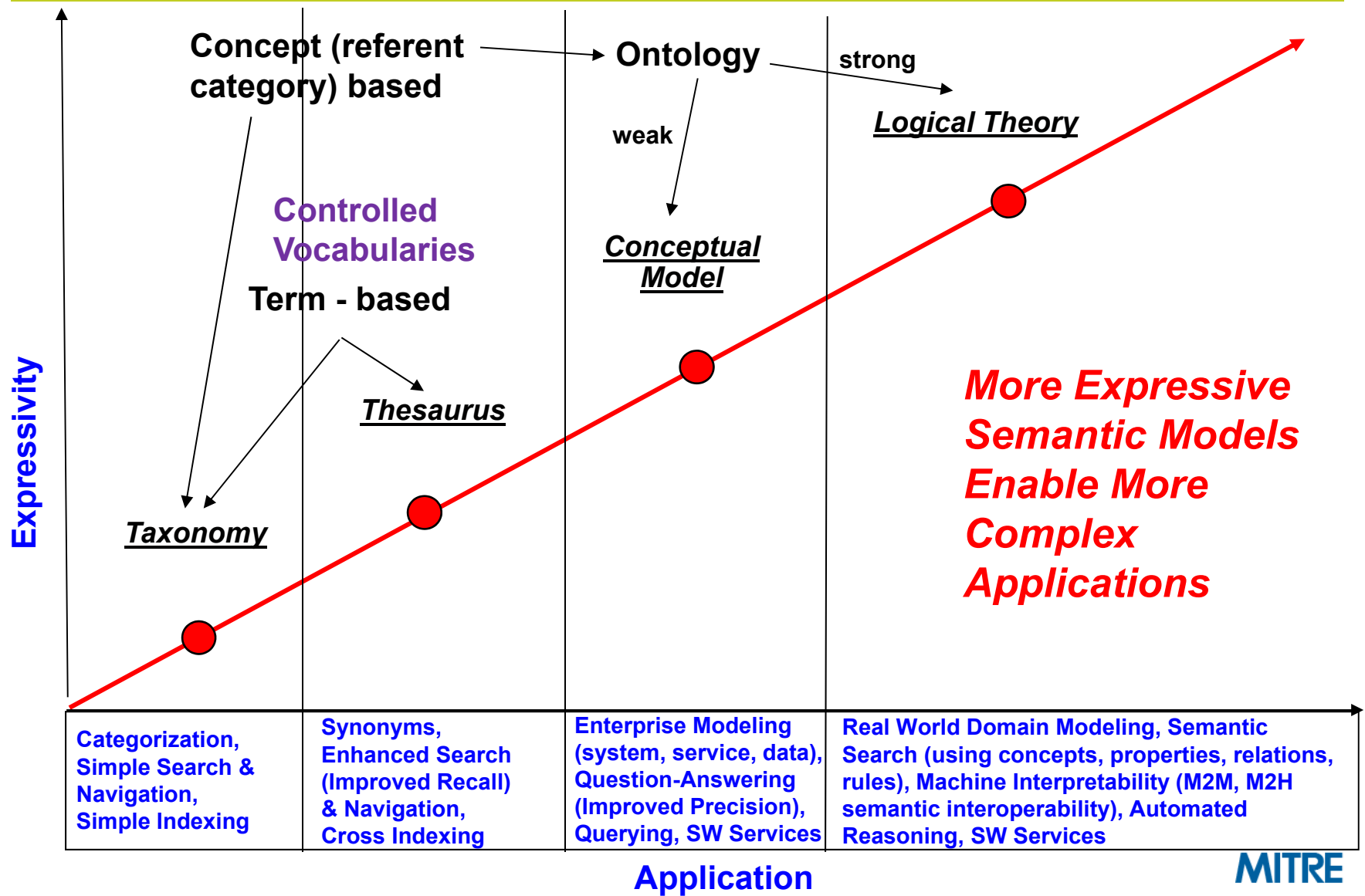
Term vs. Concept

- **Term (terminology):**
  - Natural language words or phrases that act as indices to the underlying meaning, i.e., the concept (or composition of concepts)
  - The syntax (e.g., string) that stands in for or is used to indicate the semantics (meaning)
- **Concept (ontology):**
  - A unit of semantics (meaning), the node (entity) or link (relation) in the mental or knowledge representation model
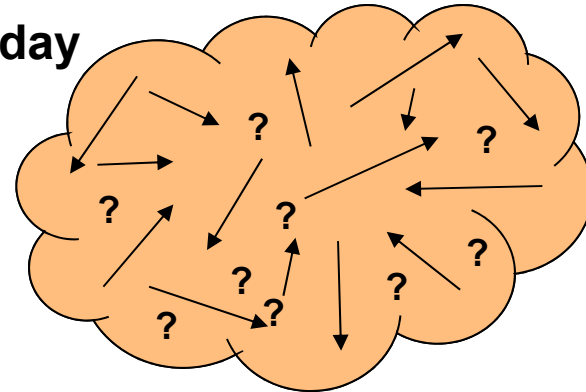
Term "Vehicle"

