

# **The 2<sup>nd</sup> Generation Web - Opportunities and Problems**

---

Dr. Uwe Aßmann

Research Center for Integrational Software Engineering (RISE)

Swedish Semantic Web Initiative (SWEB)

Linköpings Universitet

# Contents

---

- From 1st to 2nd Generation Web -The Semantic Web
  - Use Cases of the Web
  - What does "Semantic Web" mean?
- Opportunities and Problems
  - Standardized Document Processing Architecture
  - Standardized Vocabularies
  - Standardized Context Constraint Languages
- Outlook

---

## The *Semantic Web*

is an extension of the current one,  
in which information is given well-defined *meaning*,  
better enabling computers and people to work in  
cooperation.

**T. Berners-Lee, J. Hendler, O. Lassila**

# The Problem with the 1st Generation Web

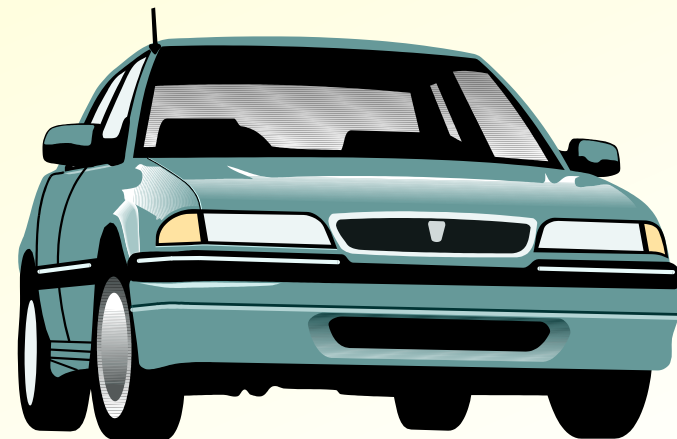
---

- Only syntax
  - Work is based on strings,
  - not concepts
- Only context free structure
  - No context dependencies

# Use Cases of the Web (1): Document Processing

---

- Car manufacturers and their suppliers need to exchange specifications of cars
- They also must pay taxes
- They need different software
- XML is not powerful enough for a uniform document processing architecture



# Use Cases of the Web (2): Search

---

- *"Find the home page of Uwe Assmann"*
- *"Find the home page of this computer scientist, Uwe - I forgot the surname - who is working in Linköping"*



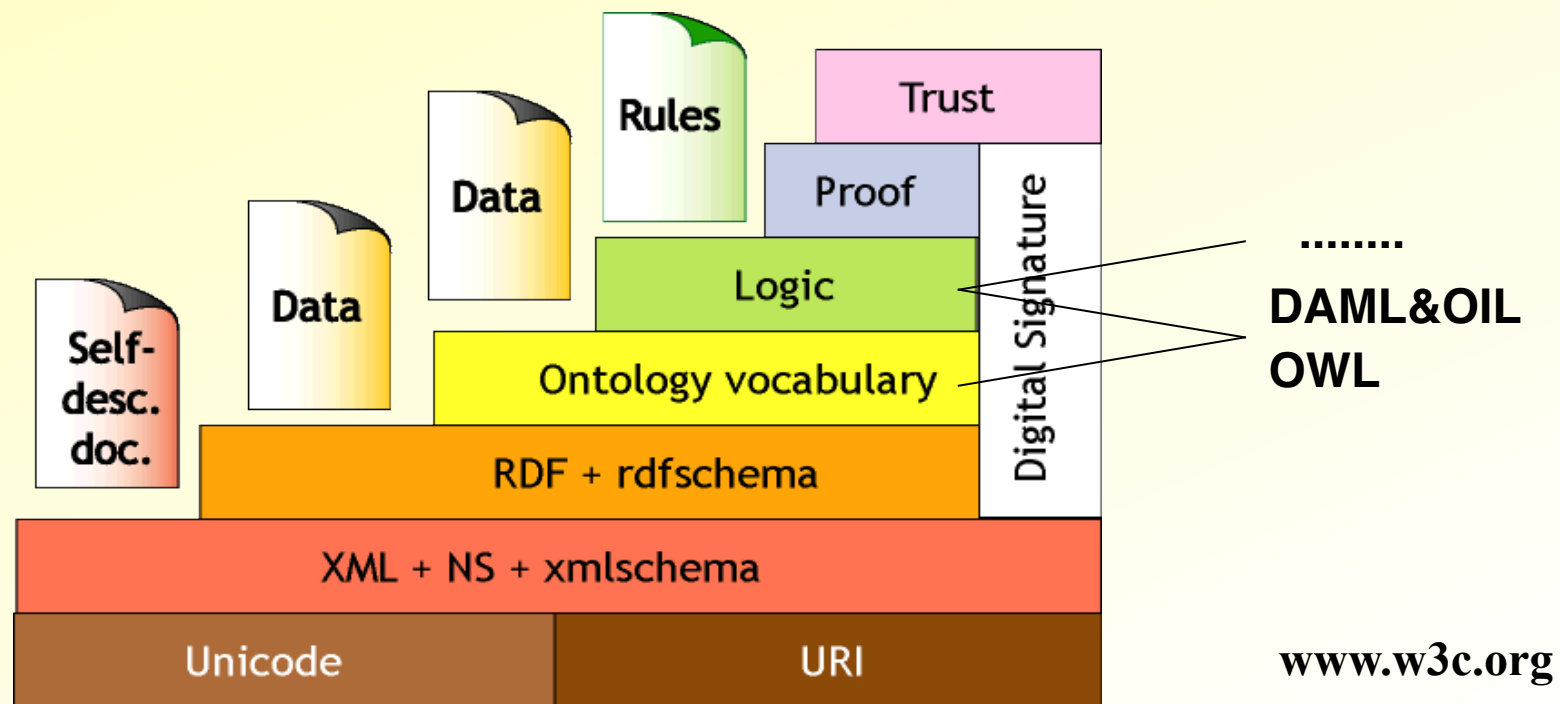
# Use Cases of the Web (3): Web Services

