

Toward a Web of Data and Programs

Steve Bratt



[Leading the Web to its Full Potential ...](#)

<http://www.w3.org/>

Speaking at the
IEEE-CS International Symposium
[Global Data Interoperability - Challenges and Technologies](#)
Sardinia, Italy
23 June 2005

Slides: <http://www.w3.org/2005/Talks/0623-sb-IEEEStorConf/>
Paper: <http://www.w3.org/2005/04/06-sb-IEEWebServData.pdf>

Summary

- *World Wide Web Consortium*
- *The challenge of global data interoperability*
- *The Semantic Web*
- *Lots of resources and examples*

W3C's Mission: Leading the Web to its Full Potential

Founded by Web inventor Tim Berners-Lee in 1994:

- **Providing the Vision to Lead.**
- **Engineering the Standards that Make the Web Work.**
 - **From a Web of Documents ...**
 - **... toward a Web of data and programs that is:**
 - **Meaning-ful**
 - **on Everything**
 - **for Everyone**
 - **... and Interoperable, Safe, Evolvable**



(Flyer: [US](#), [A4](#) / [History](#))

Toward a Web of Data and Programs

People and communities have data stores and programs to share

- *Vision: Expanding the Web to include more machine accessible resources*

Key Web technologies:

- **Web Services: *Web of Linked Programs***
 - Standards for interactions between programs on the Web
 - Easier to *Expose* and *Use* services (and data they provide)
- **Semantic Web: *Web of Linked Data***
 - Standards for data, relationships, descriptions on the Web
 - Easier to *Understand*, *Search for*, *Share*, *Re-Use*, *Aggregate*, *Extend* information

Challenge of Global Data Interoperability

Today's Communities



- Increasingly specialized ...
 - ... yet, important problems require multi-disciplinary and multi-organizational approach
- Massive quantities of data ...
 - ... yet, not readily understandable nor accessible outside of community
- Already using the Web ...
 - ... yet, mainly as a Web of documents, not as a Web of machine-processable data

Today's Web: *Searching*

Current Situation:

- Search Web for information on "Love waves"

Issue:

- Some useful, some not-so-useful hits:
 - "Love Wave Research at Brown"
 - "Courtney Love Waves the Red Flag"

- " 'Buy Faith, Hope Love: Waves and Reflections', at Amazon and SAVE"
- etc. ...

Today's Web : *Sharing and Aggregating*

Current Situation: Earthquake Monitoring:

- Web site A, B, C ...: fault, earthquake and other geo information
- Programs G, H, I ... : mapping, propagation, site response ...
- Email communications with X, Y, Z ... : gov't emergency authorities
- Want to *aggregate* data to (for R&D or realtime alerts):
 - compare network locations and compute new ones
 - examine correlations of geophysical observations
 - interface with authorities re: events and risks

Issues:

- Different databases, attributes, locations, services
- Different jargon, meanings
- Not easy to find, access, process resources
- Must build specialized systems to important, convert, update ..
 - highly inefficient

Why Can't Computers "Understand"?

Analogy:

What We Say to Dogs

- from Gary Larson cartoon ([local link](#)) ...
- "Stay out of the garbage! Understand, Ginger? Stay out of the garbage!"

What Dogs Understand

- "... blah blah blah **GINGER** blah blah blah ..."

What Computers "Understand"



" ... blah blah blah **** blah blah blah"*

- Computers must "**understand**" more:
 - Not human concept of "understanding"
 - Just useful machine processing

--

* where **** is a "link"

Toward Processable Search Semantics

Google:



- Web links are machine processable
- Minimal semantics assumed: "This" refers to "that" = popularity
- Amazing results from minimal semantics

What if Web pages had more semantics?

(See [Article by Bijan Parsia](#))

Ways to Enable Machine Processing

Smarter Machines

- Teach computers to infer the *meaning* of Web data
 - Natural language, image recognition, etc.
- ... this is the Artificial Intelligence approach
 - **WARNING:** Hard problem!

Smarter Data

- Make data *easier* for machines to find, access and process
 - Express data and meaning in standard machine-readable format
 - Support decentralized definition and management
- ... this is the *Semantic Web* approach

What Is the Semantic Web?

