Semantic Web Tool Landscape

CENDI-NFAIS-FLICC Conference
National Archives Building
November 17, 2009

Dr. Leo Obrst
MITRE
Information Semantics Group
Information Discovery & Understanding
Command and Control Center

MITRE

Copyright © 2009 The MITRE Corporation

UNCLASSIFIED

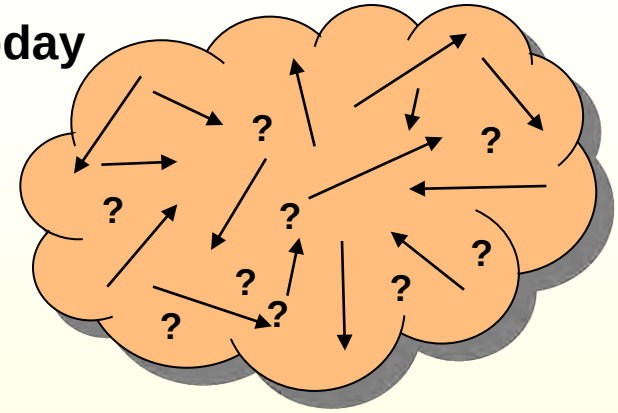
Overview

- Semantic Web
- Prefatory Comments and Caveats
- Here is What We Did 5 Years Ago
- Semantic Web Context
- Recent Semantic Web (and Linked Data) Tool Surveys
- Semantic Web Tool Landscape
- A Vertical View
- Some Tools
- Rest for Another Day
- Some Examples
- Trends and Indications
- Linked Data
- Are We There Yet?

The Semantic Web

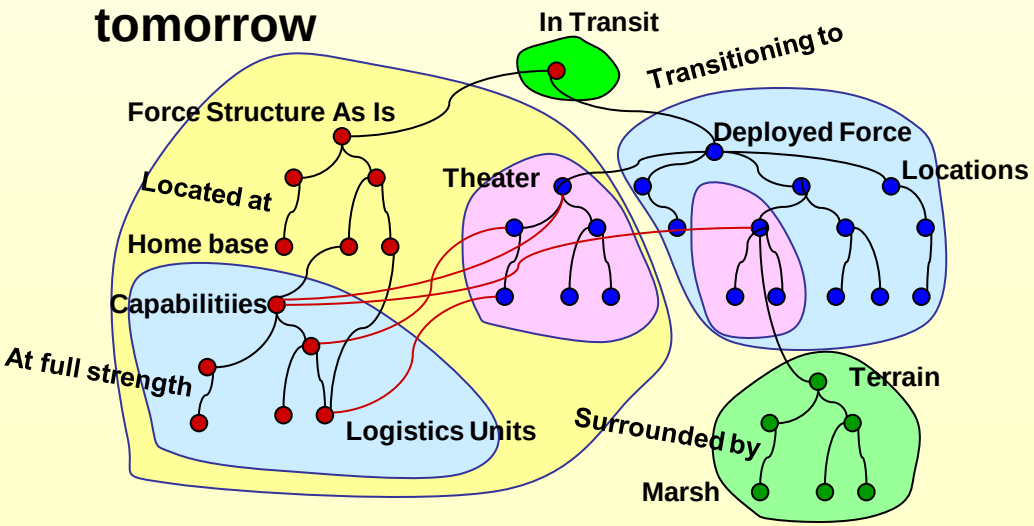
- Current Web is a collection of links and resources
  - Is syntactic & structural only
  - Excludes semantic interoperability at high levels.
  - Google of today is string based (keyword) & has no notion of the semantics (meaning) of your query
- Semantic Web extends the Current Web so information is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

Web today



Humans have to do the understanding

Semantic Web tomorrow



Machines partially understand what humans mean

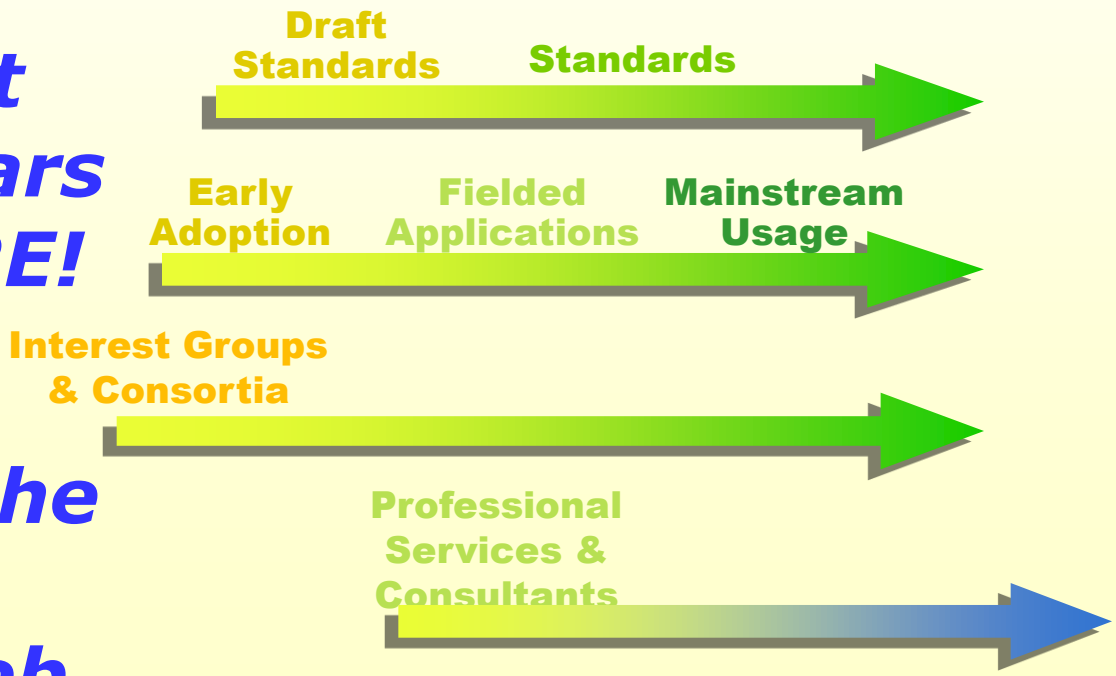
Prefatory Comments & Caveats

- Some tools span multiple functions
  - Example: some ontology development environments assist with full ontology lifecycle management
  - Example: some data integration, search tools provide ontology development environments
- Some tools blur the distinctions between vocabularies and ontologies (e.g., SKOS vs. OWL)
  - Vocabularies: focus on terms (natural language words & phrases, their lexical semantic relations)
  - Ontologies: focus on concepts (real world referents and categories of those), represent what the terms mean
- Many commercial tools are based on semantic technologies, may import/export SW content: SmartLogic, OntologyWorks, etc.
- Some tools are based on more expressive standards, such as ISO Common Logic, Horn Logic (logic programming), etc.
- For every tool I mention, there are 100 more that go unmentioned!