---

- *"Bring a doctor here - as fast as possible - who knows about fever, diabetes, and heart insufficiency"*
- Electronic Yellow Pages
  - Discovery of services
  - Execution of services
  - Composition
- CORBA has the same idea (Trader) but:
  - The CORBA trader works with keyword search
  - No preconditions, postconditions for services
  - Only simple services, no composition
- ... it failed...

# Berners-Lee's Vision with the Semantic Web

- Make web content machine understandable
  - To provide more automation and more service
- Base the web on semantics





# Problem: What Does *Semantics* Mean Here?

---

- An *interpretation function* from a syntactic to a semantic domain
  - Informally: an explanation what the syntax *means*
- Here: a function from XML syntax to an *ontology*
  - ***An explicit and shared specification of a conceptualization***
  - ***A standardized taxonomy with constraints***
- Contains:
  - Terms, partially ordered in a multiple inheritance hierarchy
  - Context constraints between the terms, specified with inference rules

# What Does Semantics Mean Here?

---

- However,...
  - Often, syntactic domain and semantic domain are mixed
- Then, the semantic language degenerates to a constraint language with inheritance
  - I.e., markup is done in a modelling language similar to UML/OCL
  - But executed in a XML processor
  - And standardized
- And the "Semantic Web" degenerates to markups in a standardized modelling language

