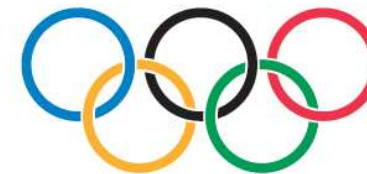


- » BOOST PERFORMANCE
- » REDUCE COST
- » INCREASE AGILITY
- » ENHANCE CRM
- » SHORTEN TIME TO MARKET
- » DRIVE INNOVATION
- » IMPROVE EFFICIENCY
- » INCREASE ADAPTIVITY
- » ENABLE BUSINESS TRANSPARENCY
- » ENSURE REGULATORY COMPLIANCE

Atos
Origin

WORLDWIDE IT PARTNER



Semantic Web. An introduction

TAO Workshop

Tomás Pariente Lobo

Paris, January 2009

TAO



- » **Introduction to the Semantic Web**
- » Semantic Web Services
- » Application of the Semantic Technologies



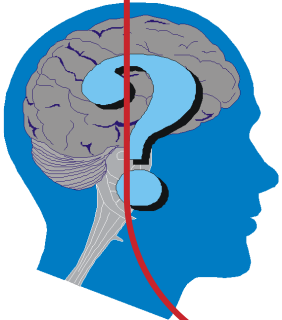
The Web has a revolutionary impact in our way of working, studying or having fun.

The Web provides access (almost) anytime anywhere to everybody

- » The Web presents its contents in multiple ways:
 - » Textual content in natural language (French, English, Spanish, Suomi...)
 - » Multimedia content (video, audio, graphs, ...)
- » We **humans are able to process** all this stuff easily
 - » Making deductions based on incomplete information
 - » Linking concepts using free association
- » And we are used to make use of several senses, taking advantage of several types of interfaces with the Web



- » However, some of the tasks we perform on Internet are the result of **combination** of data coming from different Web resources
 - » Hotels, car rental and flight info for planing personal or business trips
 - » Search on several virtual libraries
- » Besides, humans are able to **combine** all this information, even if they use different terminologies and languages



Too much information on the Web
Too human-oriented
Difficult to organize and automate



However, machines are ignorant

- » They do not know what to do with incomplete information
- » They are not able to extract the meaning of many contents (e.g. a Visio diagram or a *bande dessinée*)
- » There is software able to extract analogies between concepts automatically, but it was traditionally expensive (AI techniques)
- » They have difficulties to combine information
 - » Is it the same <foo:creator> to <other:author> ?
 - » XML is a step, but the differences on terminology still remain

A challenge for the Web

From a human-based Web, to delegate task to software agents
For that, the Web content has to be described in a way that the machines could understand



- » The current Web is based in **HTML**.
- » HTML specifies how to render a **Web page for human consumption**,
- » But **software agents** would need to “**guess**” the meaning.

Las Ketchup

Las Ketchup



Retail Price: \$13.98

Our Price: \$12.99

You Save: \$0.99(7.1%)

Readers' Advantage

Price: \$12.34 Join Now

In Stock: Ships within 2-3 days

<font face="arial, helvetica, sans-serif"
size="+1" color="#000000">Las
Ketchup

**Retail Price: \$13.98
Our Price: **

<font face="arial, helvetica, sans-serif"
color="#990000">\$12.99

■ ■ ■

While driving...

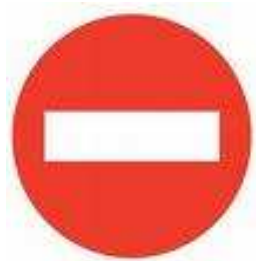
» ...different vehicles...



» ...share the same norms...



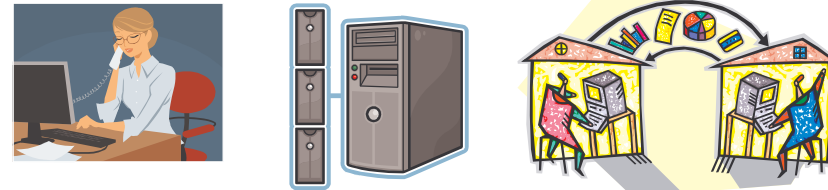
» ...and everybody knows how to proceed



➔ You know this...

Using Semantics...

» ...different agents...



» ...use the same models (ontologies)...

- nciOncology:Protein_Kind
- ▼ ● nciOncology:Technique_Kind
 - ▼ ● nciOncology:Techniques
 - nciOncology:Biological_Testing
 - nciOncology:Bioassay
 - nciOncology:Toxicity_Test
 - ▼ ● nciOncology:Computational_Technique
 - nciOncology:Computer_Modeling
 - nciOncology:Pattern_Recognition
 - nciOncology:Sequence_Analysis
 - nciOncology:Statistical_Methods
 - nciOncology:Diagnostic_Procedure



» ..and the systems knows how to proceed



➔ Logic, axioms, rules



*"The Semantic Web is a vision: the idea of having **data** on the web defined and **linked** in a way that it can be **used by machines** not just for display purposes, but **for automation, integration and reuse of data across various applications**"*

An infrastructure where machines can comprehend semantic data and extends the knowledge of humans

Meaningful information enables machines to "infer knowledge"

Tim Berners-Lee
Web inventor

Las Ketchup

[Las Ketchup](#)



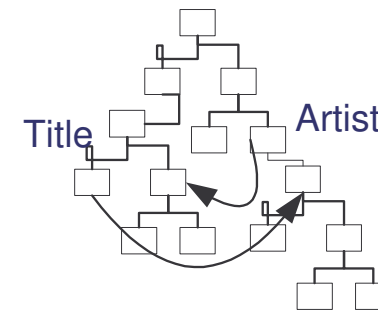
Retail Price: \$13.98
 Our Price: **\$12.99**
 You Save: \$0.99 (7.1%)
 Readers' Advantage
 Price: **\$12.34** [Join Now](#)
 In Stock: Ships within 2-3 days

```
<... xmlns:music_ontology=
"http://www.music.org/#"
xmlns:ecommerce_ontology=
"http://www.ecom.com/#" ...>
```

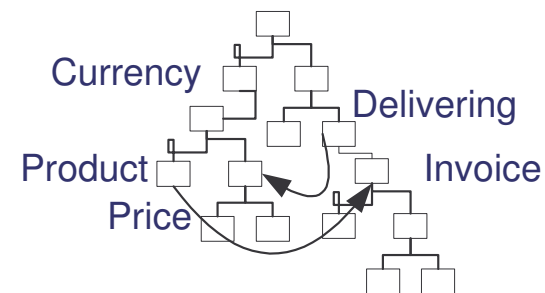
```
<Artist>Las Ketchup</Artist>
```

```
Retail Price <Price>
<currency>$</currency>13.98 </Price>
```

```
Our Price <Price>
<currency>$</currency> 12.99 </Price>
```



Music Ontology



eCommerce Ontology



*"The key notion of the semantic technology is to **represent** meanings and **knowledge separated from the contents** in a digital format **that both humans and machines could access or interpret**"*



Mills Davis
Project 10X's.
Author of the "Semantic Wave 2008 report"

Separating knowledge from
information **making it accessible for**
machine-processing



- » **What do we need?**
 - » Give meaning to the data being displayed
 - » Make it machine-understandable
 - » Tie up all meanings together
 - » Give the machine the tools and intelligence to “deduce” and correctly recognize what humans are using the web for
- » **Building blocks**
 - » **Metadata**
 - Resources are marked-up with descriptions of their content.
 - » **Ontologies**
 - Shared and common understanding of a domain that can be communicated across people and applications



Semantic Web

The Semantic Web can be seen as an infrastructure based on metadata to reason over the Web