Term "Automobile"

Term "Car"

Concept Vehicle

Concept Ground_Vehicle

Concept Automobile

Term Relations
→ Narrower than
— Synonym

Concept Relations
→ Subclass of

MITRE

Ontology Spectrum: Application

Expressivity (y-axis)

Application (x-axis)

Concept (referent category) based → Ontology → strong → *Logical Theory*

Controlled Vocabularies

weak → *Conceptual Model*

Term - based

Thesaurus

Taxonomy

More Expressive Semantic Models Enable More Complex Applications

| Categorization, Simple Search & Navigation, Simple Indexing | Synonyms, Enhanced Search (Improved Recall) & Navigation, Cross Indexing | Enterprise Modeling (system, service, data), Question-Answering (Improved Precision), Querying, SW Services | Real World Domain Modeling, Semantic Search (using concepts, properties, relations, rules), Machine Interpretability (M2M, M2H semantic interoperability), Automated Reasoning, SW Services |

MITRE

The Semantic Web

- Current Web is a collection of links and resources
  - Is syntactic & structural only
  - Excludes semantic interoperability at high levels.
  - Google of today is string based (keyword) & has no notion of the semantics (meaning) of your query
- Semantic Web extends the Current Web so information is given well-defined meaning
  - Enables semantic interoperability at high levels
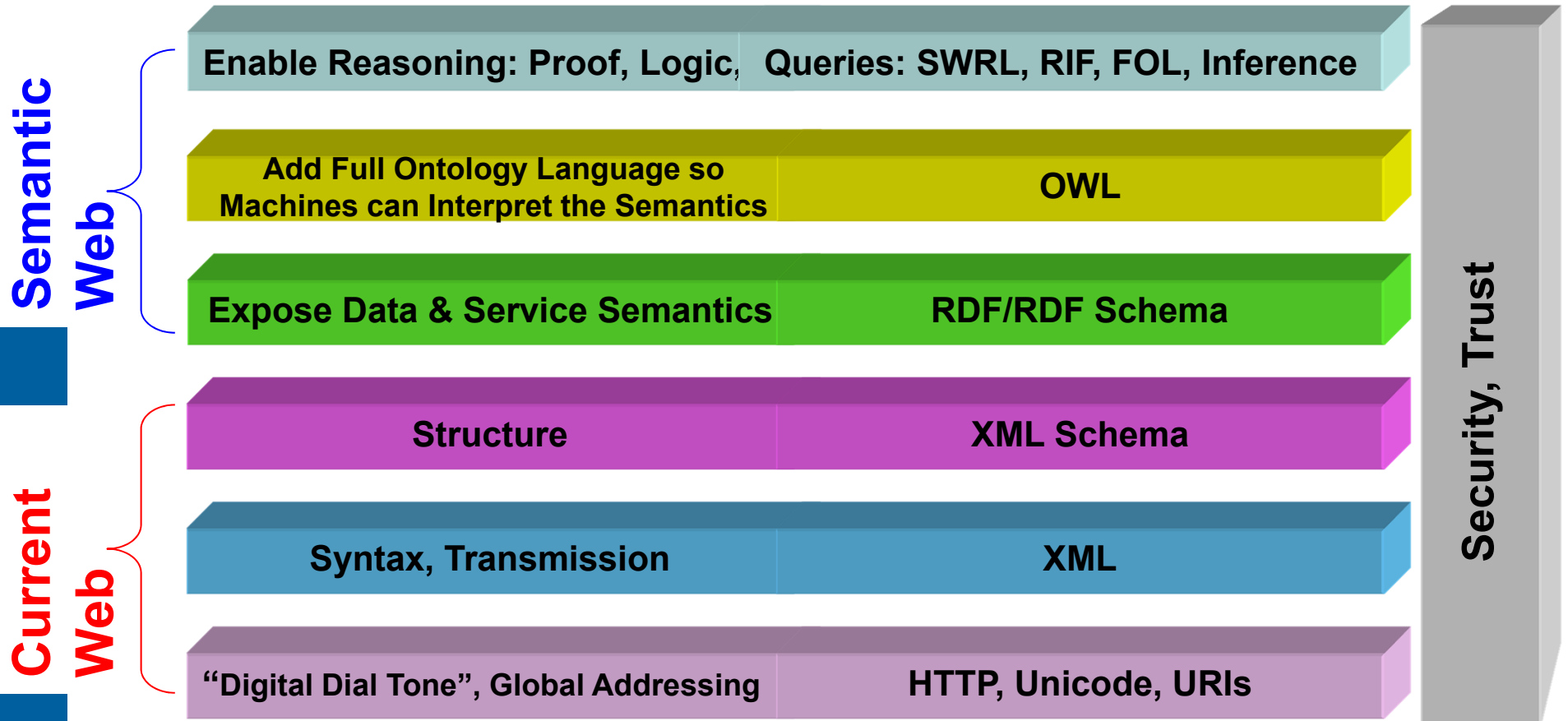 - Google Knowledge Graph is more semantic
  - Able to evaluate knowledge in context