## 3 Basic Steps Forward in the Semantic Web

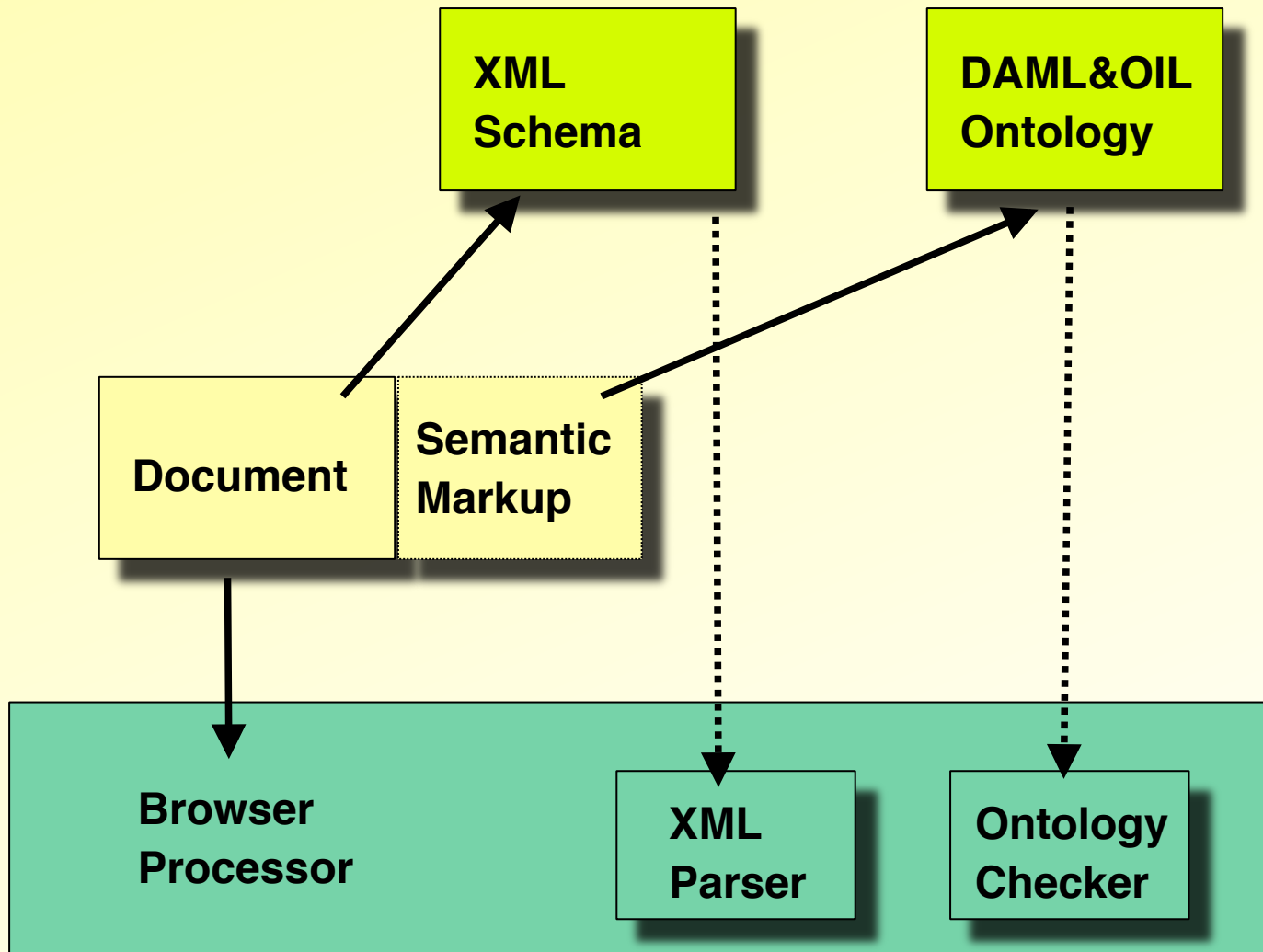
---

- Standardization of *document processing architecture*
  - Standardization of *vocabularies* for the Web (ontologies)
  - Standardization of *context constraints languages*
- 
- *The following shows their influence on the use cases*