Is a horizontal technology that collaborates with many others

Serious Problems in information

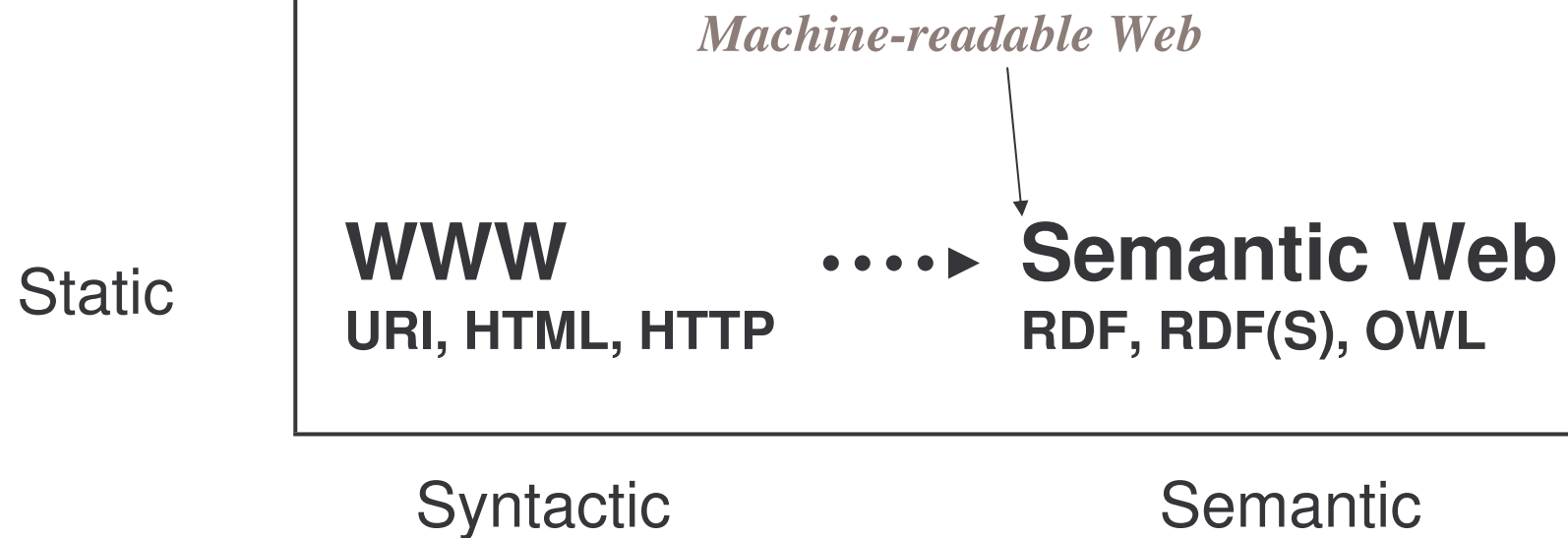
- » *finding*
- » *extracting*
- » *representing*
- » *interpreting*
- » *and maintaining*

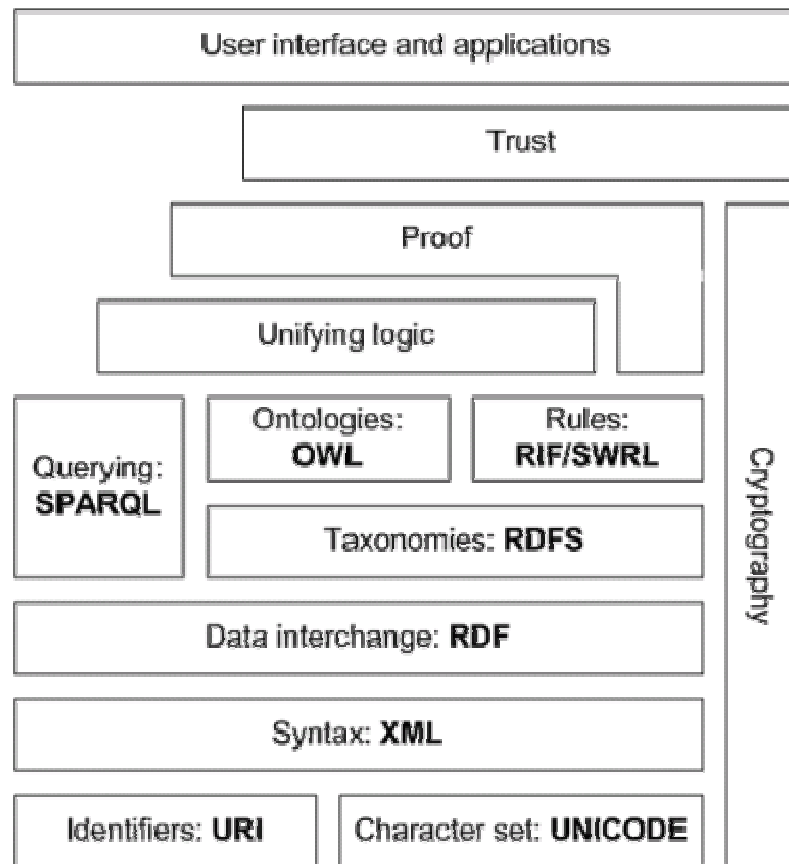
Static

WWW
URI, HTML, HTTP

Syntactic

Billions of users and
Web pages



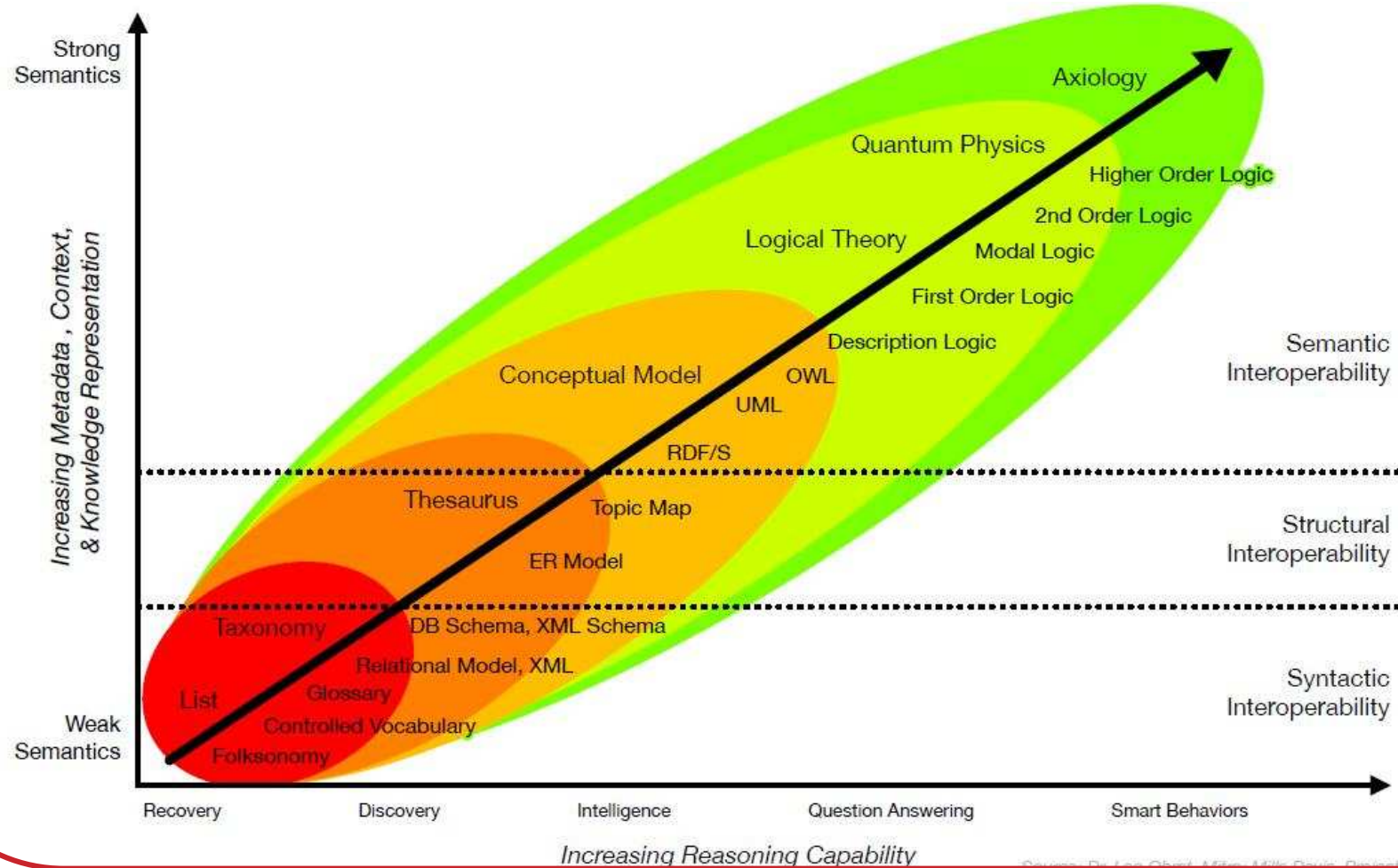


The Semantic Web Stack is an illustration of the hierarchy of languages, where each layer exploits and uses capabilities of the layers below.