Web today

Humans have to do the understanding

Semantic Web tomorrow

In Transit

Transitioning to

Force Structure As Is

Located at

Home base

Capabilitiies

At full strength

Theater

Deployed Force

Locations

Terrain

Surrounded by

Logistics Units

Marsh
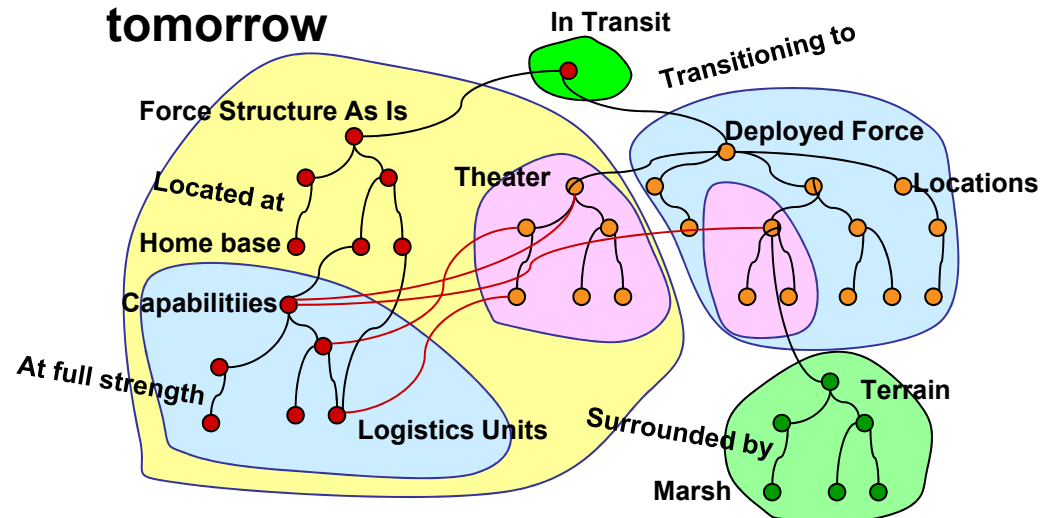
Machines partially interpret what humans mean

Semantic Web: Another View

Semantic Web

Current Web

		Security, Trust
Enable Reasoning: Proof, Logic,	Queries: SWRL, RIF, FOL, Inference	
Add Full Ontology Language so Machines can Interpret the Semantics	OWL	
Expose Data & Service Semantics	RDF/RDF Schema	
Structure	XML Schema	
Syntax, Transmission	XML	
"Digital Dial Tone", Global Addressing	HTTP, Unicode, URIs	

- Anyone, anywhere can add to an evolving, decentralized "global database"
- Explicit semantics enable looser coupling, flexible composition of services and data

MITRE

9

OWL 2 (1)

- **OWL 2 is a W3C Recommendation (27 Oct 2009)***
- **Compatible with OWL 1 (04 Feb 2004)**
- **New features**
  - Increased datatype coverage: Designed to take advantage of the new datatypes and clearer explanations available in XSD 1.1 (a recommendation 05 Apr 2012)
  - Syntactic Sugar for more easily saying things in OWL:
  - New constructs that increase expressivity
  - Simple meta-modeling capabilities
  - Extended annotation capabilities
  - Profiles

* http://www.w3.org/TR/owl2-new-features//

MITRE

OWL 2 (2)

- **Syntactic Sugar for more easily saying things in OWL:**
  - DisjointUnion:
    - DisjointUnion(:CarDoor :FrontDoor :RearDoor :TrunkDoor) : A :CarDoor is exclusively either a :FrontDoor, a :RearDoor or a:TrunkDoor and not more than one of them.
  - DisjointClasses
    - DisjointClasses(:LeftLung :RightLung) : Nothing can be both a :LeftLung and a :RightLung.
  - NegativeObject(Data)PropertyAssertion
    - NegativeObjectPropertyAssertion(:livesIn :ThisPatient :IleDeFrance) :ThisPatient does not live in the :IleDeFrance region.
  - Self-restriction on Properties: "local reflexivity"
    - SubClassOf(:AutoRegulatingProcess ObjectHasSelf(:regulate)): Auto-regulating processes regulate themselves.
  - Property Qualified Cardinality Restrictions: counted cardinality restrictions (Min, Max, Exact)
    - ObjectMaxCardinality(3 :boundTo :Hydrogen): Class of objects bound to at most three different :Hydrogen
  - Many others

MITRE

OWL 2 (3)

- **Simple meta-modeling capabilities:**
  - Punning: allows different uses of the same term and an individual
  - OWL 2 DL still imposes certain restrictions: it requires that a name cannot be used for both a class and a datatype and that a name can only be used for one kind of property; semantically names are distinct for reasoners

- **Annotations:**
  - AnnotationAssertion: for annotation of ontology entities
  - Annotation: for annotations of axioms and ontologies
  - Etc.