State of Research → Practice
A Maturity Continuum

MITRE

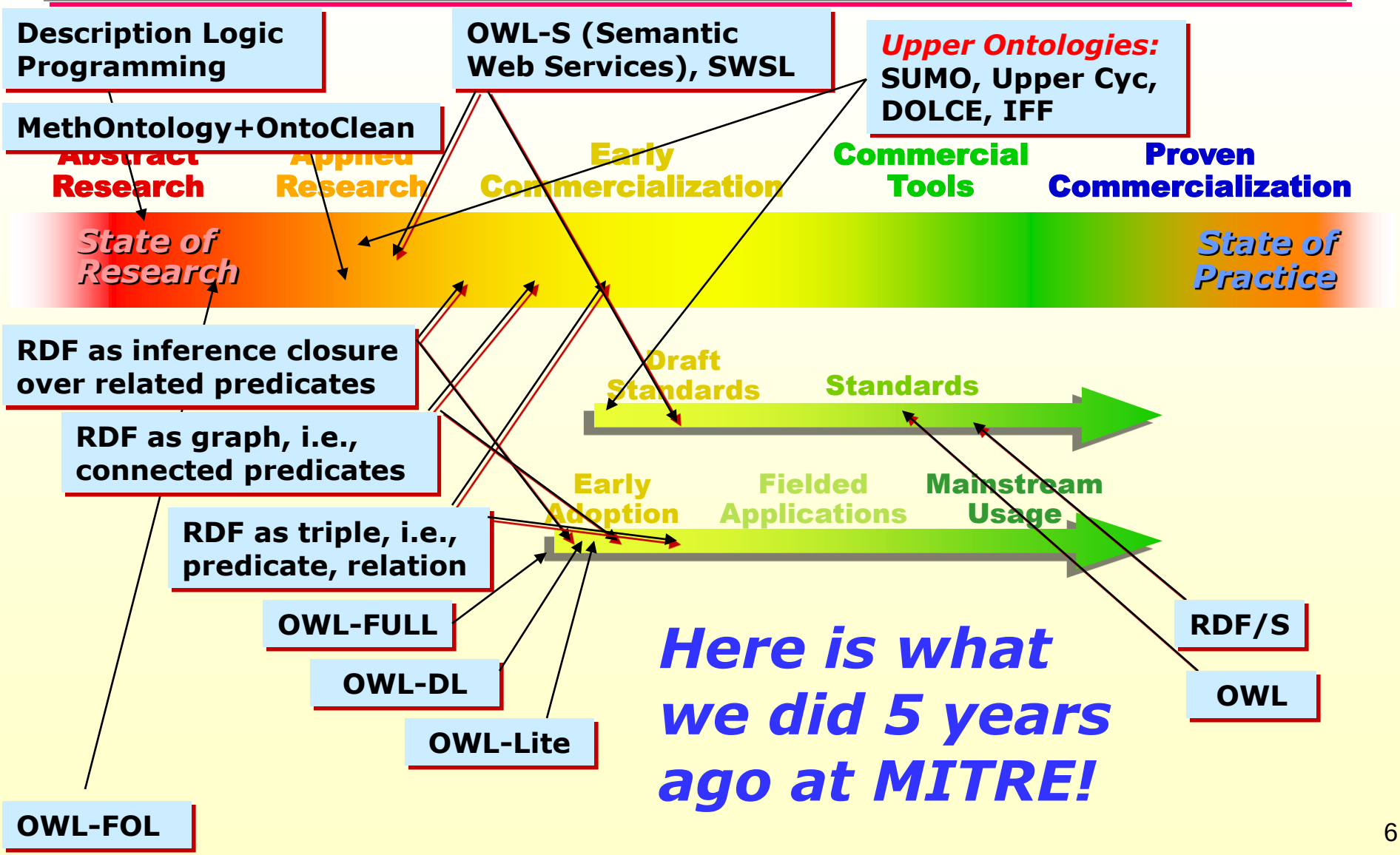
Here is what we did 5 years ago at MITRE!