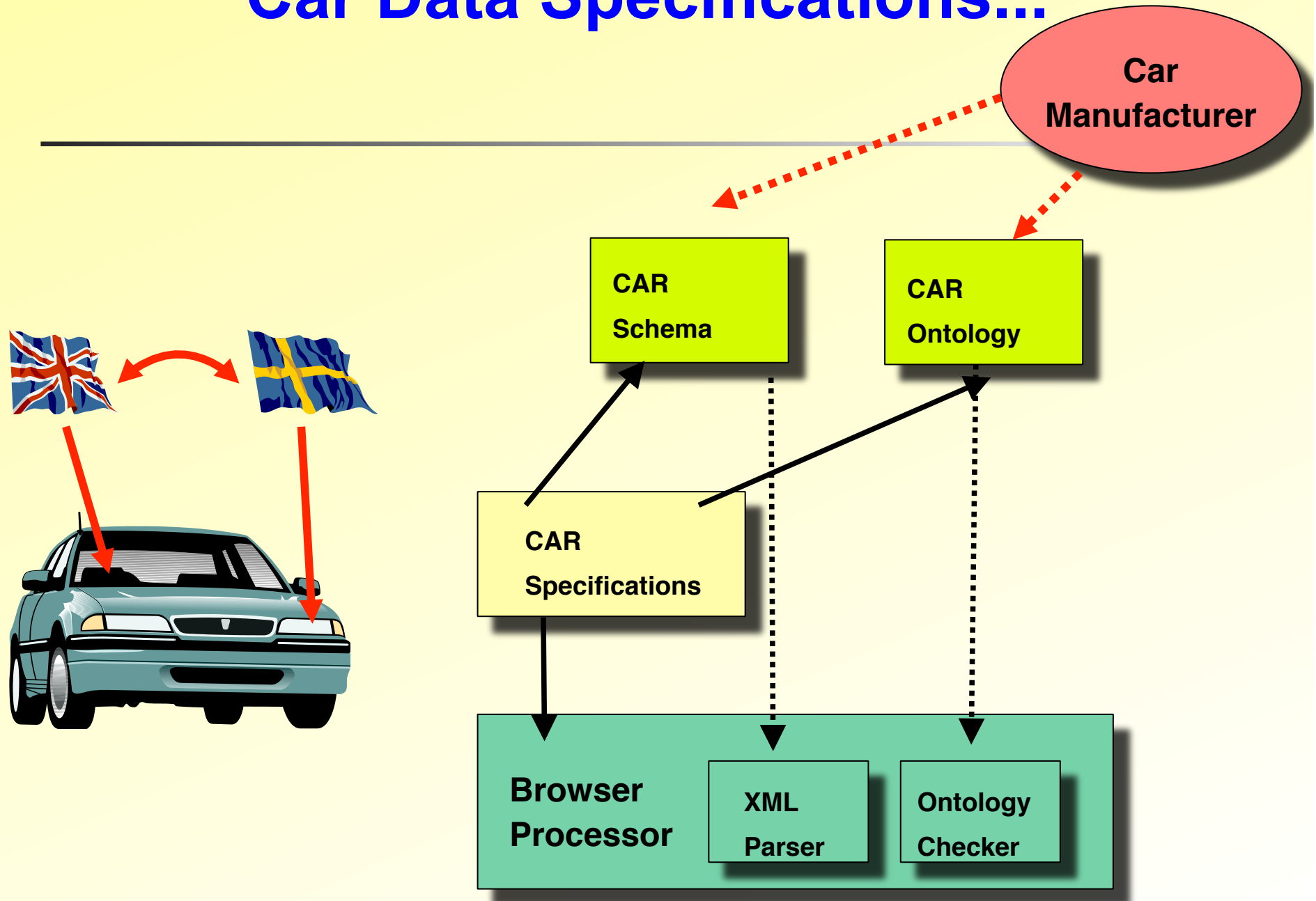
# **Standardized Document Processing Architecture**

---

# Standardized Document Processing Architecture



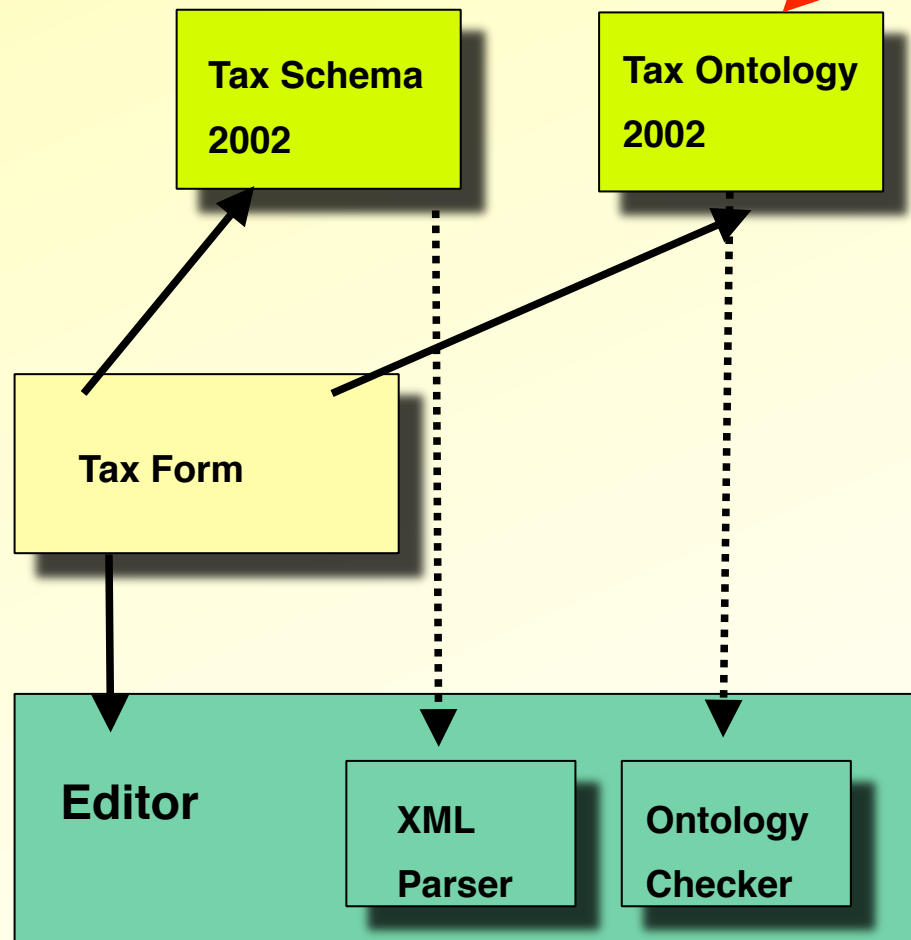
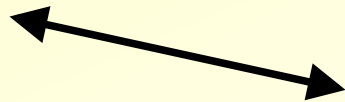
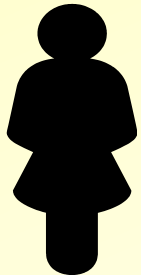
# Car Data Specifications...