It shows how technologies that are standardized for Semantic Web are organized to make the Semantic Web possible. It also shows how Semantic Web is an extension (not replacement) of classical hypertext web.

The illustration was created by Tim Berners-Lee. The stack is still evolving as the layers are concertized





Source: Dr. Les Gallet, MIT, Mike Davis, Project101

Tom Gruber definition

“An ontology is a formal, explicit specification of a shared conceptualization”

Machine-readable

**Concepts, properties, relations,
functions, constraints, axioms
are explicitly defined**



**Consensual
knowledge**

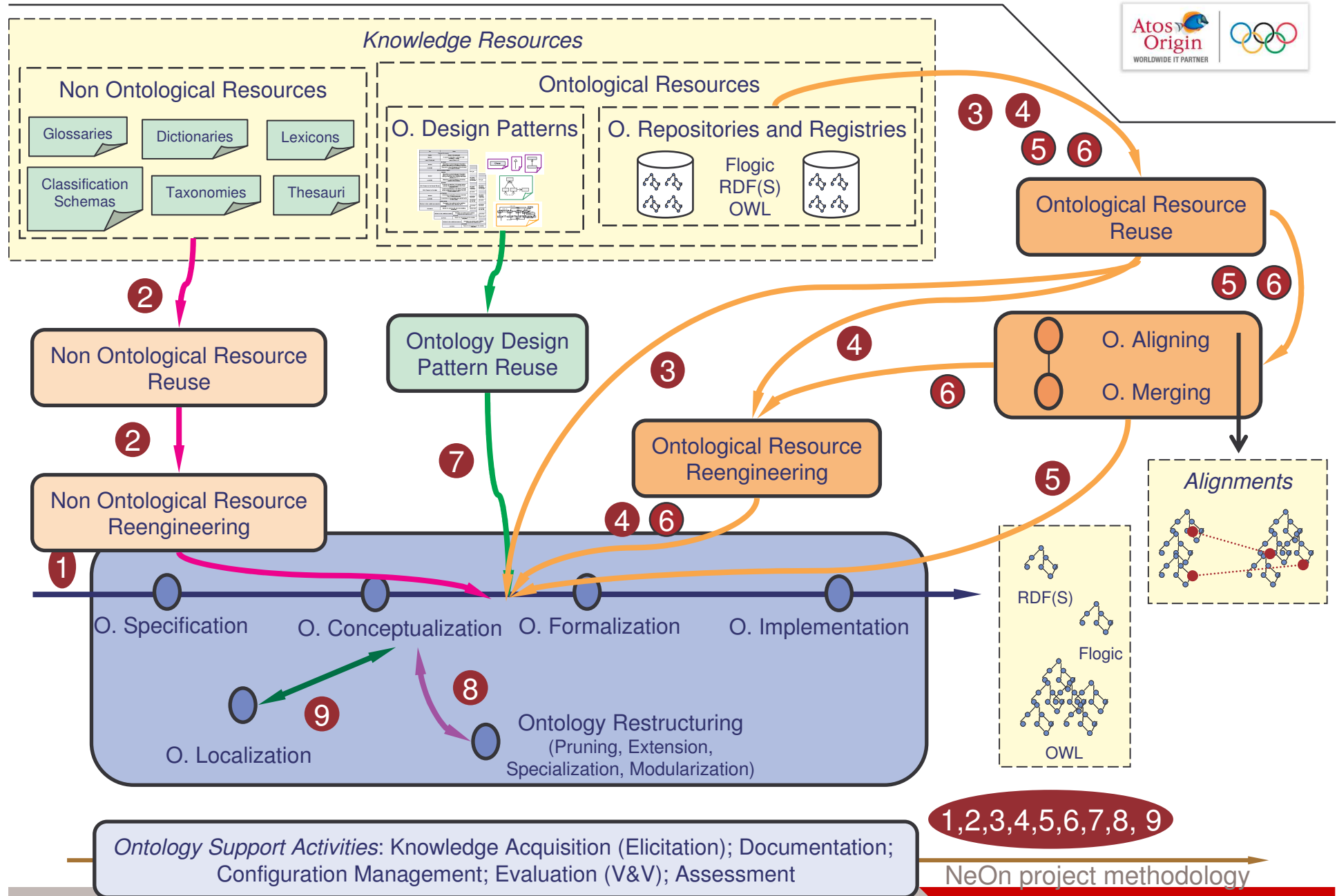
**Abstract model and simplified view
of some phenomenon in the world
that we want to represent**

***"One Ring to rule them all,
One Ring to find them,
One Ring to bring them all and
in the darkness bind them."***

Inscription on the One Ring



- **Concepts:** Basic ideas that are being formalised
- **Relations:** Represents the interactions and links among concepts: subclass-of, part-of...
- **Functions:** Specific type of relation where an element is identified by using a function that considers several ontology elements: *assign-date, apply-vat...*
- **Instances:** Represent specific individuals of a given concept: "Antoine" is an individual of "Person"...
- **Axioms:** Logic theorems declared over relations that the elements of the ontology must comply with: "If A and B are of the class C, then A is not a subclass of B", "For all A fulfilling condition C1, A is B"...



But

- » Manual engineering of ontologies is a very time consuming task!
- » (Semi)automatic support needed to reduce the burden of engineering
 - ... e.g. with Ontology and Instance Learning.
- » Why semi-automatic
 - » A lot of tacit background knowledge, experiences, social conventions, etc, is involved in the modeling process.
 - » Simplifies knowledge adquisition.
 - » However, in order to obtain high quality results,
a human has to be in the loop.
 - » If this were not the case, the Semantic Web would be superfluous!
- » TAO take:
 - » Semi-automatic domain ontology creation from documentation and legacy content



- » Types of document annotation
 - » Marking up contained information
 - Portions of documents associated to objects in ontologies
 - Enable ontology-driven processing
 - » Adding free text annotation
 - » Adding knowledge to documents
 - Document enrichment: helping connecting the document to the rest of the knowledge
- » (Semi)automated annotation
 - » To help manual annotation or to replace human annotators
 - » Simplifies maintenance
 - » TAO take: Knowledge augmentation



- » Introduction to the Semantic Web
- » **Semantic Web Services**
- » Application of the Semantic Technologies