"The Semantic Web is an

... extension of the current web in which

... information is given well-defined meaning,

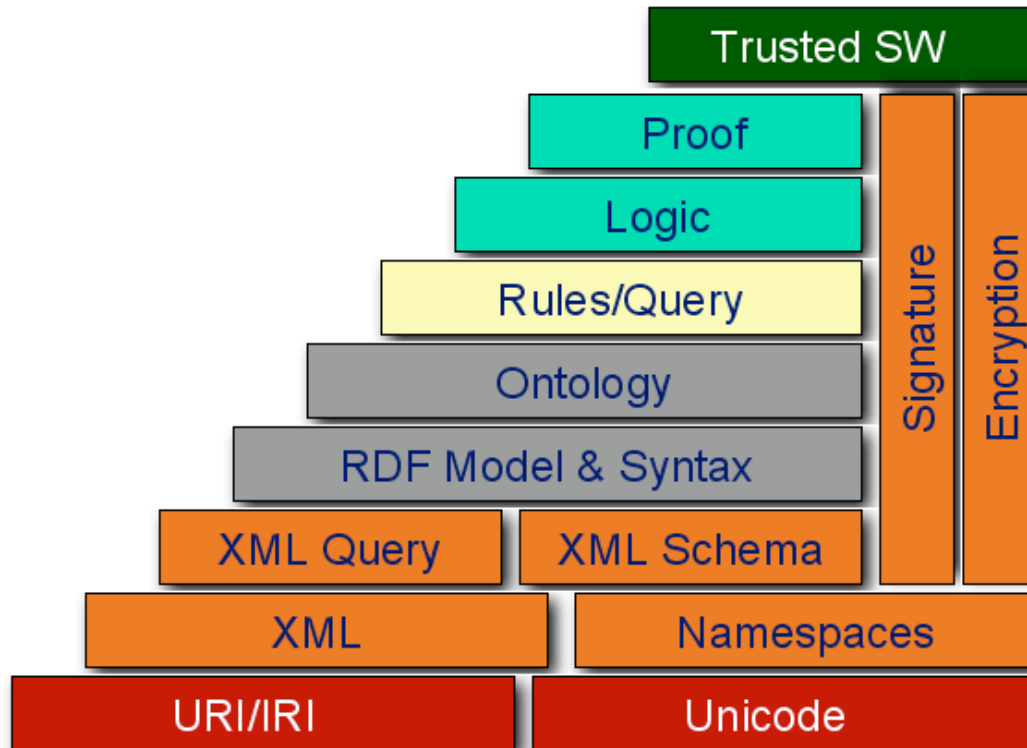
... better enabling computers and people to work in cooperation."

[The Semantic Web](#)

Tim Berners-Lee, James Hendler and Ora Lassila

Scientific American, May 2001

Semantic Web Standards



Machine-processable, global Web standards:

- Assigning unambiguous names (URI)
- Expressing data, especially metadata (RDF)
- Clear, rich ontologies (OWL)
- Query, rules, logic, proofs, trust (in progress)

Accessible wherever you want it to be ...

... on the Web.

Uniform Resource Identifiers (URI)

- URI has two different uses:
 1. Unambiguous name for something
 2. Location of a document
- URIs can be used to identify definitions for concepts
 - Especially useful for ontologies & metadata
- Examples:
 - <http://www.example.org/index.html#section2>
 - <mailto:steve@w3.org>

Resource Description Framework (RDF)

- W3C standard ([Primer](#))
 - Statements about things (concepts, objects, etc.)
 - Metadata+
- Descriptive statements expressed as *triples*:

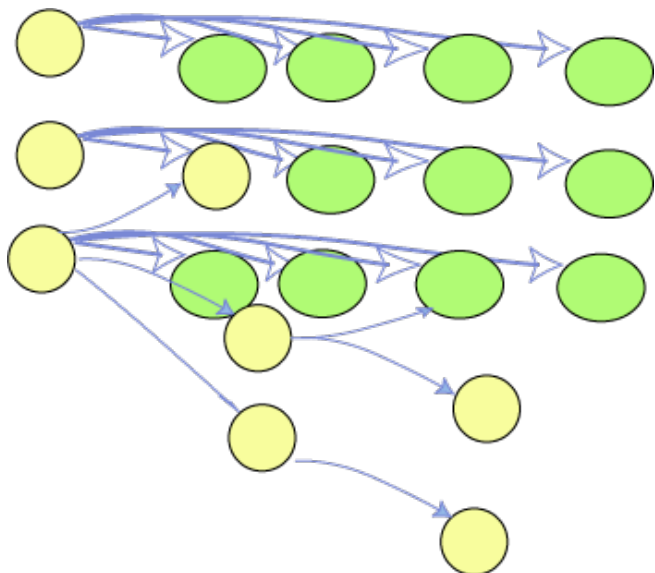
(*Subject, Predicate, Object*) or (*Subject, Property, Value*)

- All 3 of these could be URIs



Flexible Representation

Linking table and tree data ...