We gauged the state of the Semantic Web

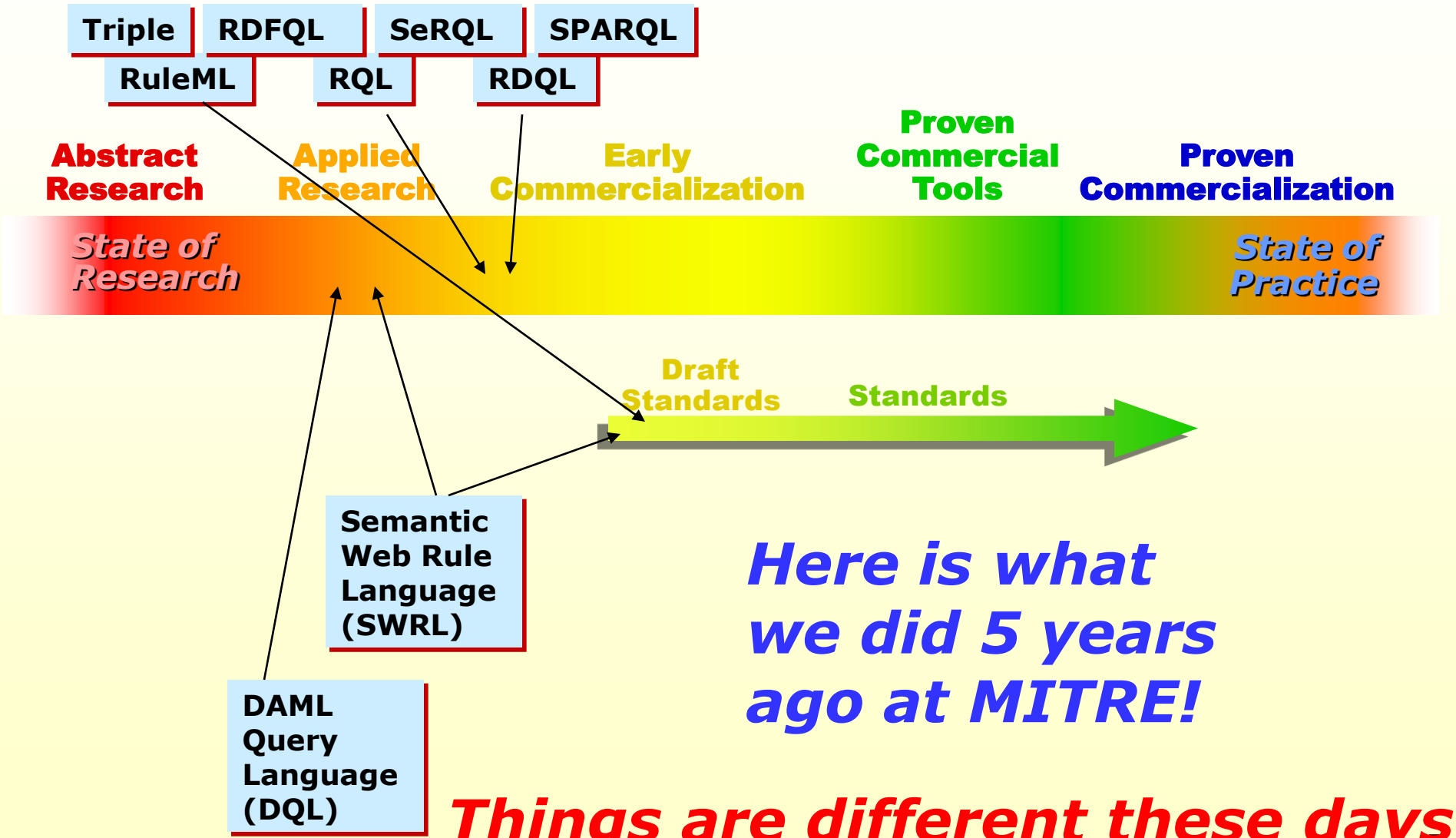
State of Research → Practice
Languages, Ontologies, Methodologies

MITRE



State of Research → Practice
Rules / Query Languages

MITRE

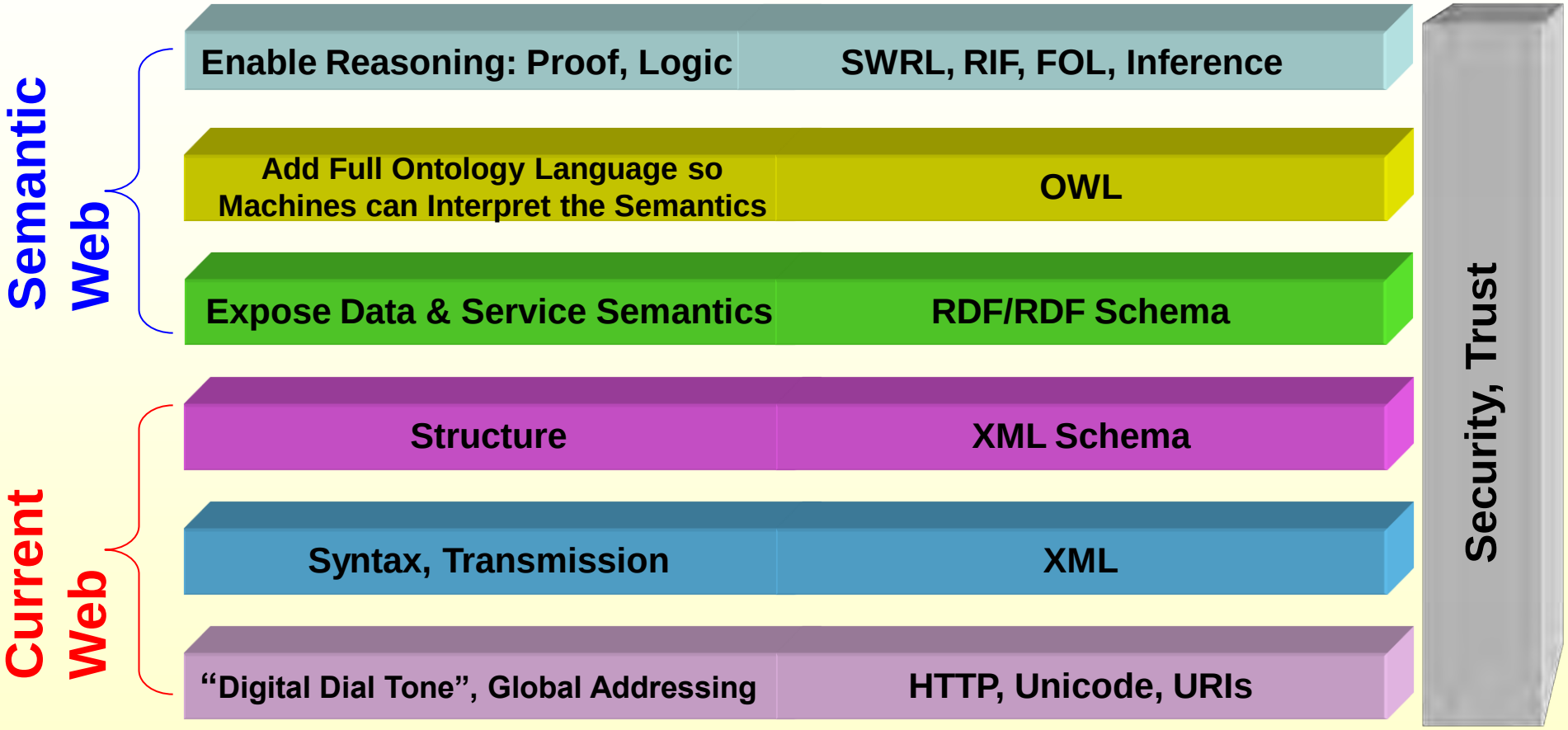


Here is what we did 5 years ago at MITRE!