*Bringing the computer back
as a device for computation*

Dynamic

Web Services

UDDI, WSDL, SOAP



Static

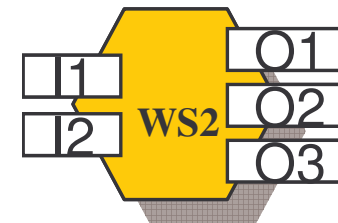
WWW

URI, HTML, HTTP

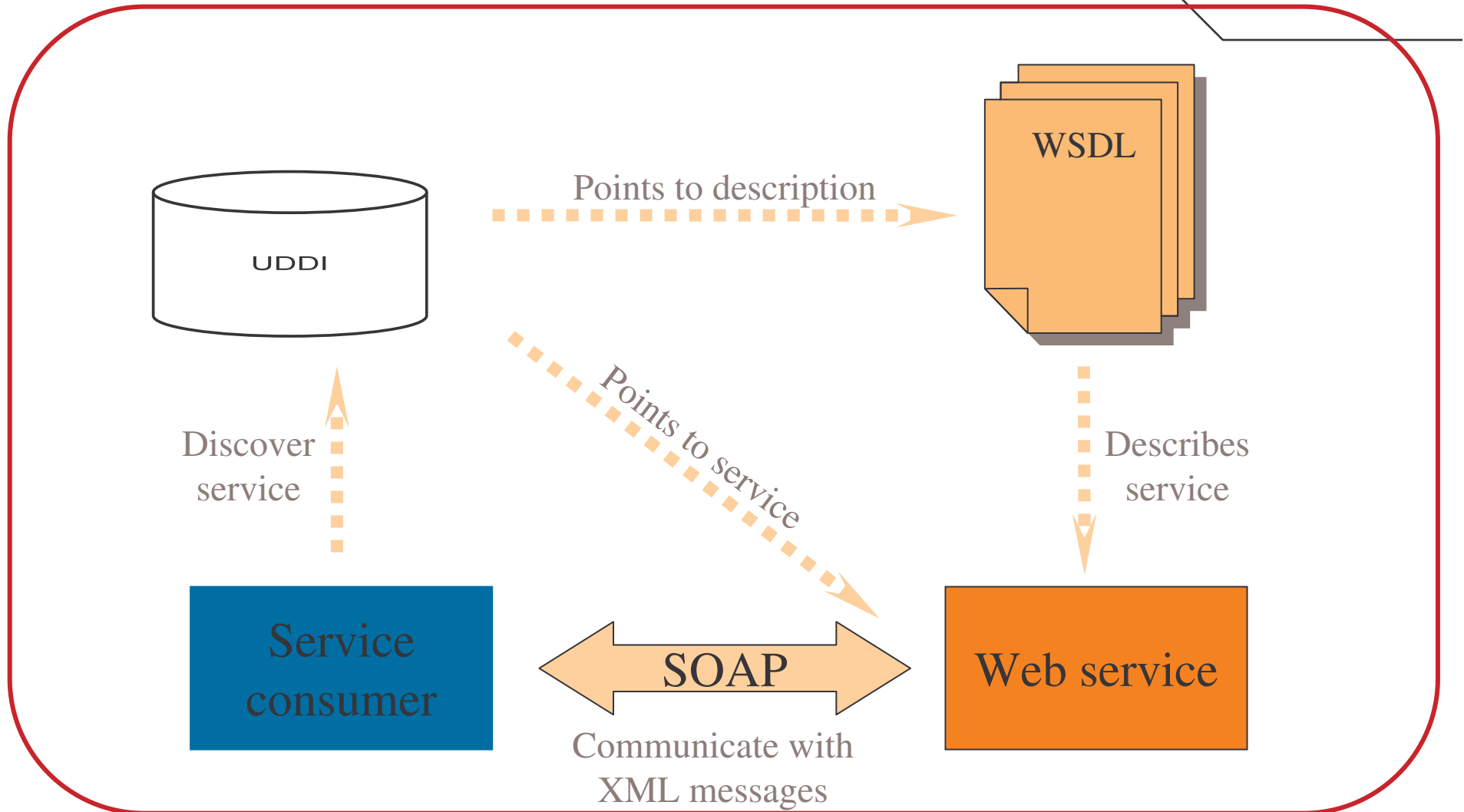
Syntactic

Semantic

- » A web service is a piece of software that is made available on the Internet and utilizes a standardized XML messaging system.
- » In other words ~ A web service is a remote procedure call over the Internet using XML messages.
- » Loosely coupled, reusable components
- » Distributed
- » Add new level functionality on top of the current web

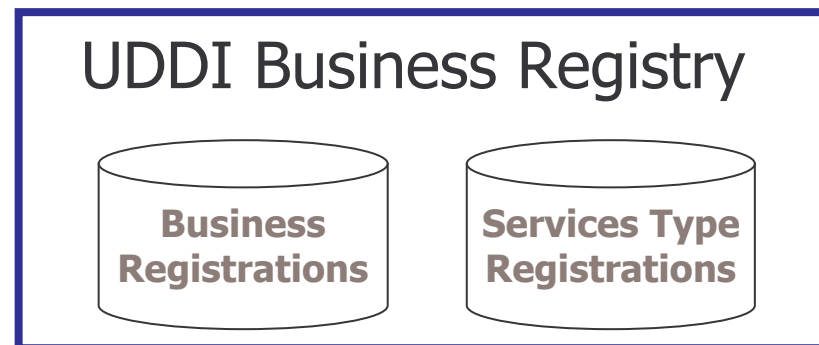


Web services - Framework

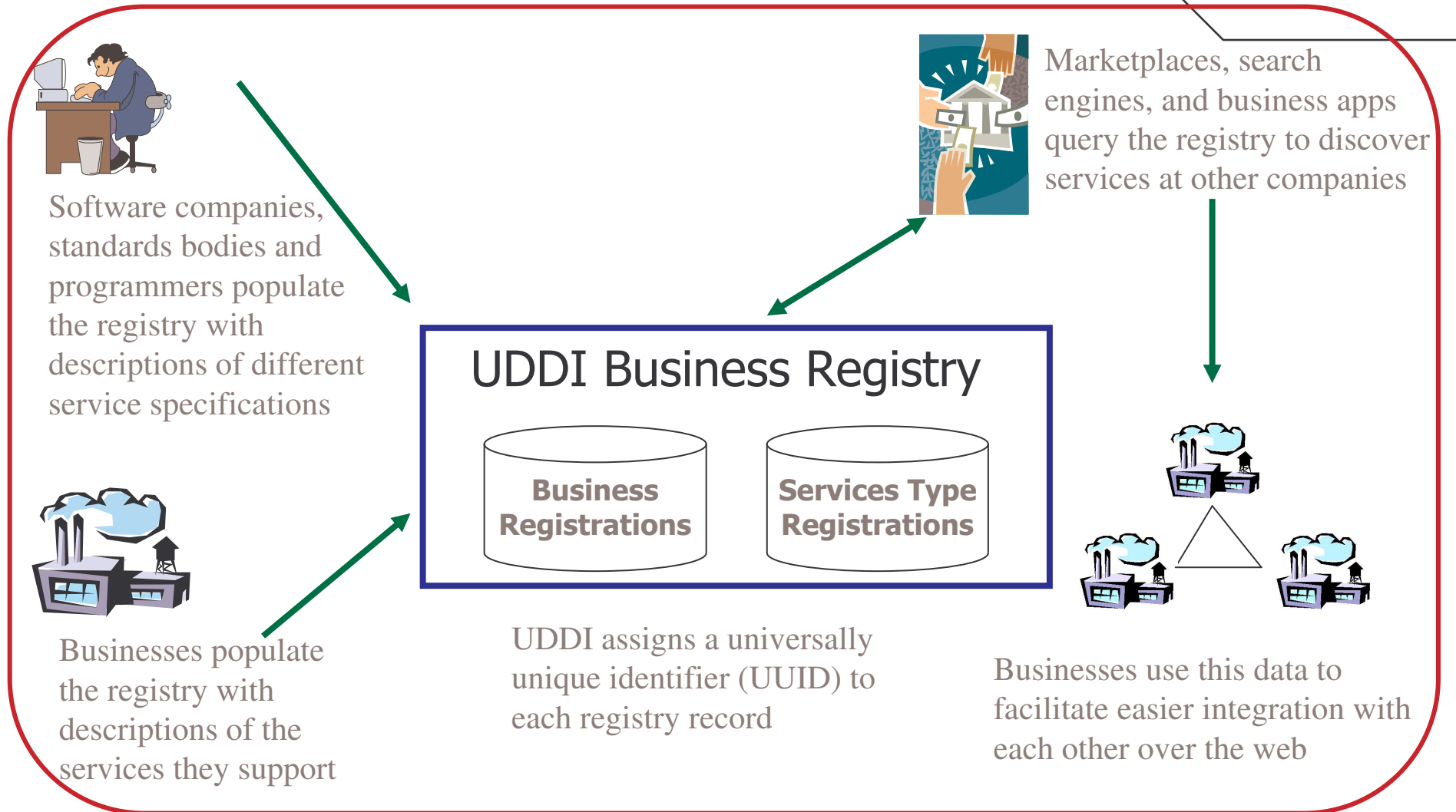




- » **Universal Description, Discovery, and Integration**
- » A project to speed interoperability and adoption for web services
 - » Standards-based **specifications** for **service description and discovery**
 - » Shared **operation** of a **business registry** on the web