Web Ontology Language (OWL) +

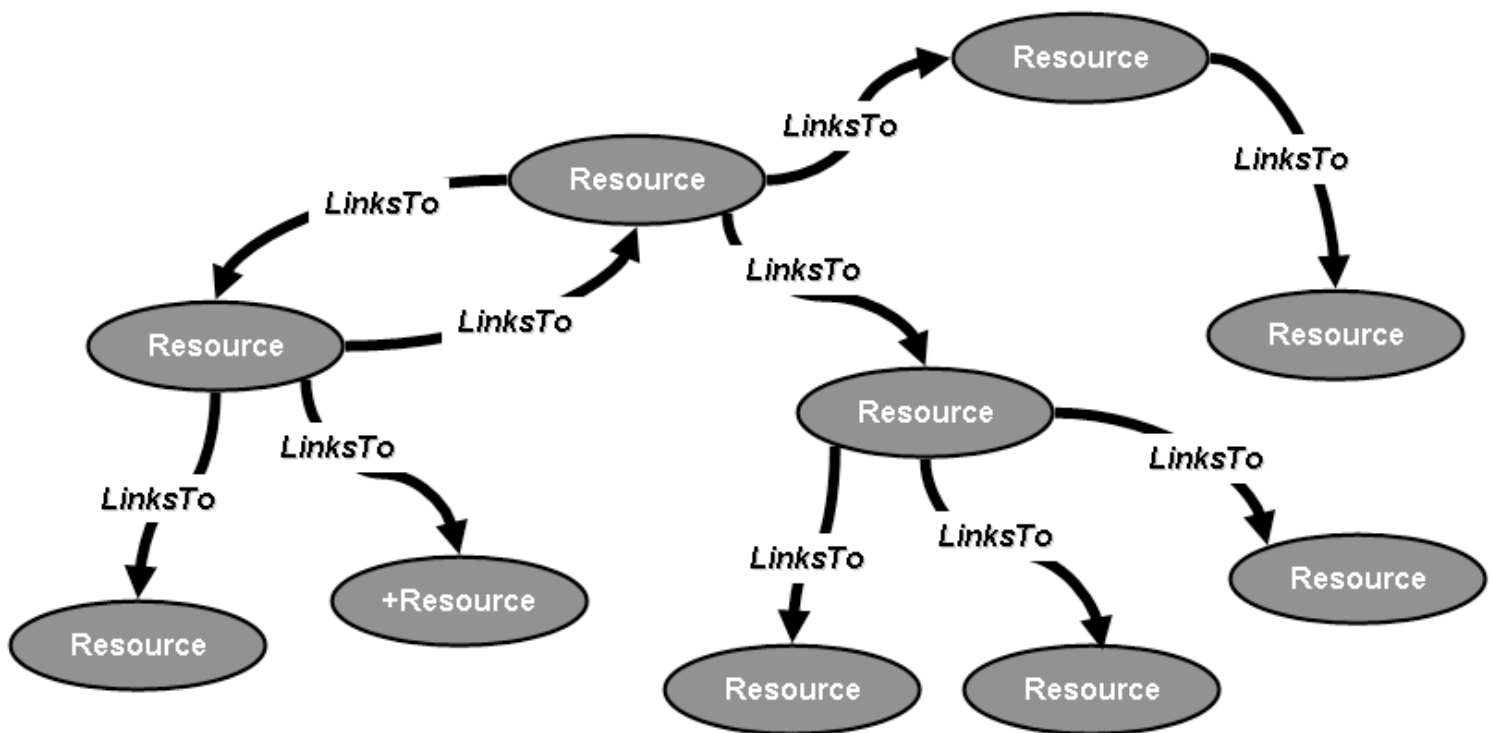
- W3C standard ([Guide](#), [Overview](#))
 - Defining concepts & relationships within area of knowledge
- Like a DBMS data dictionary, *but ..*
 - Defines more complex relationship
 - Global, standard syntax based on RDF
 - Using URIs as keys
 - On the Web
- Includes standard vocabulary for describing [properties and classes](#).:
 - Datatypes (e.g., integer, date, string ..)
 - Relations (e.g. subclass of, property of ..)
 - Cardinality (e.g. exactly one ..)
 - Characteristics of properties (e.g. symmetry)
 - ... and more