Things are different these days

MITRE

Semantic Web Context



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

Recent Semantic Web (and Linked Data) Tool Surveys 1

MITRE

Semantic Web Tools: http://esw.w3.org/topic/SemanticWebTools

1. Semantic Web Development Tools: Introduction
2. General Development Environments, Editors, Content Management Systems, ...
3. RDF Triple Store Systems
4. Programming Environments
5. Reasoners (OWL or rule based) and related tools
6. Modularization and diffing tools
7. RDF Generators
8. On-line Validators
9. SPARQL "Endpoints"
10. Special Browsers
    1. RDF and/or OWL Browsers
    2. Semantic Web Browsers
11. Search Engines
12. Tools for Converting RDF Data between different Serializations
13. Tools for Converting Classifications and Other KOS into RDF-S or OWL Ontologies
14. Tagging and Semantic Web Vocabularies

Recent Semantic Web (and Linked Data) Tool Surveys 2

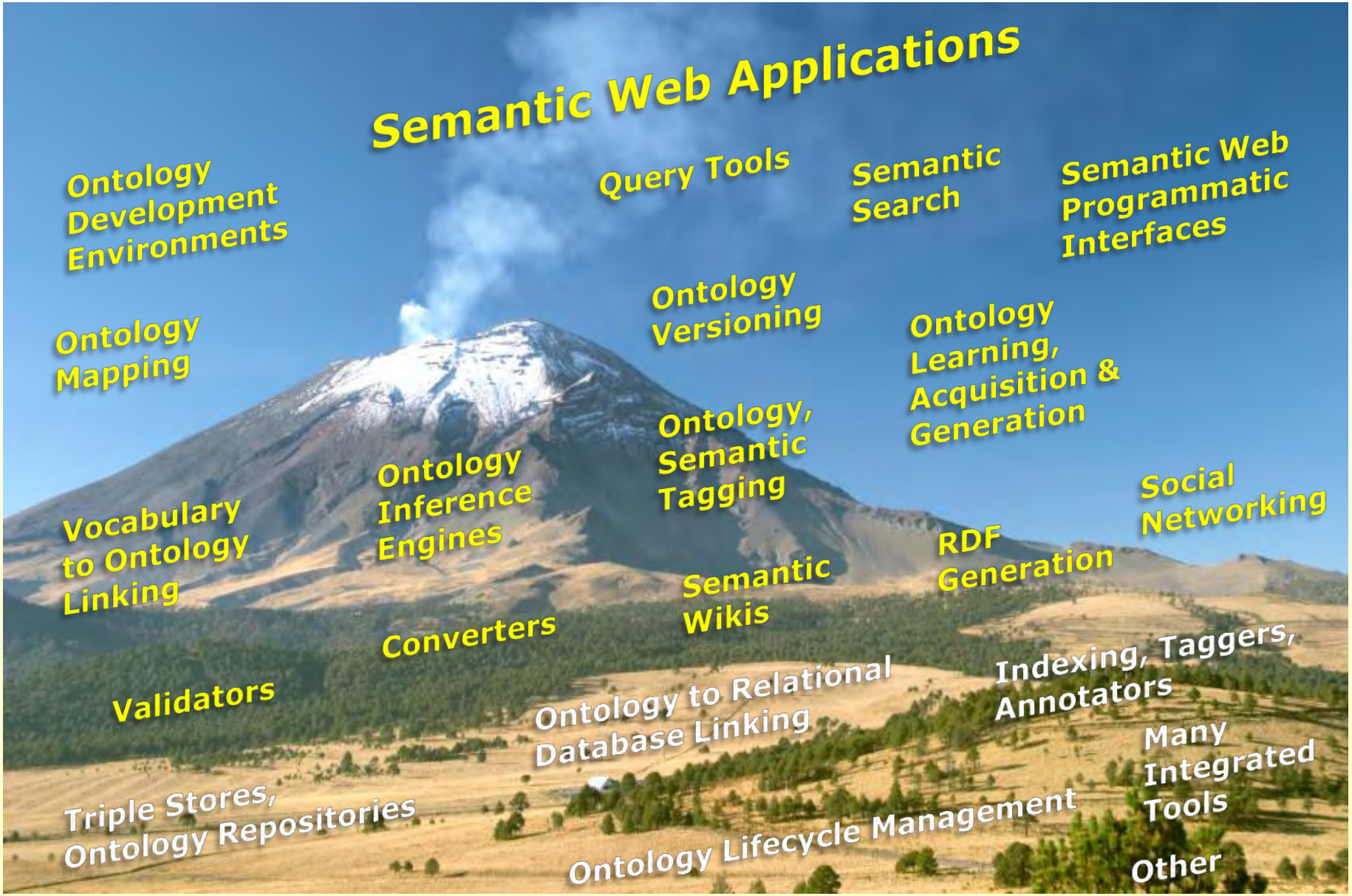
MITRE

- Lee Feigenbaum's 2009 Semantic Web Landscape: Technologies Tools, and Projects: http://www.slideshare.net/LeeFeigenbaum/semantic-web-landscape-2009
- Semantic Web Tools: http://semanticweb.org/wiki/Tools. About 120 tools.
- W3C Semantic Web Frequently Asked Questions: http://www.w3.org/RDF/FAQ

- These provide extensive references!