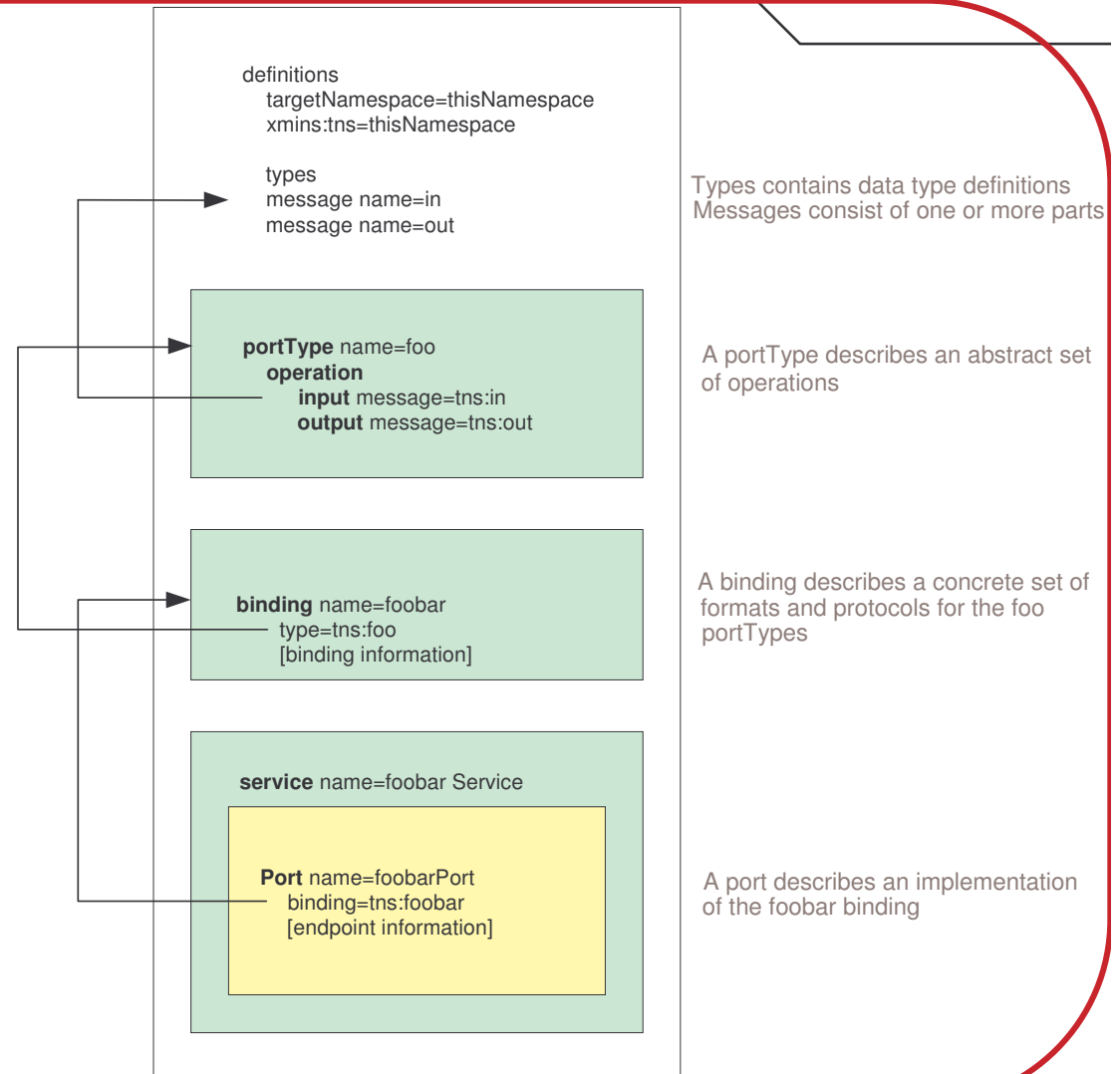
- » Partnership among industry and business leaders





Web services - WSDL

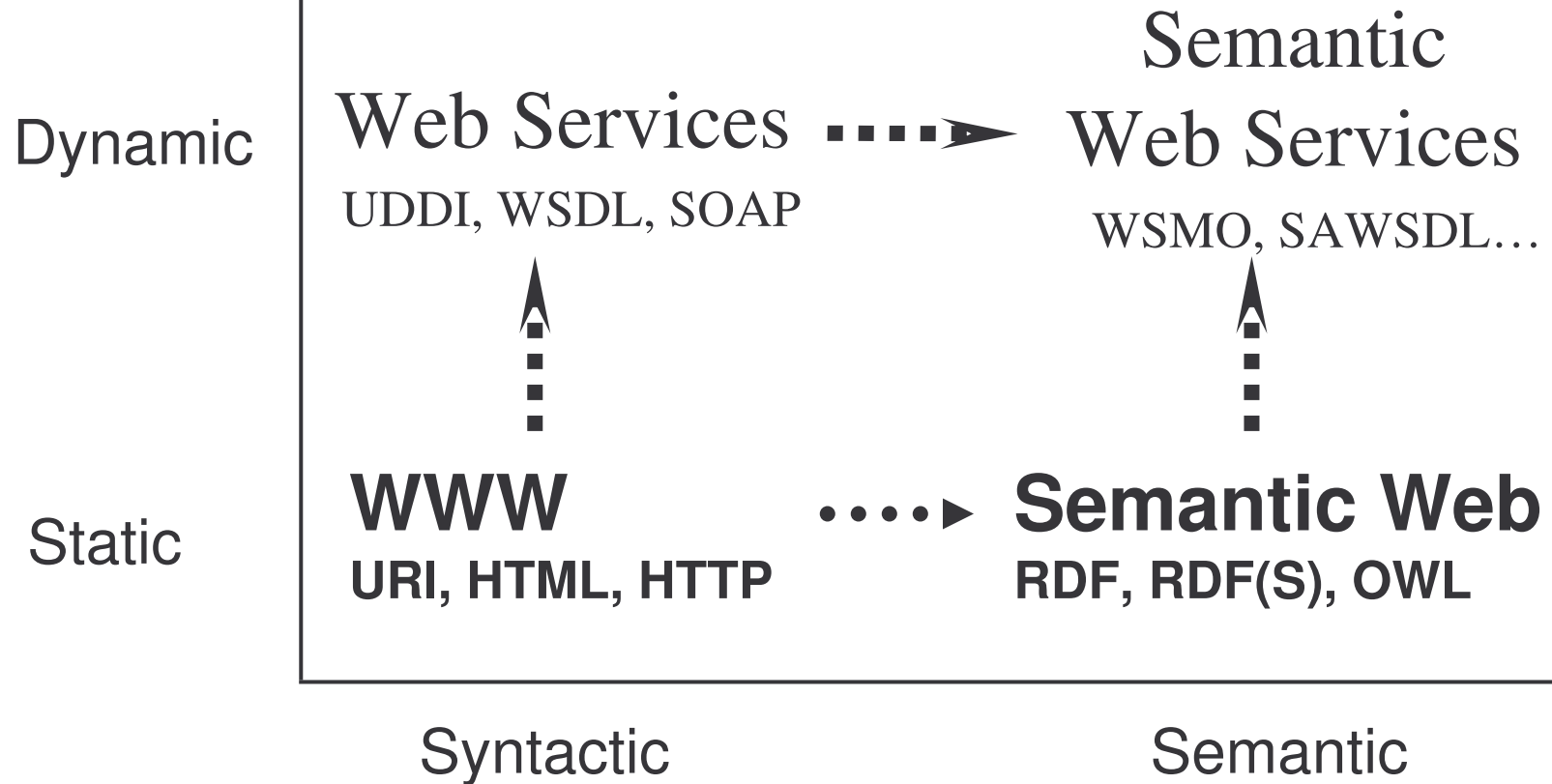
```
<?xml version="1.0" encoding="UTF-8"?>
<definitions>
  <types>
    <!-- define the types here using XML Schema -->
  </types>
  <message>
    <!-- XML messages the web service uses are defined here -->
  </message>
  <portType>
    <!-- define the input and output parameters here -->
  </portType>
  <binding>
    <!-- define the network protocol here -->
  </binding>
  <service>
    <!-- location of the service -->
  </service>
</definitions>
```