Standards En Route

- Query
 - [W3C RDF Data Access Working Group](#)
 - [SPARQL](#) - RDF Query Language
- Rules
 - [W3C Workshop on Rules Languages for Interoperability](#) (Apr 2005)
- Logic, proof, trust, etc.
- Tools and applications
 - [Semantic Web Services](#)
 - *See extra slides for many references*
 - growing efforts underway within industry, gov't, academia

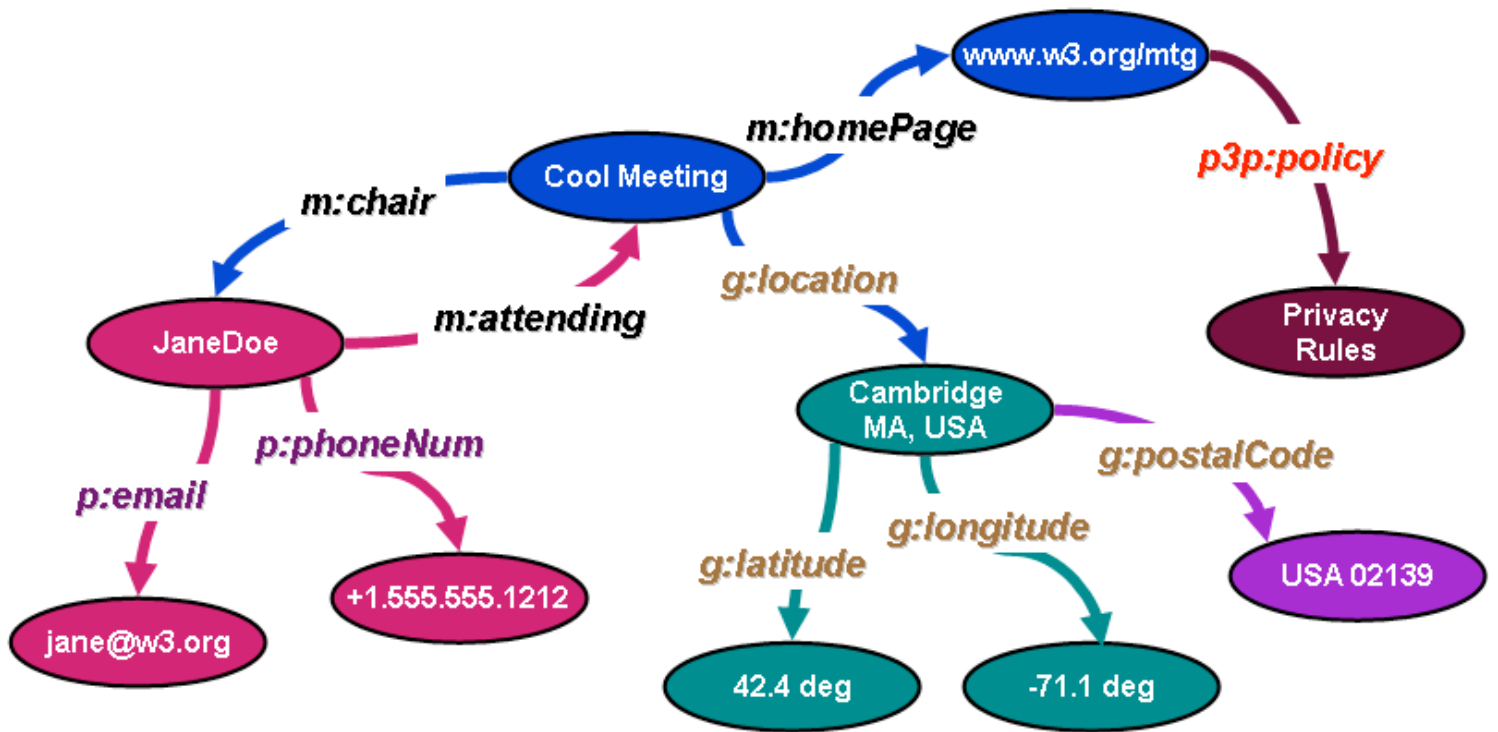
Current Web

- Minimal machine-processable information -- dumb links



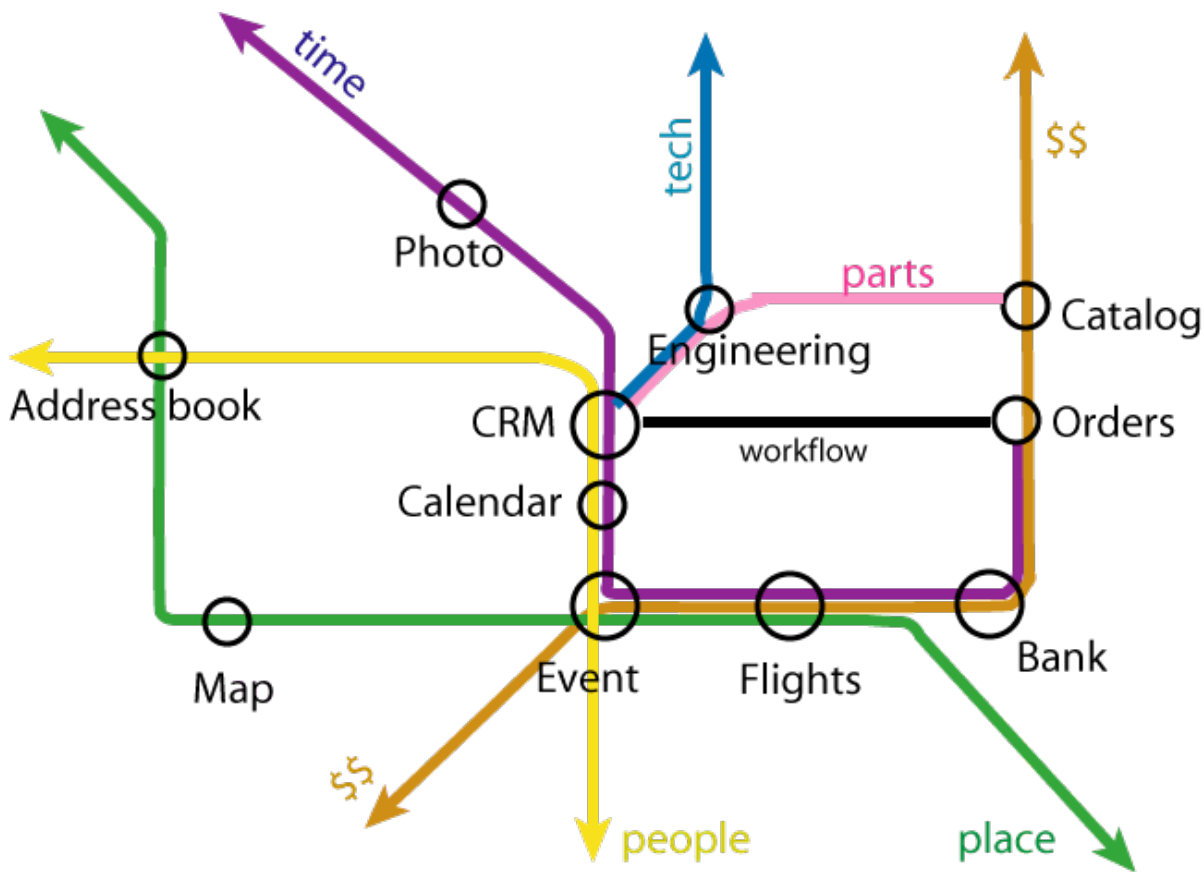
Semantic Web - Enhancement of Current Web

- More machine-processable information - applications connected by concepts



[ellipses = resources; ellipse/line color = one data source; x: = one ontology]

Applications Connected by Concepts



RDF as a Global Integration Bus

Tim's Slides?

Industry Taking Notice

- Pioneering domains
 - Life sciences
 - Enterprise integration
 - Personal information management
- Applications
 - Nokia, HP, IBM, Oracle
 -

Bottom-Up Evolution for Semantic Web

- Follow [Web standards](#) developments
- Participate in community efforts (e.g., [GEON](#), [Life Sci](#))
- Explore existing tools (see extra slides)
- Pick a few community applications and databases

- Semantically-enable existing [RDBMSs](#) and [XML stores](#)
- Develop key community ontologies
- Develop shared ontologies
 - Widely used vocabularies: math & sci, people, calendar, geospatial ...
 - Other links ontologies
- Exponential growth?
 - More machine accessible data
 - Better tools
 - Killer apps
 - Network effects

Global Data and Service Interoperability

Machine usable applications and processable data -- on the Web:

- Unique names for data, applications, descriptions, relationships (URIs)
- Web Service-enabled applications
- Web standards for data representation, ontologies, query, etc. (RDF, OWL, SPARQL, etc.)
- Possible benefits
 - Build, share, find, re-use, aggregate, extend data and systems
 - Build, share tools
 - Network effects
 - Increase opportunities for communication, collaboration, directed and serendipitous discovery



We are happy to help ...