MITRE

Semantic Web Tool Landscape

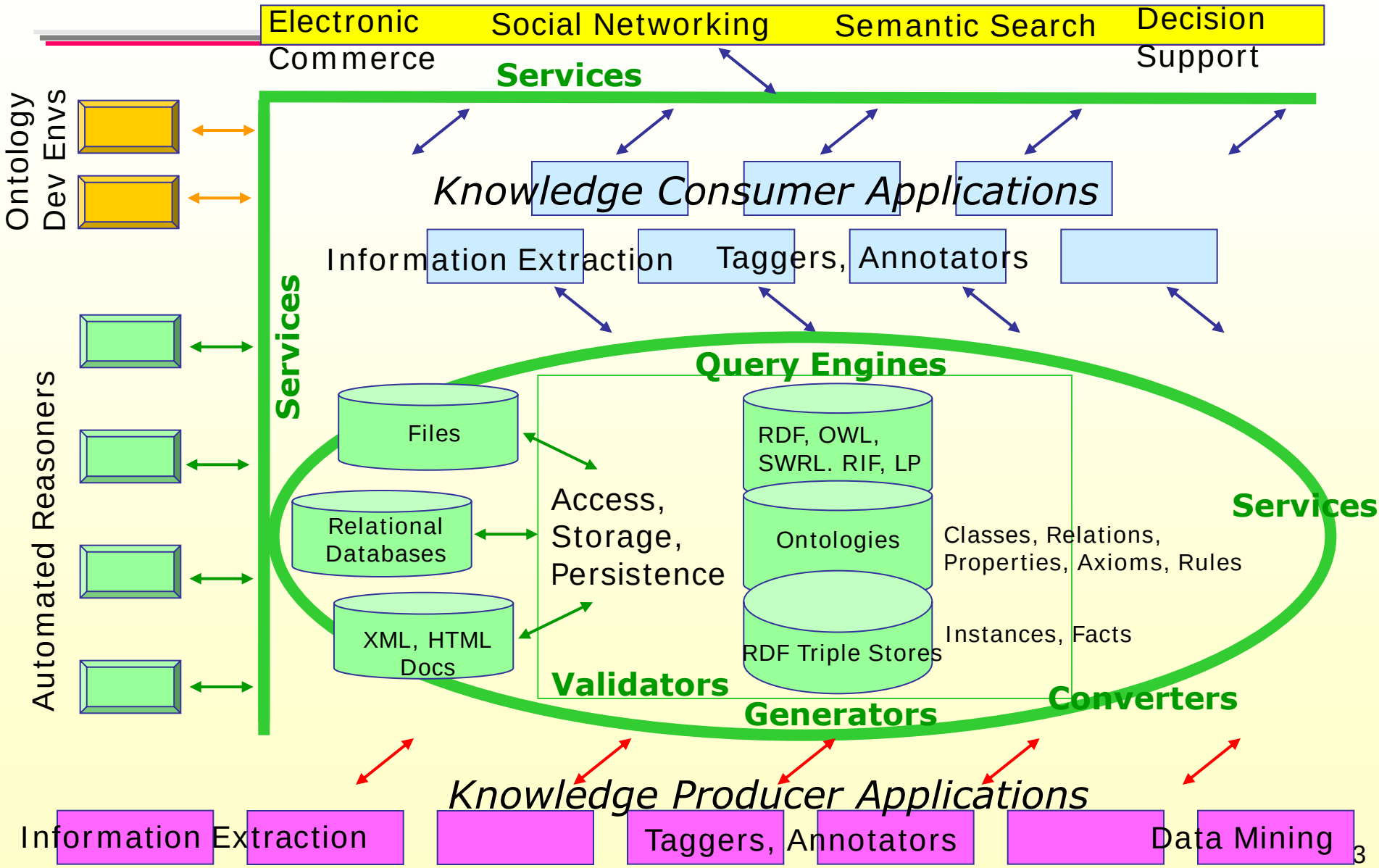


MITRE

Semantic Web Tool Landscape

Let a Thousand Flowers Bloom!

A Vertical View



Tool Categories

- Ontology Development Environments
- Ontology Reasoners
  - Description Logic Reasoners
  - Rule Reasoners
  - Probabilistic Reasoners
  - Hypothetical Reasoners
- Triple Stores (generally RDF instances in a graph), Ontology Repositories
- Query Languages, Engines, Semantic Search Tools
- Semantic Web Programmatic Interfaces
- Ontology Modularity Tools
- Ontology Versioning Tools
- Ontology Mapping Tools
- Controlled Vocabularies to Ontologies Linking Tools
- Ontology/Semantics/Metadata Tagging Tools
- RDF Generation Tools
- Ontologies to Relational Databases Linking Tools
- Ontology Lifecycle Management Tools
- Validators
- Converters
- Indexing, Taggers, Annotators
- Semantic Wikis, Social Networking Tools

Ontology Development Environments

MITRE

- Protégé 4.0, Collaborative Protégé
- TopQuadrant's TopBraid Suite
- NeOn: http://www.neon-toolkit.org/
  - Reference implementation of the NeOn architecture
  - Supports ontology engineering and management
  - Supports complete ontology lifecycle
  - Supports different ontology languages (OWL, F-Logic)
  - Supports networked ontologies (modules, mappings)
  - Free academic; commercial; eval license
- Altova's SemanticWorks
- KAON2 (KArlsruhe ONtology)

Ontology Reasoners

- Description Logic Reasoners
  - FaCT
  - FaCT++
  - Racer
  - RacerPro
  - Description Logic Programming (DLP)
  - Pellet
- Rule-based, Horn Logic, and First-Order Logic Reasoners
  - Jena
  - JESS (Java Expert System Shell)
  - SILK (Semantic Inferencing on Large Knowledge): http://silk.semwebcentral.org/

Storage: Triple Stores, Ontology Repositories

MITRE

- Triple Stores
  - AllegroGraph
  - Oracle Spatial 11g
  - OWLIM
  - SESAME
  - Mulgara
  - Drupal
  - Garlic 4Store
- Ontology Repositories
  - Swoogle
  - BioPortal
  - XMDR
  - Open Ontology Repository (OOR)
    - Federation of registries/repositories: BioPortal, XMDR, Semantic MediaWiki, etc.
- Other (vocabulary, etc.)
  - Synaptica Taxonomy Manager: includes Zthes, SKOS

Query Tools, Semantic Search

- SPARQL
- Twinkle
- TopBraid's SPIN (SPARQL Inference Notation): constraint checking, user-defined SPARQL functions, rules
  - SPARQLMotion: a scripting language for importing, exporting, processing data in RDF
- Sindice
- Many tools listed under "Trends and Indications"