# ... Look Similar to Tax Declarations

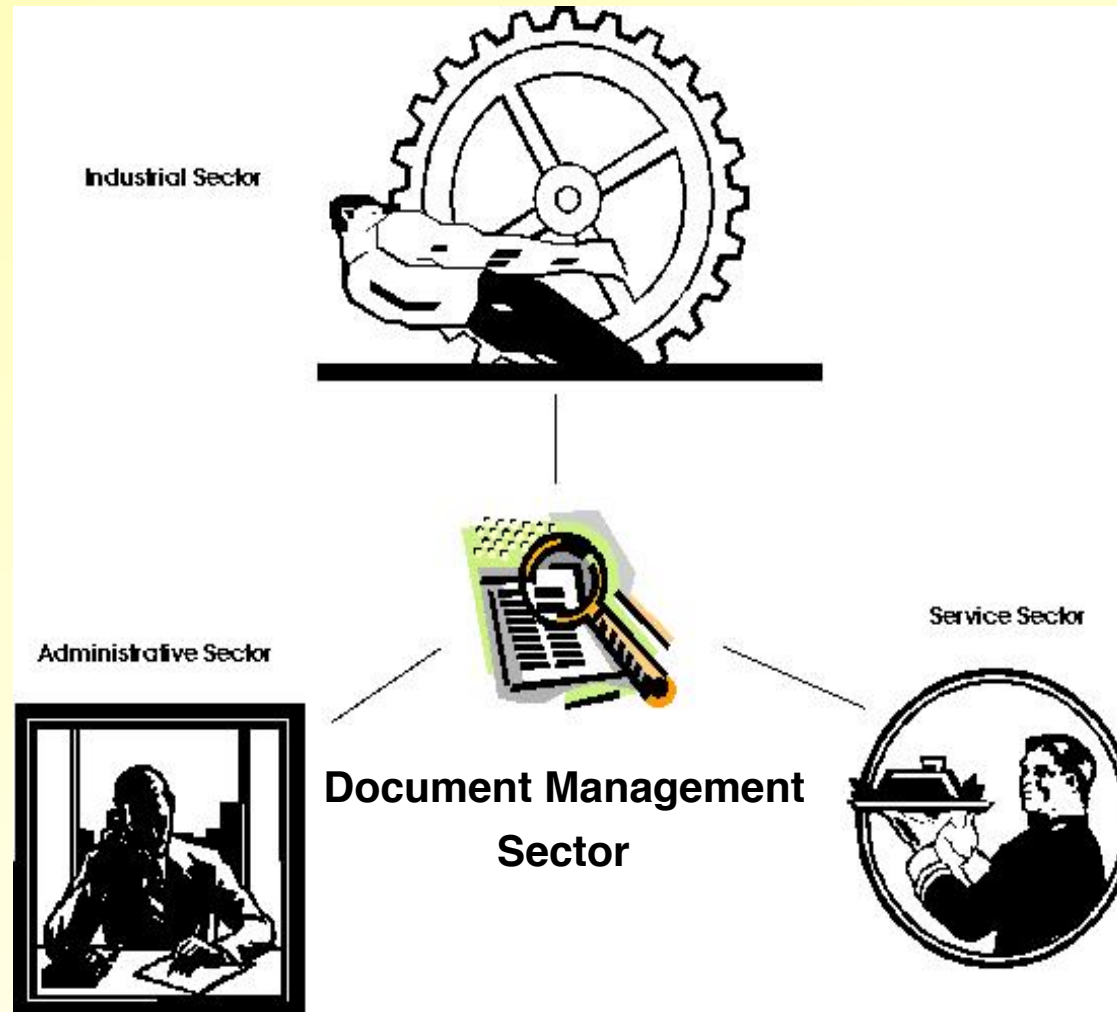
Tax Authorities

*"If you did not earn more interest than 3000SEK, you need not fill in the appendix"*