- » Problems
 - » Descriptions are syntactic
 - » All tasks associated with web services application development have to be carried out by humans
 - Discovery, composition and invocation
 - » Poor UDDI adoption
 - » Problems of scalability

Bringing the computer back as a device for computation *Bringing the web to its full potential*





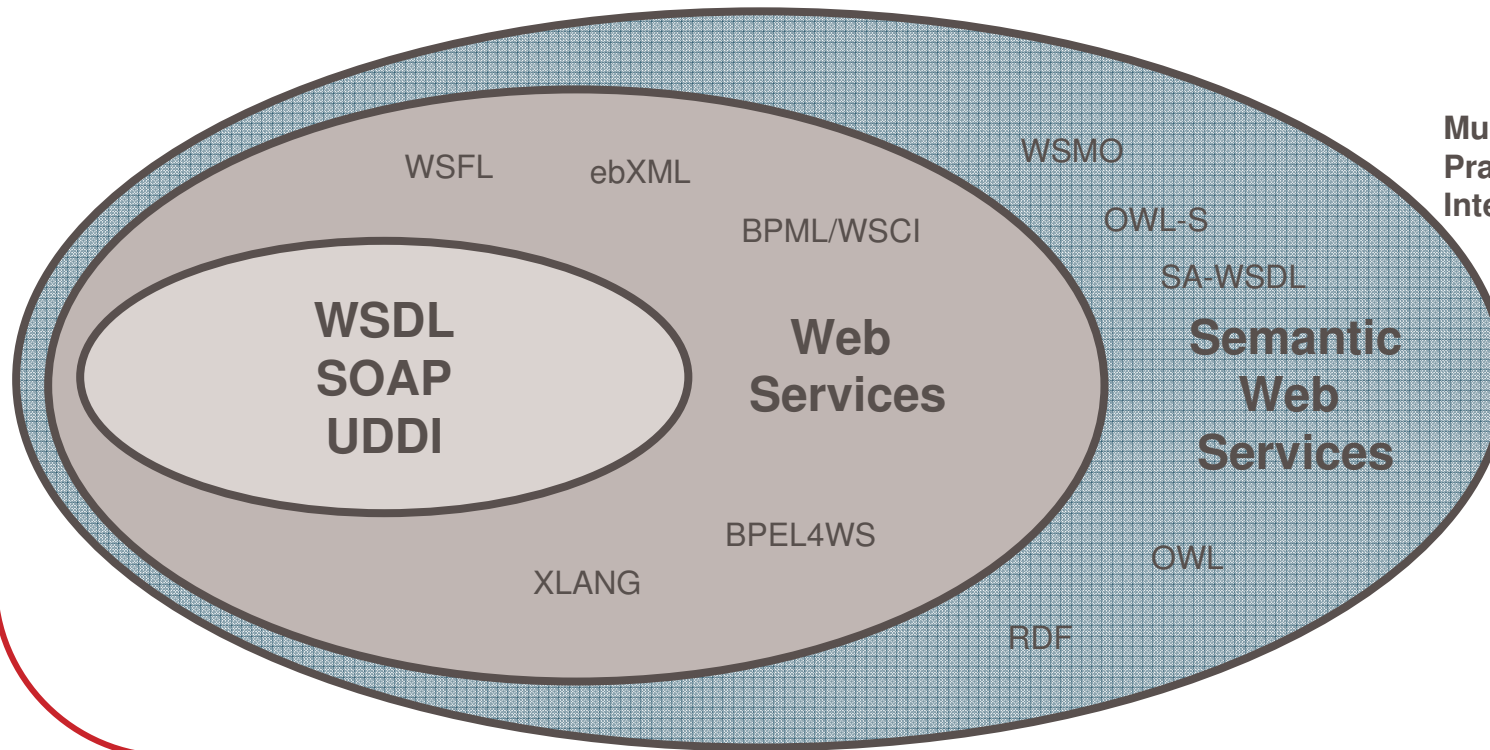
Semantic Web services

- » **Semantic web technology**
 - » Machine readable data
 - » Ontological basis

Applied to

- » **Web Services technology**
 - » To be able to describe service properties
 - » and later search for services according to their properties
 - » giving the possibility of create clever service compositions
 - » and execution of new composed services
- » This search and composition needs to be done in a machine processable and interoperable manner
- » This in turn is possible only by describing the semantics of Web services through ontology languages

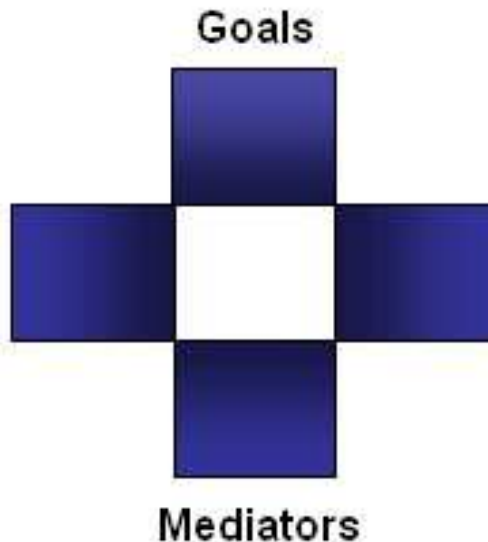
"Self-contained, self-describing, semantically marked-up software resources that can be published, discovered, composed and executed across the Web in a task driven semi-automatic way"



Munindar P. Singh. 2004.
Practical Handbook of
Internet Computing.

Provide the formally
specified terminology of the
information used by all other
components