- **New constructs that increase expressivity**
  - Declarations: a declaration signals that an entity is part of the vocabulary of an ontology. A declaration also associates an entity category (class, datatype, object property, data property, annotation property, or individual) with the declared entity
  - Declaration(NamedIndividual(*:Peter*)): *Peter* is declared to be an individual

MITRE

OWL 2 (4)

- **Profiles:**
  - OWL 1 defined two major dialects, OWL DL and OWL Full, and one syntactic subset (OWL Lite)
  - Needs:
    - Some large-scale applications (e.g., in the life sciences) are mainly concerned with language scalability and reasoning performance problems and are willing to trade off some expressiveness in return for computational guarantees, particularly w.r.t. classification
    - Other applications involve databases and so need to access such data directly via relational queries (e.g., SQL)
    - Other applications are concerned with interoperability of the ontology language with rules and existing rule engines
  - Therefore, 3 profiles (sublanguages, i.e., syntactic subsets of OWL 2) are defined: OWL 2 EL, OWL 2 QL, and OWL 2 RL*

- **And more!**

*http://www.w3.org/TR/owl2-profiles/

MITRE

Protégé 4.1: OWL Pizza Ontology: Local Property Restrictions

Local Property Restrictions: for example, vegetarian pizzas should not have fish or meat toppings

Protégé 4.1: OWL Pizza Ontology: Caprina Pizza
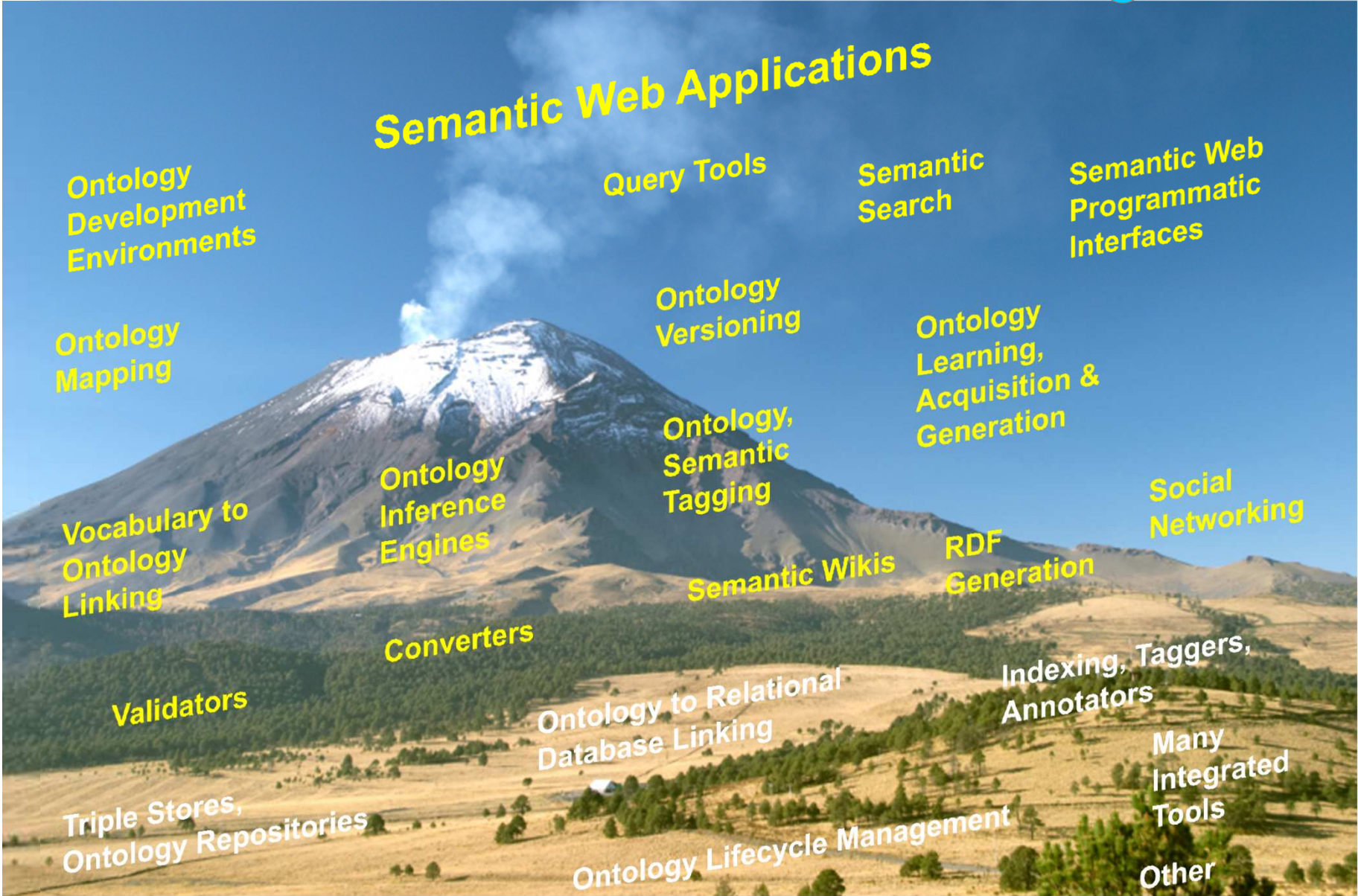

Caprina pizza is a named pizza having only these toppings.