Semantic Web Programmatic Interfaces

MITRE

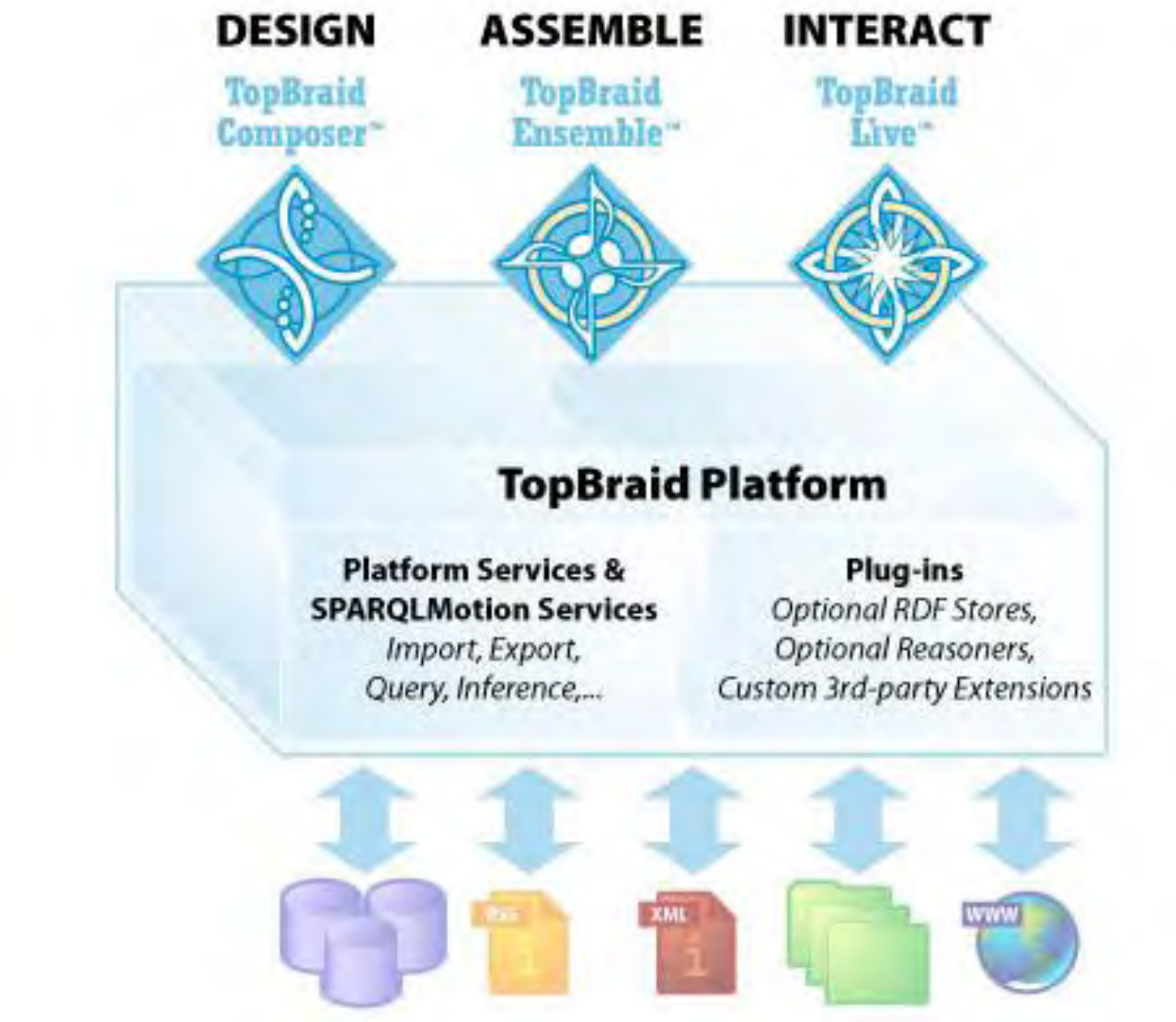
- See Semantic Web Tools: http://esw.w3.org/topic/SemanticWebTools
  1. Multi-language environments
  2. Java Developers
  3. Python Developers
  4. C Developers
  5. C++ Developers
  6. C# and .Net Developers
  7. Flex/ActionScript Developers
  8. JavaScript Developers
  9. Tcl/Tk Developers
  10. PHP Developers
  11. Lisp Developers
  12. Obj-C Developers
  13. Prolog Developers
  14. Perl Developers
  15. Ruby Developers
  16. Haskell Developers