# This is a Huge Market

---





# Technical Problem

---

- **I want to process some documents, but it takes too long**
  - Evaluation of large ontologies and large documents hard
  - Advanced compiler and generator techniques required

# Stakeholder Problem

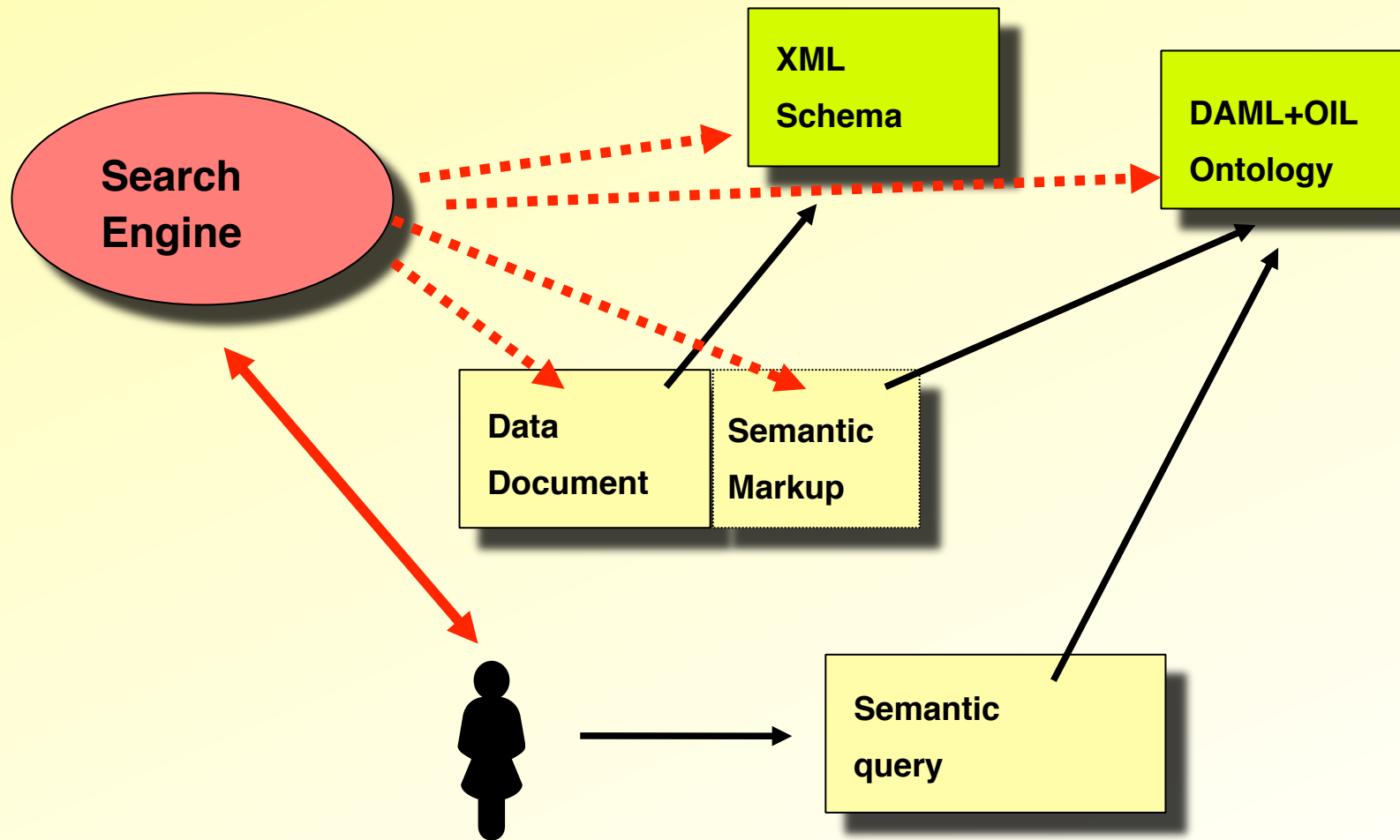
---

- **I want to share things with my friends in private**
  - Intranet vs Extranet is too simple
  - No definition of "groups" on the web possible so far
- **...but society must be secure**
  - September 11 problem: crimes must be prevented
  - P2P networks cannot be controlled at the moment