Ontologies



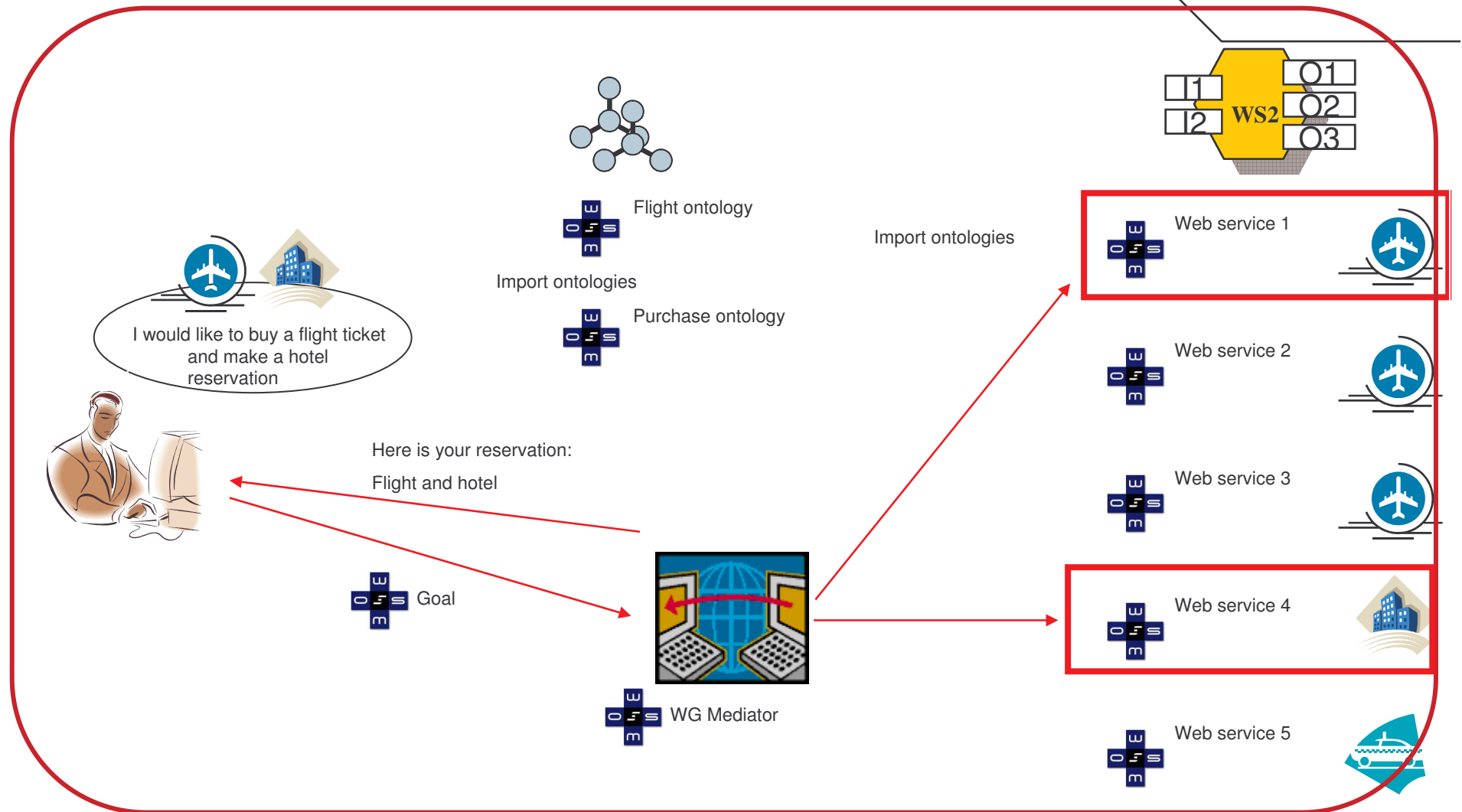
Objectives that a client wants to
achieve by using Web Services

Connector between components with mediation
facilities for handling heterogeneities

Semantic description of
Web Services:

- Capability (functional)
- Interfaces (usage)

Semantic Web Services example



Semantic Web: A bigger example



REAL WORLD

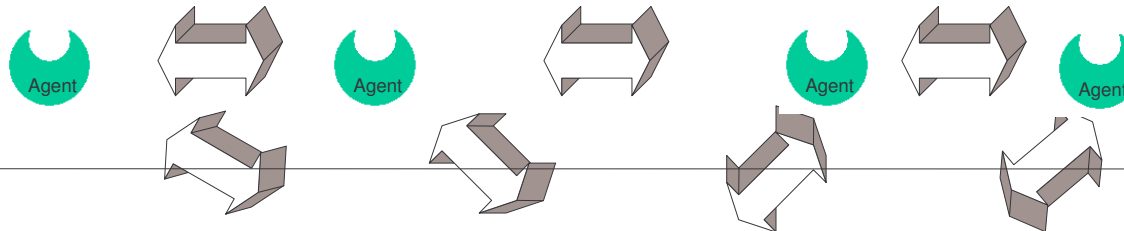
Users



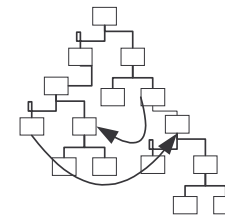
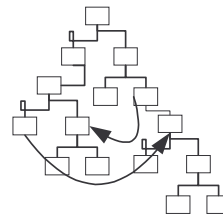
Businesses



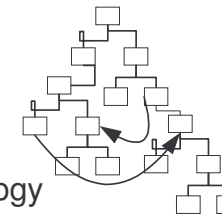
Intelligent
Agents



Knowledge representation
The Semantic Web



Ontology



Web Services

Web Services

Web Services



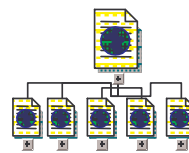
Business
applications



Documents



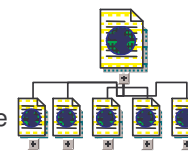
Emails



Static Websites



Online Database
(Schema
unknown)



Dynamic
Websites



Password protected
Websites

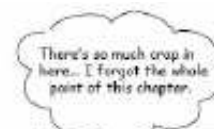


Web Services



- » Introduction to the Semantic Web
- » Semantic Web Services
- » **Application of the Semantic Technologies**

- » Semantic interoperability: **understanding**
- » Transforming information into knowledge: **learning**
- » Content annotation and extraction: **discovery**
- » Semantic search and navigation: **seeing beyond**
- » Semantic desktop: **knowing what I have**
- » Content, service, and application composition: **automating complexity from simple things**
- » Systems that know: **adding intelligence**
- » **Integrated and interoperable** architectures, **collaboration**, ontology engineering, information sharing, **assessment systems** (semantic analysis), **information and knowledge management** (semantic wiki), **content publication and retrieval**