Resources Slides

Subsequent slides provide information on ...

- W3C
- Web Services
- Web Services and Semantic Web tools and applications

International Web Standards Organization

W3C's expanding international basis of operations ([map](#))

- 3 **Hosts**: [Massachusetts Institute of Technology](#) | [European Research Consortium for Informatics and Mathematics](#) | [Keio University](#)
- 14 **Offices (locations)**: [Australia](#) | [Benelux/Bénélux](#) | [Suomi](#) | [Deutschland und Österreich](#) | [Ελλάδα](#) | [香港](#) | [Magyarország](#) | [ישראל](#) | [Italia](#) | [한국](#) | [المغرب](#) | [España](#) | [Sverige](#) | [United Kingdom and Ireland](#)

Announcing: New India Office at [CDAC](#) (more tomorrow)
Office discussions underway: China, Latin America.

- ~65 vendor-neutral [Technical Staff](#)

W3C's Members Power Our Work

~ 370 [Members](#) ([png](#), [map history](#))

- World's leading technology organizations
 - "Third-class companies make products; second-class companies develop technology; first-class companies set standards." - saying in China
- Commercial, academic, research institutions, government agencies, non-profits
- Developing Web products, leveraging the Web, for your work, conducting Web research, developing specifications built upon W3C's work

Values of W3C Membership

- **Leadership**
 - Introduce ideas through submissions & workshops
 - Influence standards through Working Group participation, review, implementation
 - Provide strategic direction through seat on Advisory Committee
- **Early insight into market trends**
 - Access Member & Team technical experts
 - Track emerging technologies & markets through Member-confidential access
 - Implement standards & plan for procurement of new technologies ahead of market
- **Promoting image as innovator**
 - Participate in international media activities, press releases, testimonials
 - Display organization and logo on W3C site (250K visits/day)
 - Display W3C Member logo on organization's page
 - Ensure & enjoy benefits of royalty-free implementation of Web standards

W3C Brings the World Together to Engineer the Web

- ~50 [Working, Interest and Coordination Groups](#)
 - Producing Web standards
 - Driven primarily by Members
 - Strongly supported by Team
 - Invited Experts

- **Liaisons**
 - with ~[40 international standards bodies](#)
- Accountable to the **Global Public**
 - Public lists, some public Interest Groups
 - Groups must address all submitted issues

Developing Standards at W3C

- Clear and effective **Process**, which encourages:
 - Member-**neutrality**
 - Consistent **architecture**
 - [Architecture of the World Wide Web, Vol 1](#)
 - **Coordination, consensus and interoperability**
 - No "rubber stamping"
 - Must engineer dependencies (inside and outside of W3C)
- Clear **Patent Policy**, with royalty-free objective

Views of the Process: [Life of Working Group](#), [Recommendation Track](#)

W3C Technologies

Over 80 Standards (Recommendations) have been released ([List](#) / [SVG-by-yr](#) / [Translations](#))



(Organization of work: [SVG-from-RDF](#), [PNG](#))