# **Standardized Vocabularies**

---

# Better Search with Standardized Vocabularies



# Better Search on the Web

---

- Queries can utilize standardized ontologies
  - domain-independent ontologies such as Dublin Core (<http://www.dublincore.org>)
  - domain-specific ontologies
- the vocabularies
  - *"Find the **home page** of **Uwe Assmann**"*
- and their relations
  - *"Find the home page of this **computer scientist**, Uwe - I forgot the surname - who **is working** in Linköping"*
- [www.dmoz.org](http://www.dmoz.org), the free Yahoo-like portal, builds on RDF metadata already
- Search engines from European projects (OntoKnowledge, IBROW)

# Stakeholder Problem

---

- **I want to communicate more efficiently**
  - I'd like to mark up my email
    - so that it can be classified better
- **but I'm too lazy to mark up...**
  - Mark up will slow down my writing
  - Solution: *Markup mining* of documents
    - Specialized knowledge mining
    - Then interactive improvement

# Stakeholder Problem

---

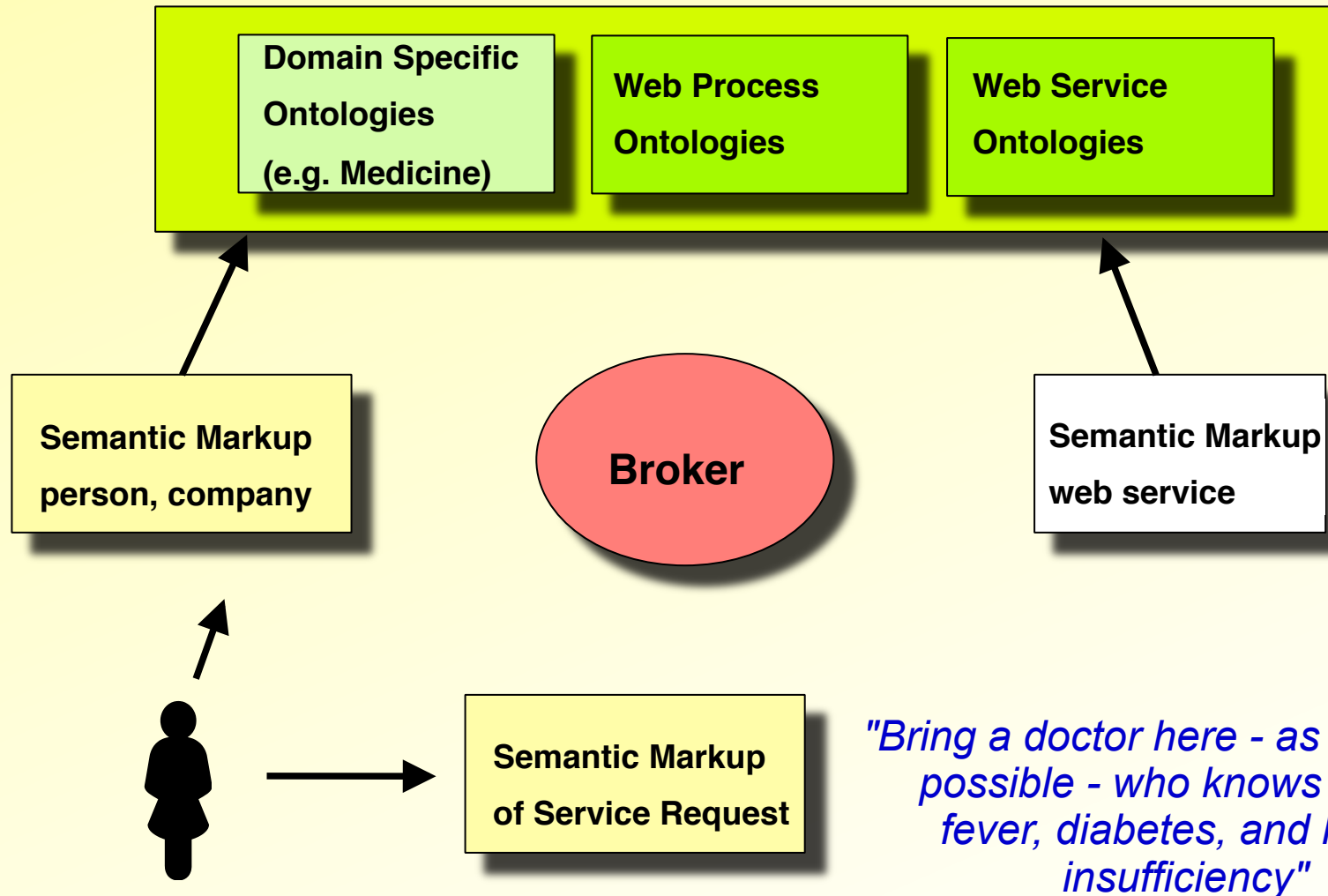
- **Vendor X uses a slightly different ontology than vendor Y**
  - The "Tower of Babel" problem does not vanish
  - Use public standard ontologies such as Dublin Core
  - Mapping and equivalences required to map synonyms in different ontologies onto each other
  - Advanced translation techniques required