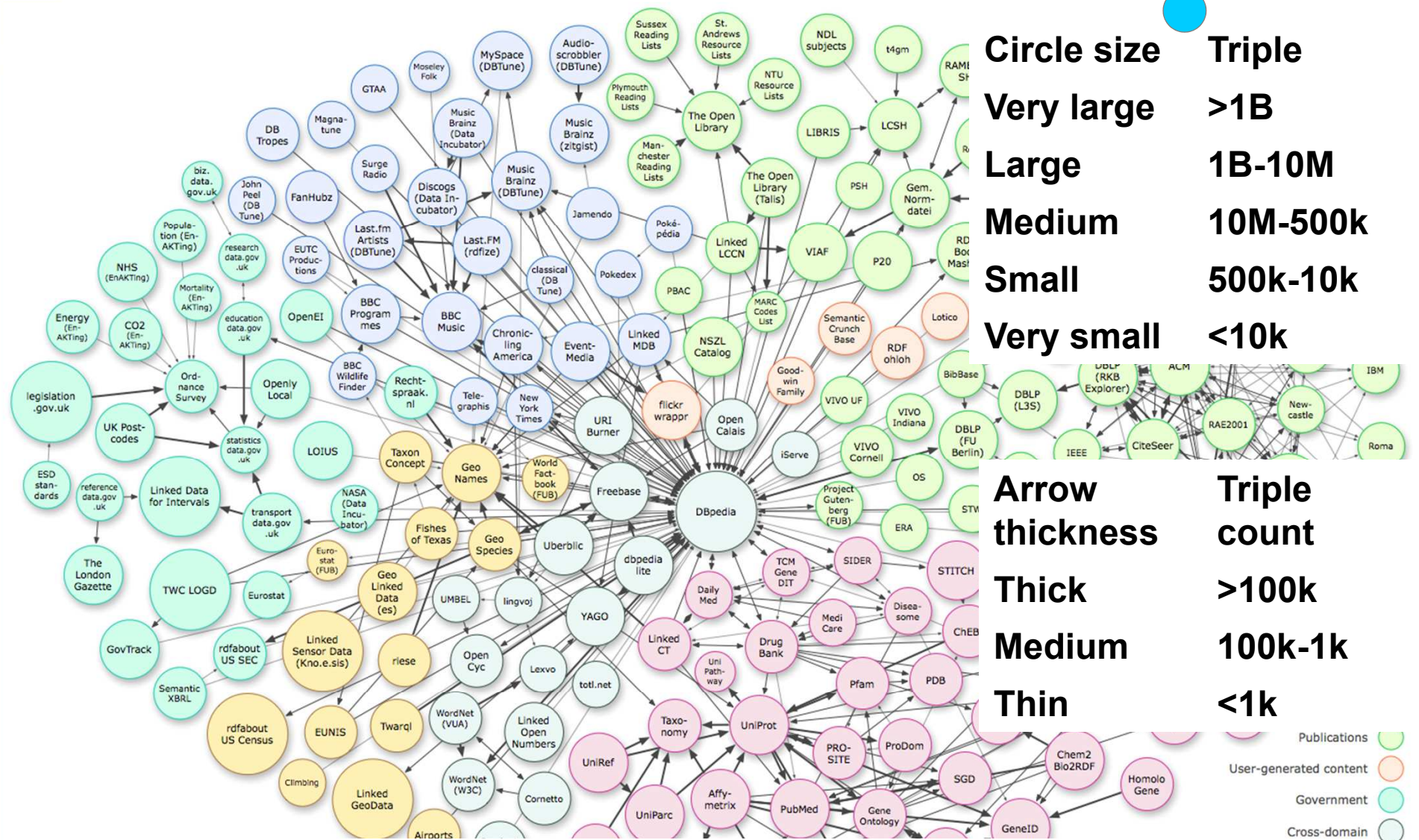# Protégé 4.1: OWL Pizza Ontology: Inferred Model Prior to HermiT 1.3.4 Reasoning, & After