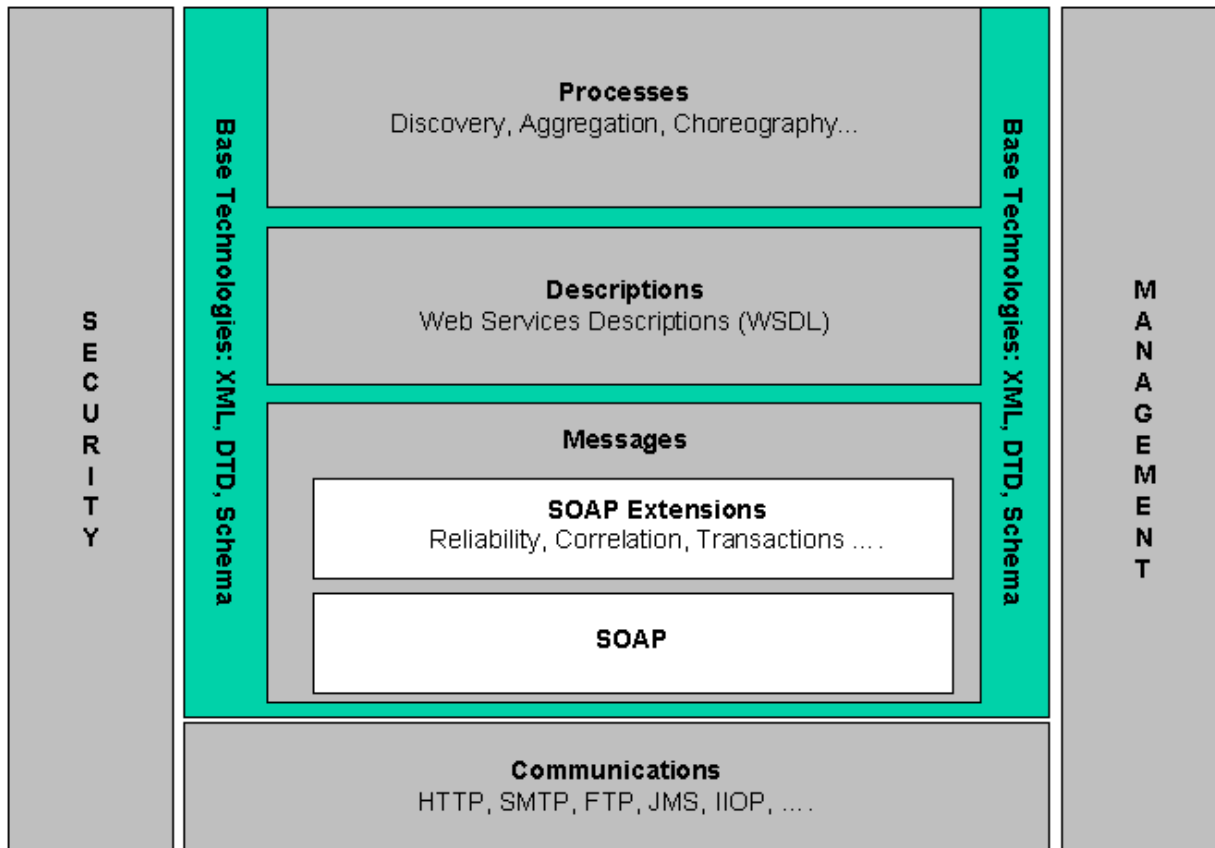
What's Next?

Recently started work, and ideas for new work ...

- **Web of Meaning**
 - **Semantic Web**: [Best Practices](#)*, [Query](#)*, [Rules](#), [Life Sciences](#), Filtering
 - **Web Services**: [Performance](#)*, [Addressing](#)*, [SemWeb Services](#)*, [Policy](#), Security, Trust
 - **XML**: [Binary](#)*, [Schema User Experience](#)
- **Web on Everything**
 - [Mobile Web Initiative](#)
 - [Compound Doc Formats](#)*, [Web Applications](#), [Ubiquitous Web?](#)
- **Web for Everyone**
 - Web Accessibility Initiative ([new charters](#))
 - Expanded [Internationalization](#) Activity
 - [Expanding global participation in W3C](#)

* Recently started

Web Services Stack



... from the [W3C Web Services Architecture Working Group](#)

What is a Web Service?