The Rest: For Another Day

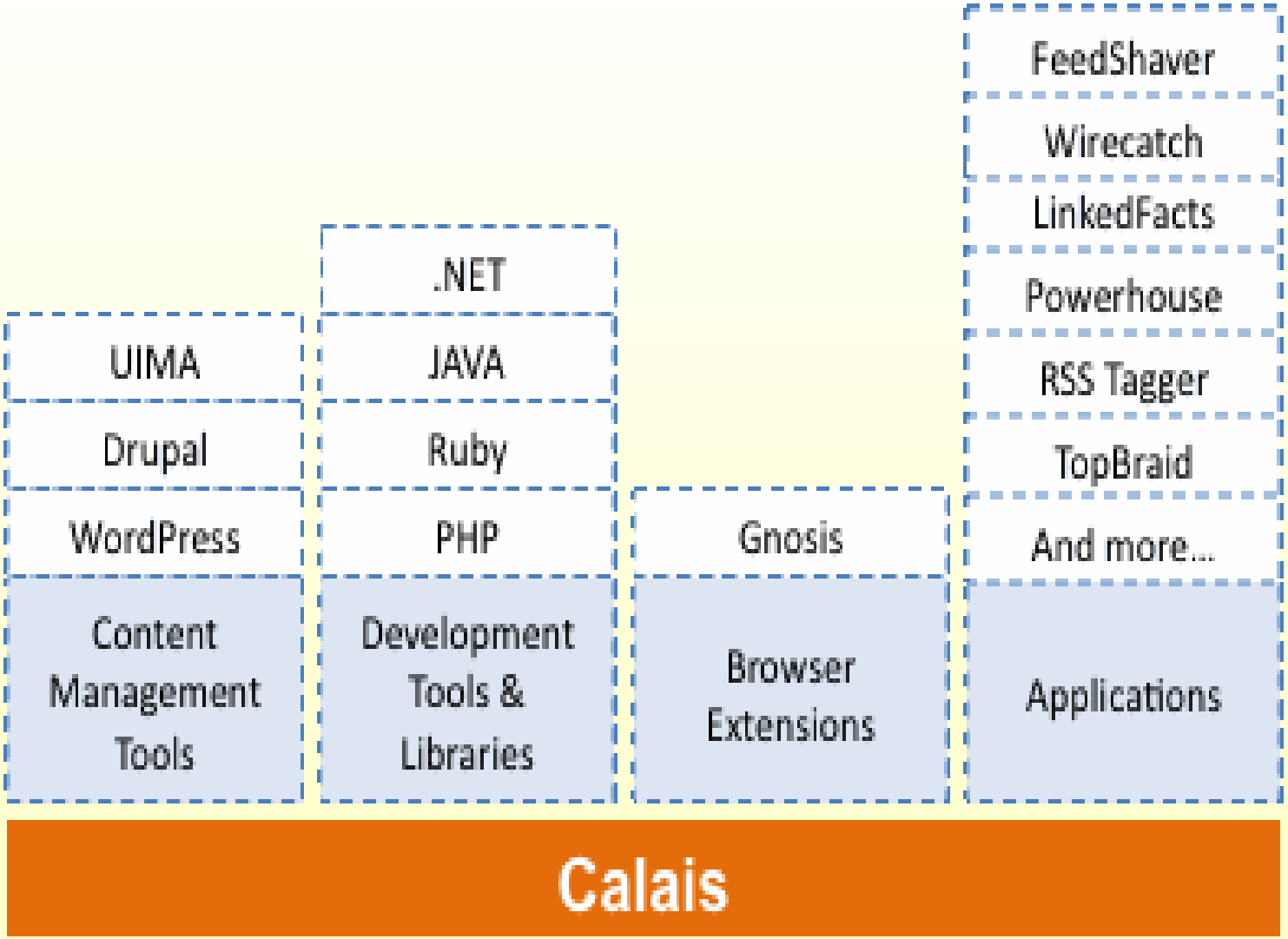
- Ontology Modularity Tools
- Ontology Versioning Tools
- Ontology Mapping Tools
- Vocabularies to Ontologies Linking Tools
- RDF Generation Tools
- Ontologies to Relational Database Linking Tools
- Validators
- Converters
  - UML-OWL Generator
- Semantic Wikis, Social Networking Tools
  - Revelytix's Knoodl
  - Semantic MediaWiki
  - OntoWiki
- Other
  - Large Knowledge Collider (LarKc)
  - Good Relations Annotator