Prior to reasoning: bare inferred model

After reasoning: extensive inferred model & Caprina

MITRE
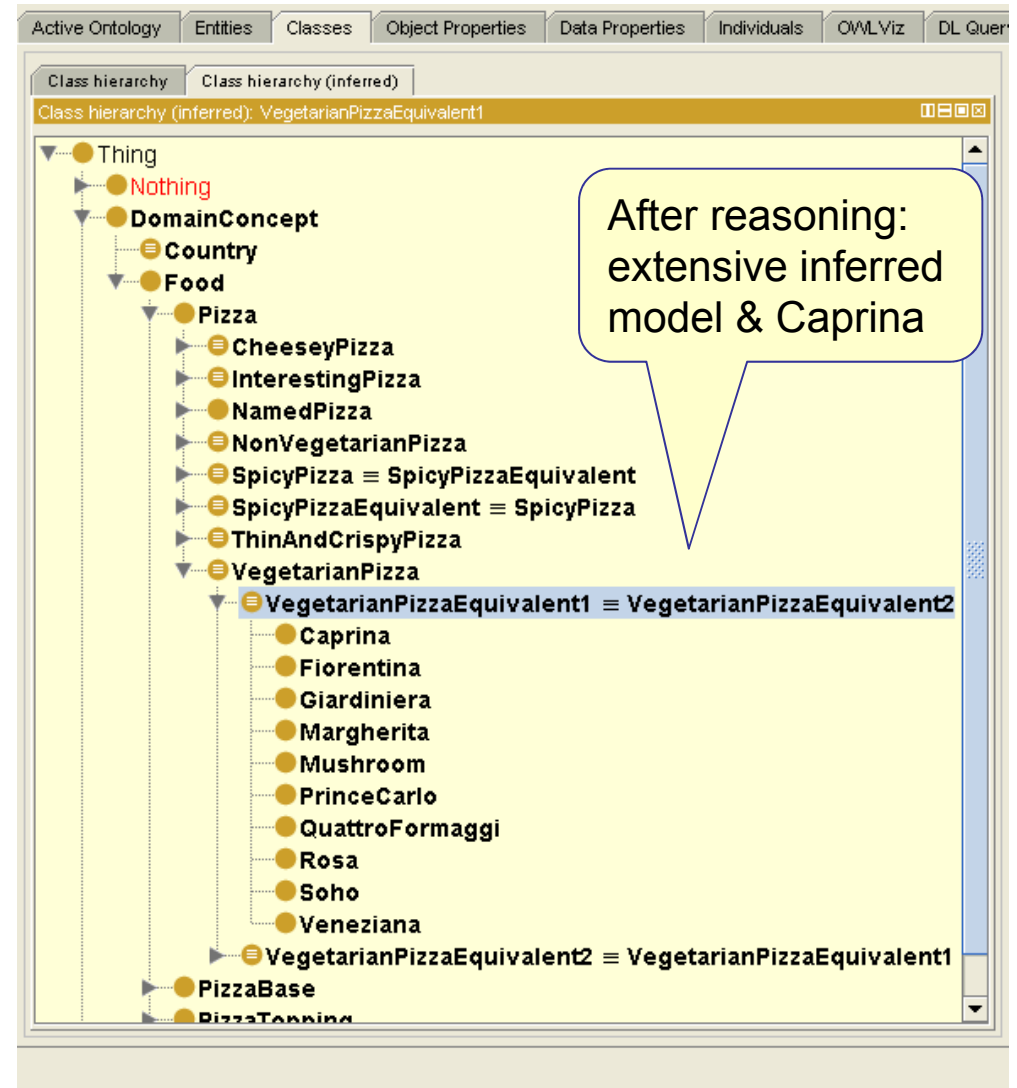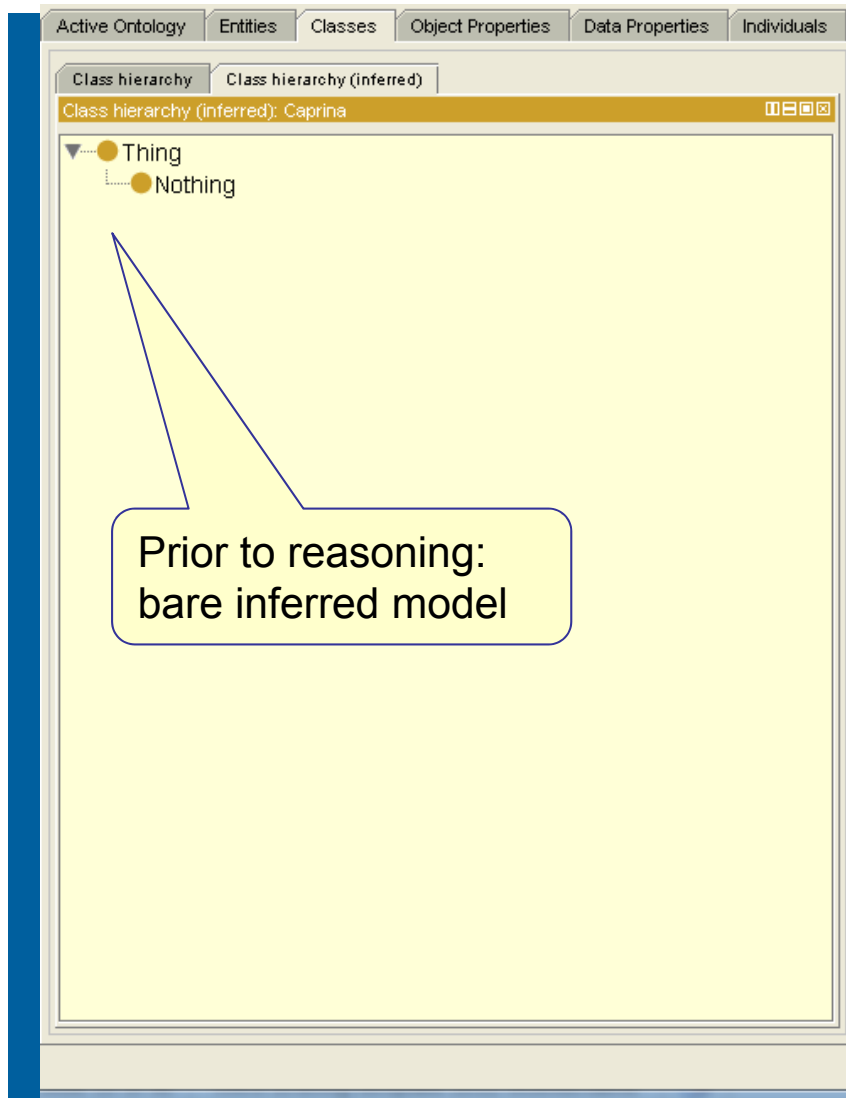
Semantic Web Tool Landscape