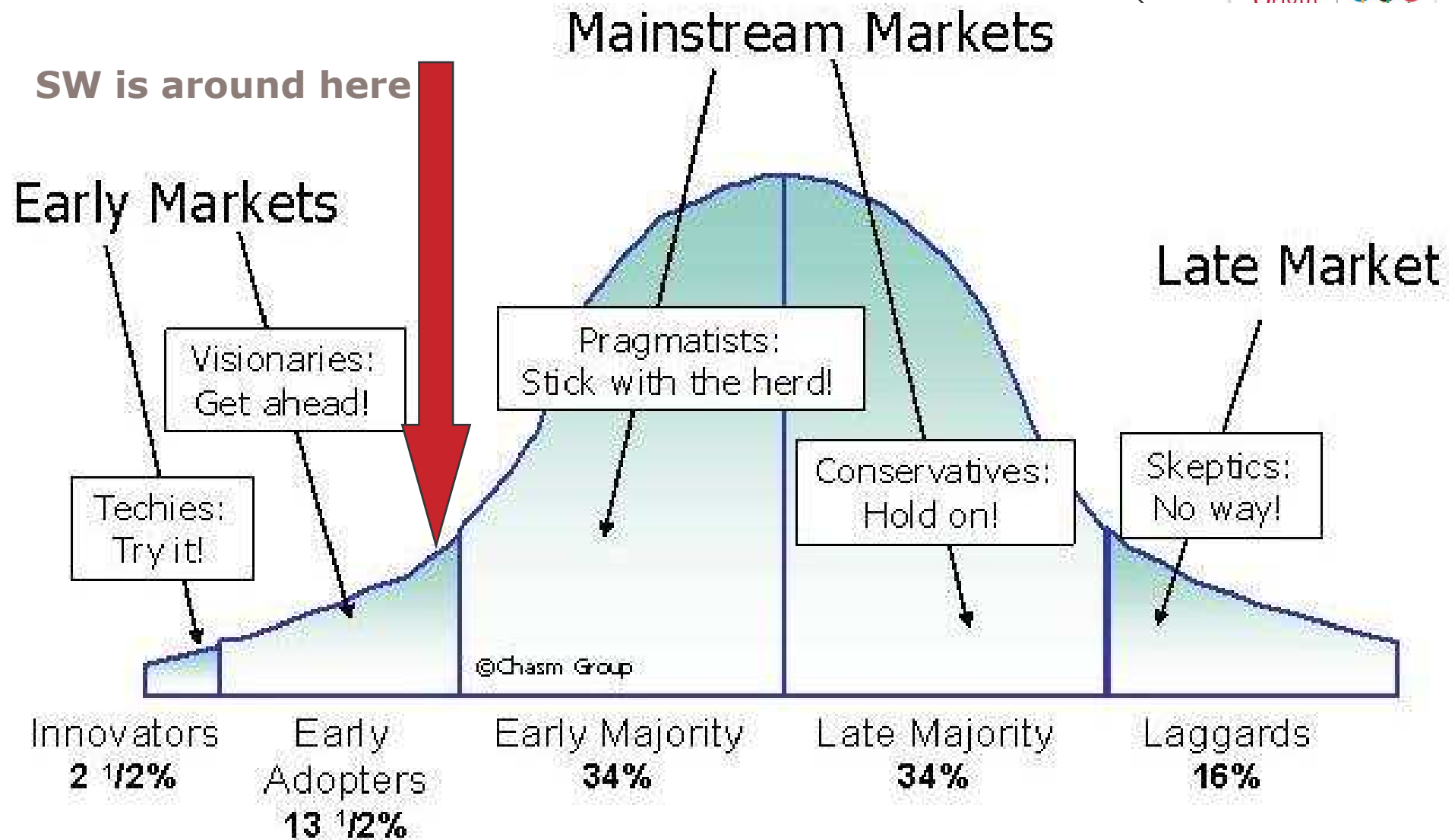
Fuente: Mills Davis

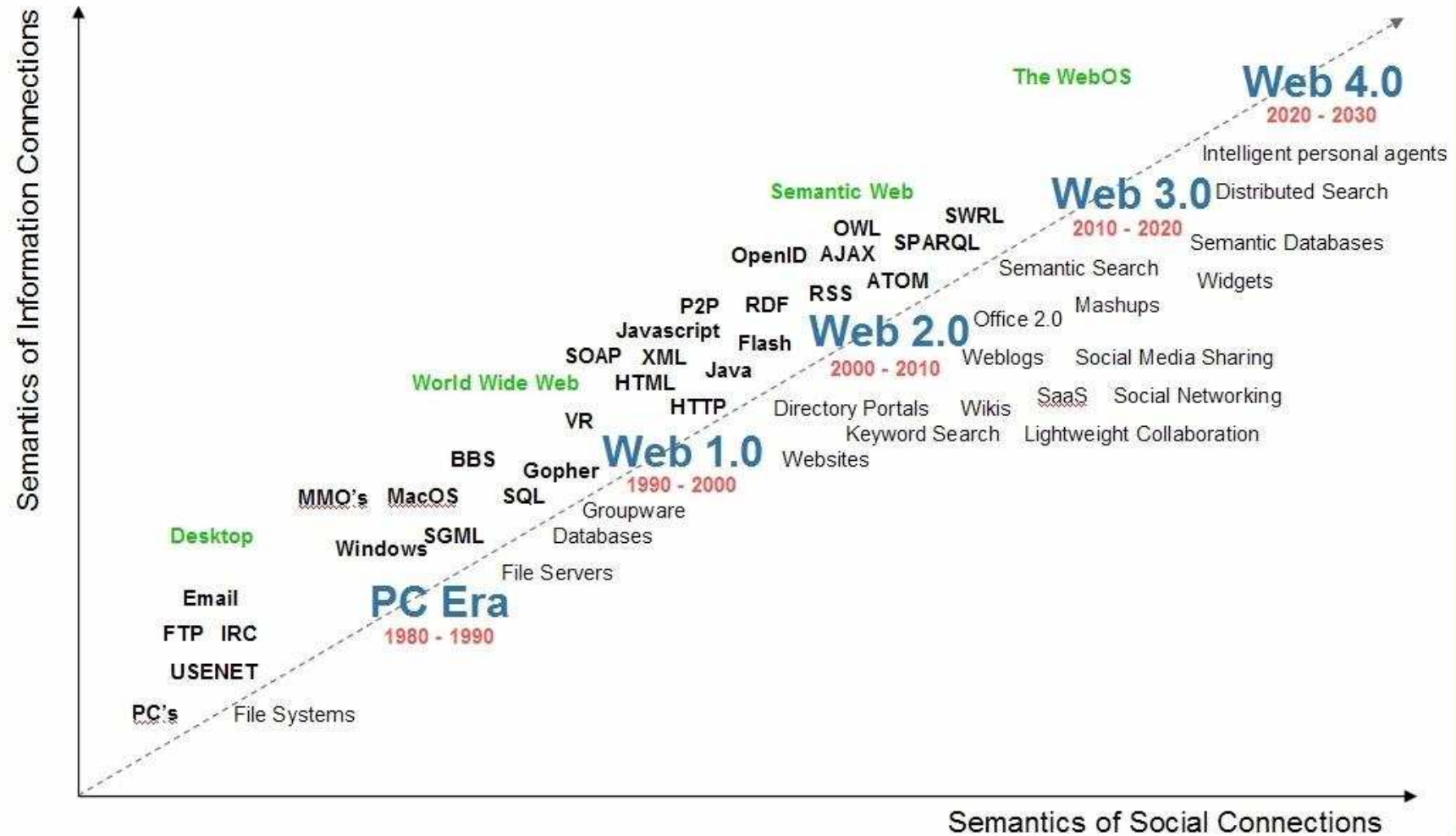
(11th Annual Lockheed Martin Information Technology Trends Conference 2006)





- » The Semantic Web has indeed a strong foundation in research results
- » But remember:
 1. the Web was born at CERN...
 2. ...was first picked up by high energy physicists...
 3. ...then by academia at large...
 4. ...then by small businesses and start-ups...
 5. “big business” came only later!
 6. network effect kicked in early...
- » Semantic Web is now at #4, and moving to #5!





Source: Radar Networks & Nova Spivack, 2007 – www.radarnetworks.com

CEO guide for SW: the “DO-s” and “DON’t-s”



» DO-s

- » **Start small:** Test the Semantic Web waters with a pilot project [...] before investing large sums of time and money.
- » **Check credentials:** A lot of systems integrators don't really have the skills to deal with Semantic Web technologies. Get someone who's savvy in semantics.
- » **Expect training challenges:** It often takes people a while to understand the technology. [...]
- » **Find an ally:** It can be hard to articulate the potential benefits, so find someone with a problem that can be solved with the Semantic Web and make that person a partner.

» DON'T-s

- » **Go it alone:** The Semantic Web is complex, and it's best to get help. [...]
- » **Forget privacy:** Just because you can gather and correlate data about employees doesn't mean you should. Set usage guidelines to safeguard employee privacy.
- » **Expect perfection:** While these technologies will help you find and correlate information more quickly, they're far from perfect. Nothing can help if data are unreliable in the first place.
- » **Be impatient:** One early adopter at NASA says that the potential benefits can justify the investments in time, money, and resources, but there must be a multi-year commitment to have any hope of success.

Source: BusinessWeek Online, April 2007



- » Semantic Web technologies **integrate data, content, applications, and processes via a shared ontology** of concepts, properties, constraints, logic and rules.
- » Semantic Web technologies are “**meaning-centered**”. The building blocks for semantics are data, **metadata**, context, and **ontologies**.
- » By using ontologies, Semantic Web technologies **support data integration, enterprise interoperability, and the discovery and composition of Web Services**.
- » Based on W3C standards, semantic web technologies are used to **create, discover, represent, organize, process, manage, reason with, present, share, and utilize meanings to achieve enhanced integration, interoperability, intelligent content management and knowledge** enabled business capabilities.

SEMANTICS ARE HERE TO STAY

THANKS FOR YOUR ATTENTION



THE SEMANTIC WEB

A new form of Web content
that is meaningful to computers
will unleash a revolution of new abilities

by
TIM BERNERS-LEE,
JAMES HENDLER and
ORA LASSILA

PHOTO CREDIT HERE

Tomás Pariente Lobo – tomas.pariantelobo@atosresearch.eu