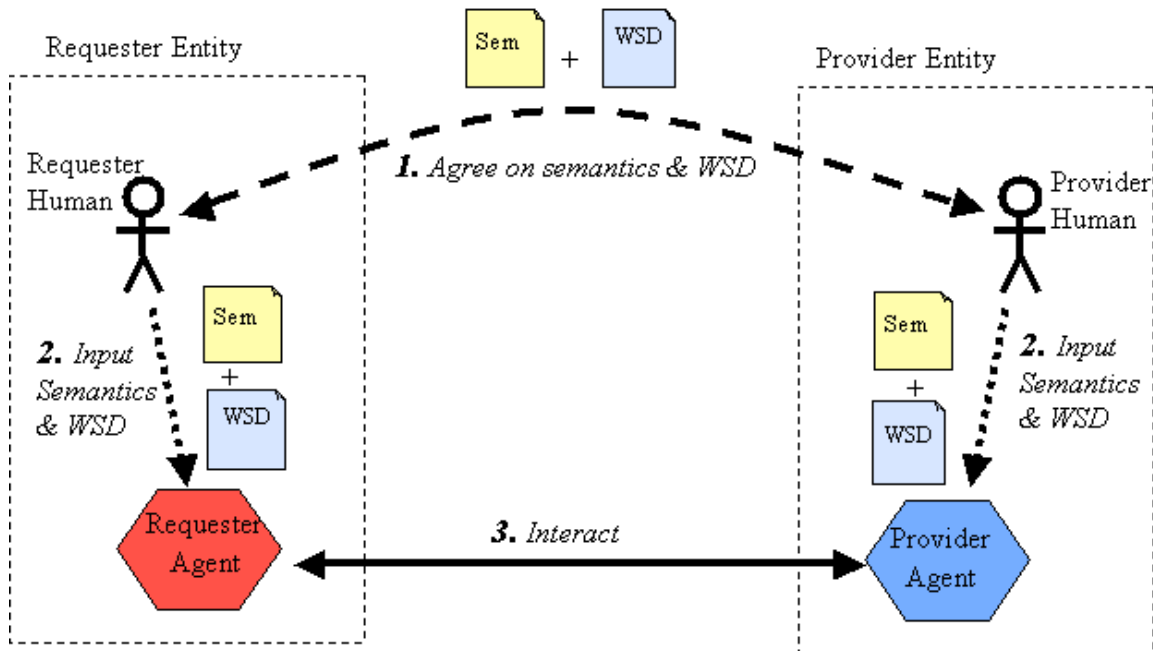
- XML based
- Interoperable machine-to-machine interaction over network
- Interface described in machine-processable format ([WSDL](#))
- Interaction with it using [SOAP](#) messages
- Typically using HTTP

... from "[Web Services Architecture](#)", W3C Web Services Architecture Working Group, 11 February 2004

The Process of Engaging a Service

"Client"

"Service"



- SOAP 1.2 is a W3C standard ([Primer](#))
- WSDL 2.0 being standardized at W3C ([Core](#))

(example: [SOAP 1.1 request \(no header\) to Google search interface](#))

(example: [WSDL 1.1 for interface for a Google search](#))

Other Web Services and Semantic Web References

Standards, specifications, profiles:

- [W3C Web Services Standards](#)
- [W3C Semantic Web Standards](#)
- [Web Services specification development at OASIS](#)
- [Web Services profile development at WS-I](#)

Integration of Web Services and Semantic Web

- [DAML-S \(Ontology for Web services\)](#)
- RDF mapping for WSDL ([W3C](#), [IBM](#))
- W3C Semantic Web Services ([Interest Group](#), [Workshop](#))

Semantic Web Research and Development Groups

- [Massachusetts Institute of Technology, DIG](#)
- [University of Bristol, ILRT](#)
- [University of Maryland, MINDSWAP](#)
- [Digital Enterprise Research Institute](#)

Some Web Services Tools

- [Web Services Project](#) (Apache) apps
- [SOAPy for Python](#) (sourceFORGE)
- [Java Web Services](#) (Sun)
- [.net](#) (Microsoft) [apps](#)
- [webLogic](#) (BEA) [apps](#)
- [Web Services Technology Center](#) (Oracle)
- [WebSphere](#) (IBM) [apps](#)
- ... and many more (Iona, Systinet, ETC.)

* *Most products not yet using new W3C standard versions of SOAP (1.2), WSDL (2.0), etc., but they will in the future.*

Some Semantic Web Tools

RDF Environments (interpreters, etc)

- [Jena](#) (from HP; for Java, includes OWL reasoning)
- [RDFLib](#) (for Python)
- [Redland](#) (in C, with interfaces to Tcl, Java, PHP, Perl, ...)
- [Sesame](#) (RDF Schema based repository and querying facility)
- [cwm](#) (general-purpose data processor for the semantic web)
- [Protege](#) (ontology editor)
- [Access to Relational Databases](#) (white paper)
- [Annotea](#) (annotation of Web documents)
- [Longwell](#) (browse and search RDF datasets)
- [Semantics Toolkit](#) (from IBM; for storage, manipulation, query, and inference of ontologies)

SemWeb Graphical Editors

- [IsaViz](#) (Xerox Research/W3C),
- [RDFAuthor](#) (Univ. of Bristol)

Web Service & Semantic Web Applications

Rapid growth of both ...

- Selected compilations of Web Service apps:
 - <http://www.xmethods.com/>
 - http://directory.google.com/Top/Computers/Programming/Internet/Web_Services/
- One compilation of Semantic Web apps:
 - <http://www.w3.org/2003/11/SWApplSurvey>
- Both
 - <http://wsindex.org/>