Example: TopQuadrant's TopBraid Suite

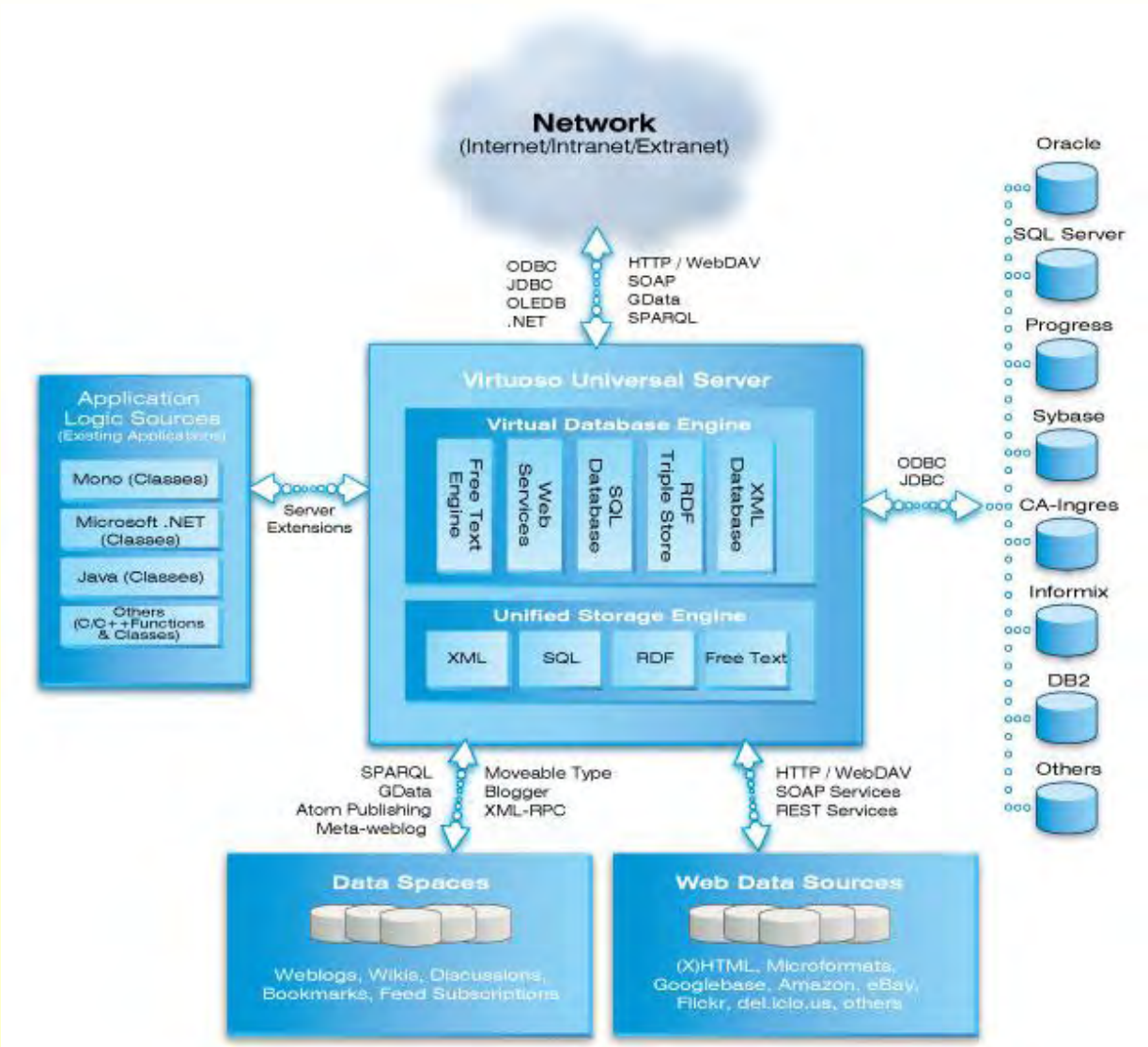
MITRE



Example: Open Calais Architecture



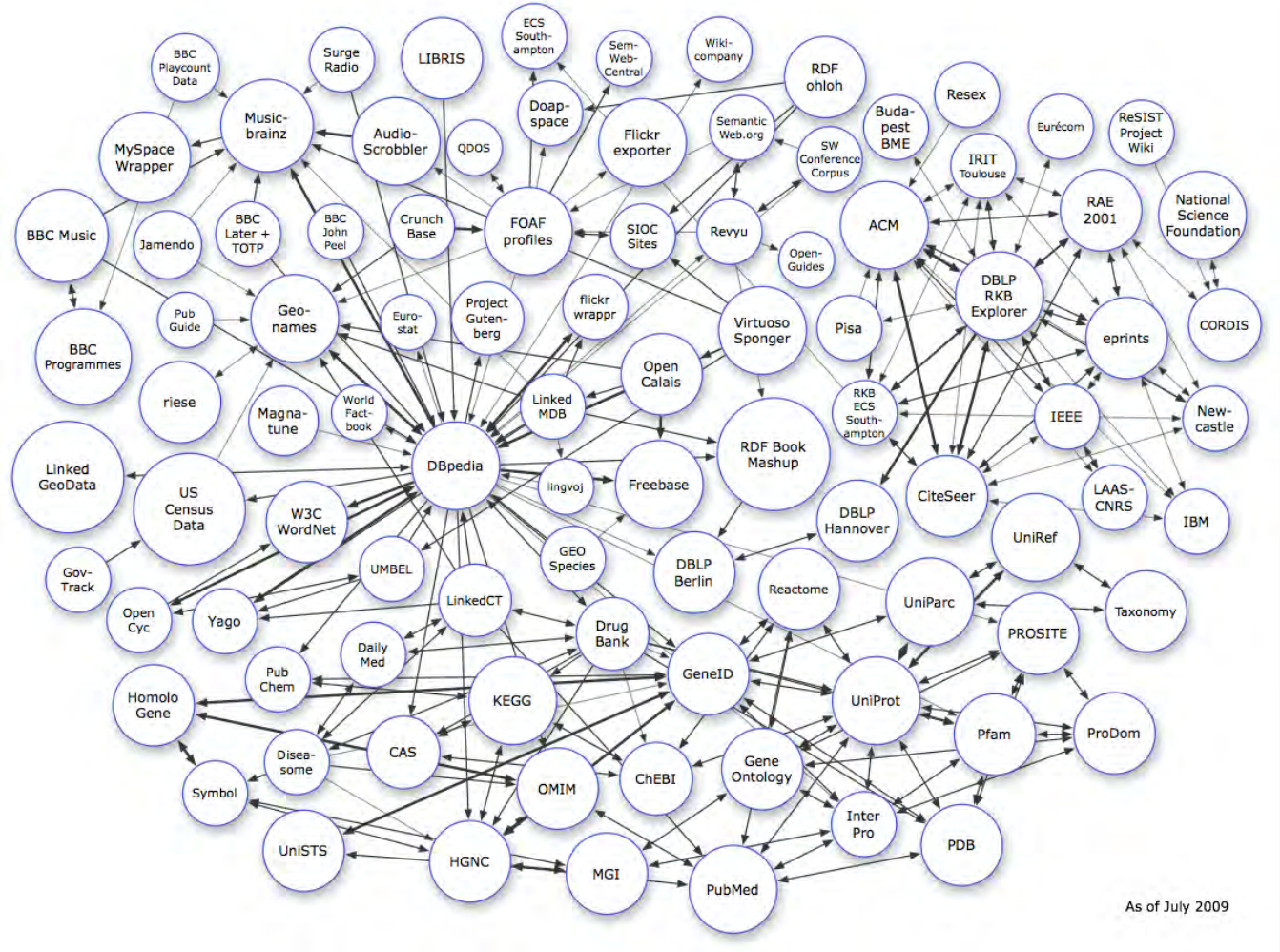
Example: OpenLink's Virtuoso



Trends and Indications

- Linked Data: strong movement to expose structured and unstructured data to Web using RDF and URIs
- RDF Generators
- Semantic Web Scripting & Mash-up Languages (or scripting languages with Semantic Web APIs, Libraries)
- Wider Semantic Technology applications
  - E.g., Infrastructure, Semantic Search, Computed Answers, etc.
  - UIMA, Wolfram Alpha, Google Fusion Tables, Open Calais,
  - OpenLink's Virtuoso, Zepheira, Freebase, Talis Group
  - Radar Network's Twine, Siri,
  - Metatomix, Hakia, True Knowledge, SmartLogic's Semaphore,
  - ThoughtWeb, Cambridge Semantics, Collibra, Inbenta
  - Semantic Discovery Systems, Structured Dynamics
- Large organizations exposing their data

Linked Data http://linkeddata.org/



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

Linked Data

- Exposing data structurally using dereferenceable URIs
- One might call the core of “Linked Data” the Structural Web, being constituted by RDF graphs
- Semantic Web ontologies describe those RDF graphs
- Need to expose the data + its semantics!
- Large organizations exposing data: NY Times, BBC, etc.
  - NY TIMES: http://open.blogs.nytimes.com/2009/10/29/first-5000-tags-released-to-the-linked-data-cloud/

Are We There Yet?

- RDF/s & OWL became W3C Recommendations, Feb. 2004
- Commercial vendors investing heavily
- Huge growth in number of conferences: ISWCs, SemTechs, etc.
- More ontologies being posted
- Research trend: SOA, Web services and Semantic Web convergence
- Tools & methodologies are maturing
  - Proof/trust/pragmatic web: still early
- Are we there yet?
- According to Tim Berners-Lee, not yet, but getting there!

http://gcn.com/articles/2009/10/30/berners-lee-semantic-web.aspx

MITRE

The Real Semantic Web Landscape

Copyright: http://crafttutorials.files.wordpress.com/2008/04/earth.jpg

Thanks!

- Questions?
- lobrst@mitre.org