Linked Data: Sept. 2010: 13 billion triples

http://richard.cyganiak.de/2007/10/lod/

| Circle size	Triple
Very large	>1B
Large	1B-10M
Medium	10M-500k
Small	500k-10k
Very small	<10k

Arrow thickness	Triple count
Thick	>100k
Medium	100k-1k
Thin	<1k

Publications
User-generated content
Government
Cross-domain

In May, 2009, 4.7 billion RDF triples, interlinked by around 142 million RDF links, reported by W3C's Linking Open Data Project. In Sept, 2010, 13 billion triples.

Linked Data: Sept. 2010: 31.5 billion triples

http://wifo5-03.informatik.uni-mannheim.de/lodcloud/state/

As of September 2011
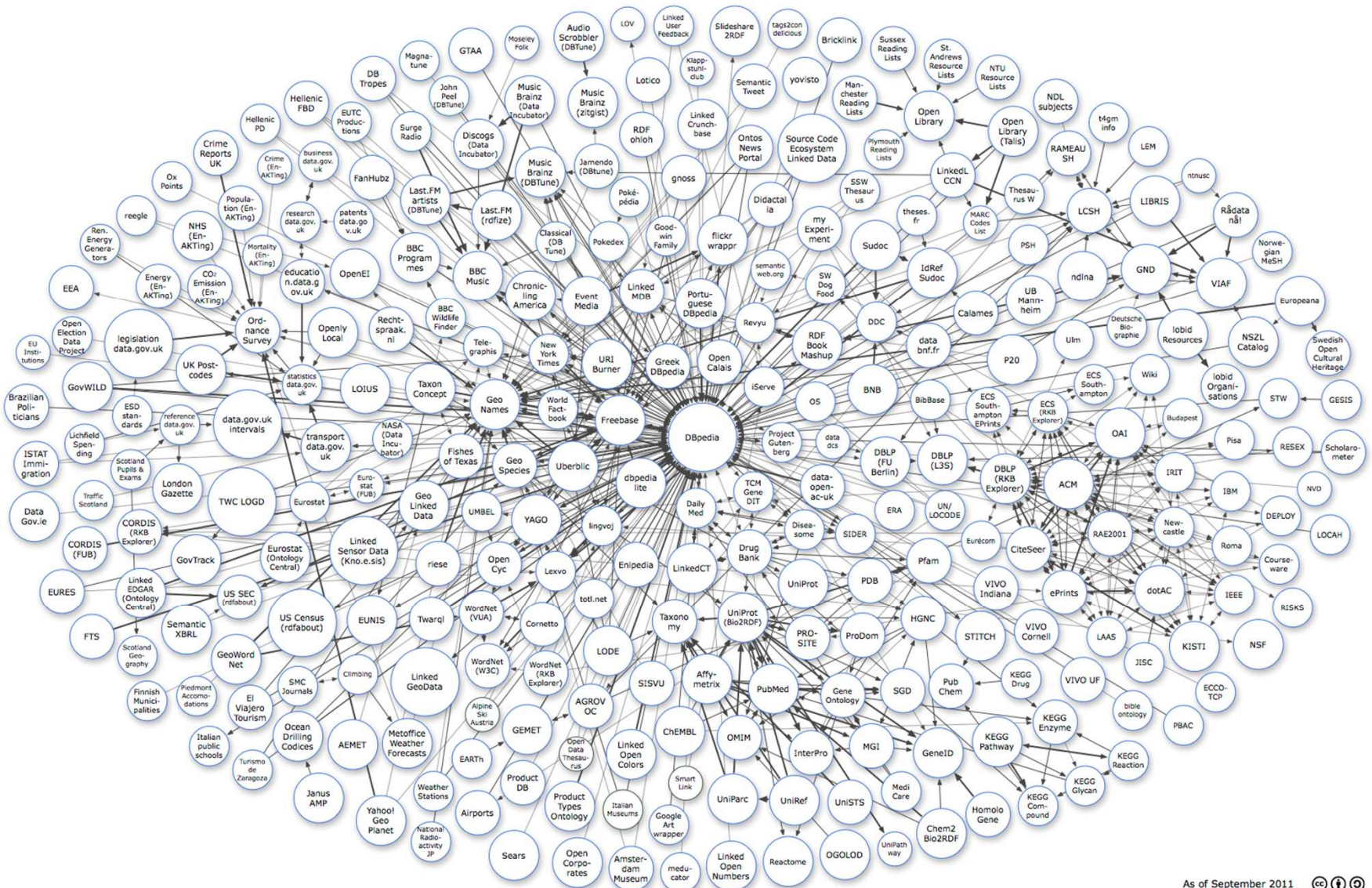
Tomorrow: Ontologies for Big (Complex) Systems & Big Data: Millions of Semantic Interactions
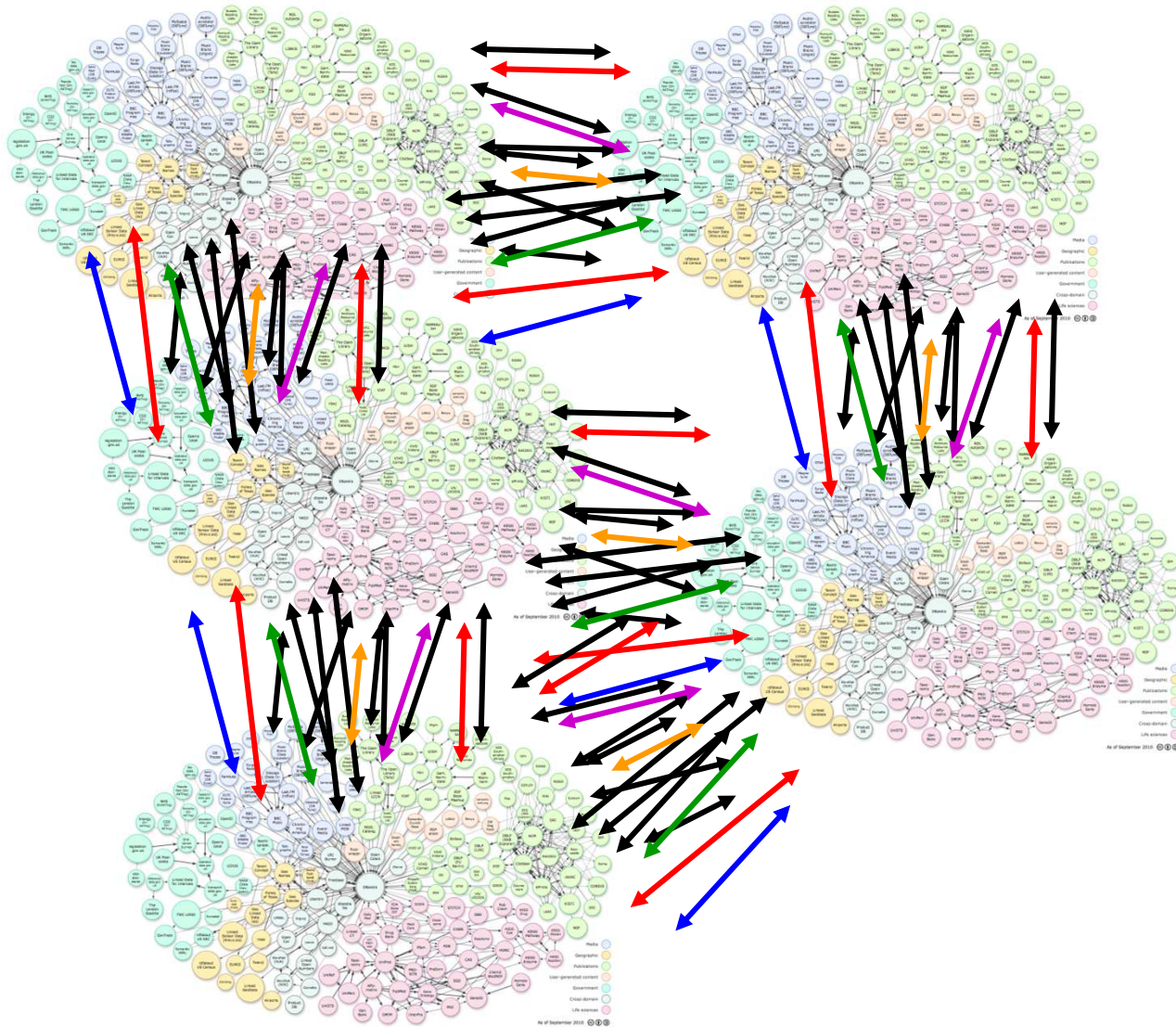
And more: RIF, SPARQL

More is coming: Common Logic

Thanks! Questions?