# **Standardized Context Constraint Languages**

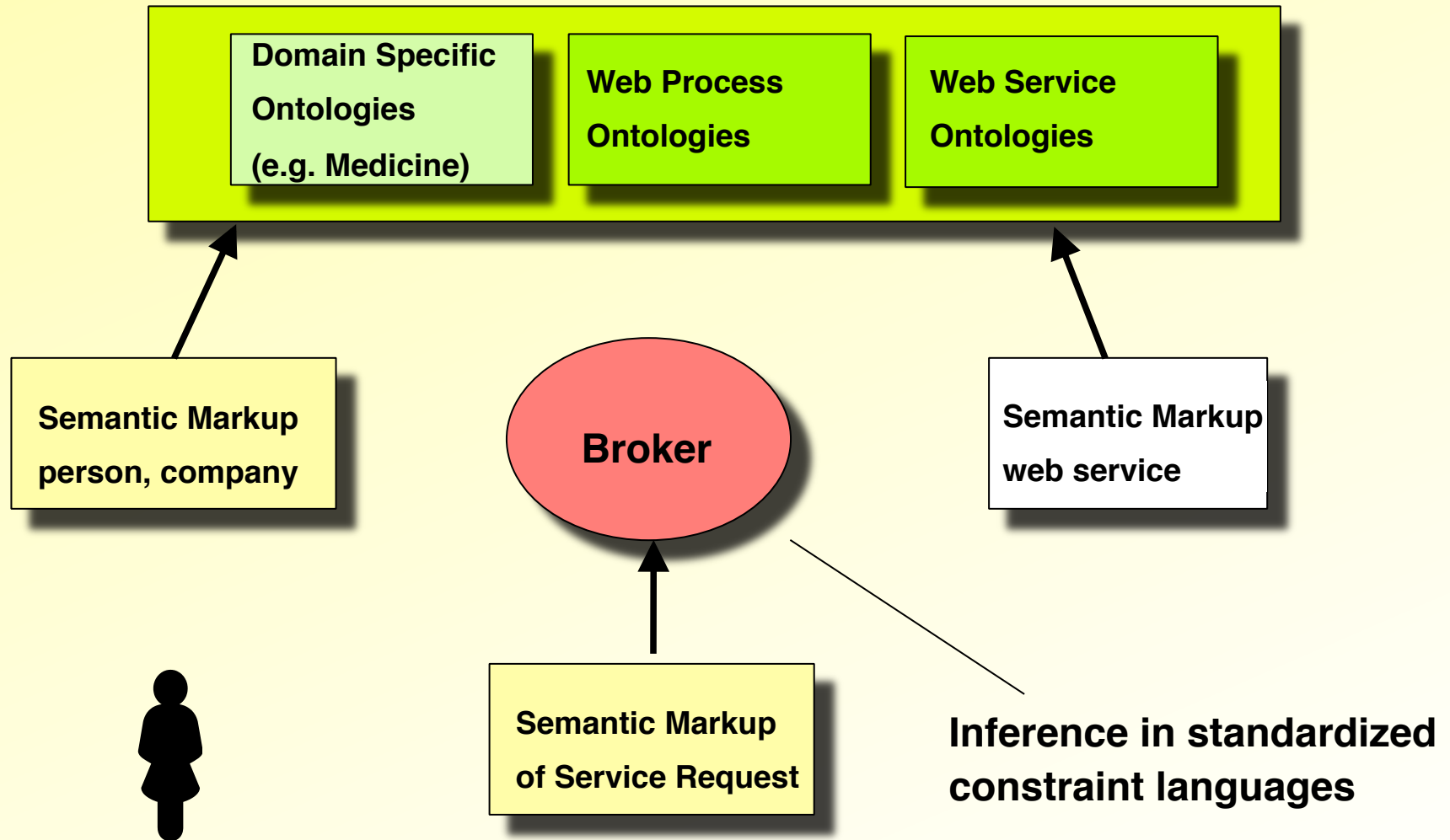
---



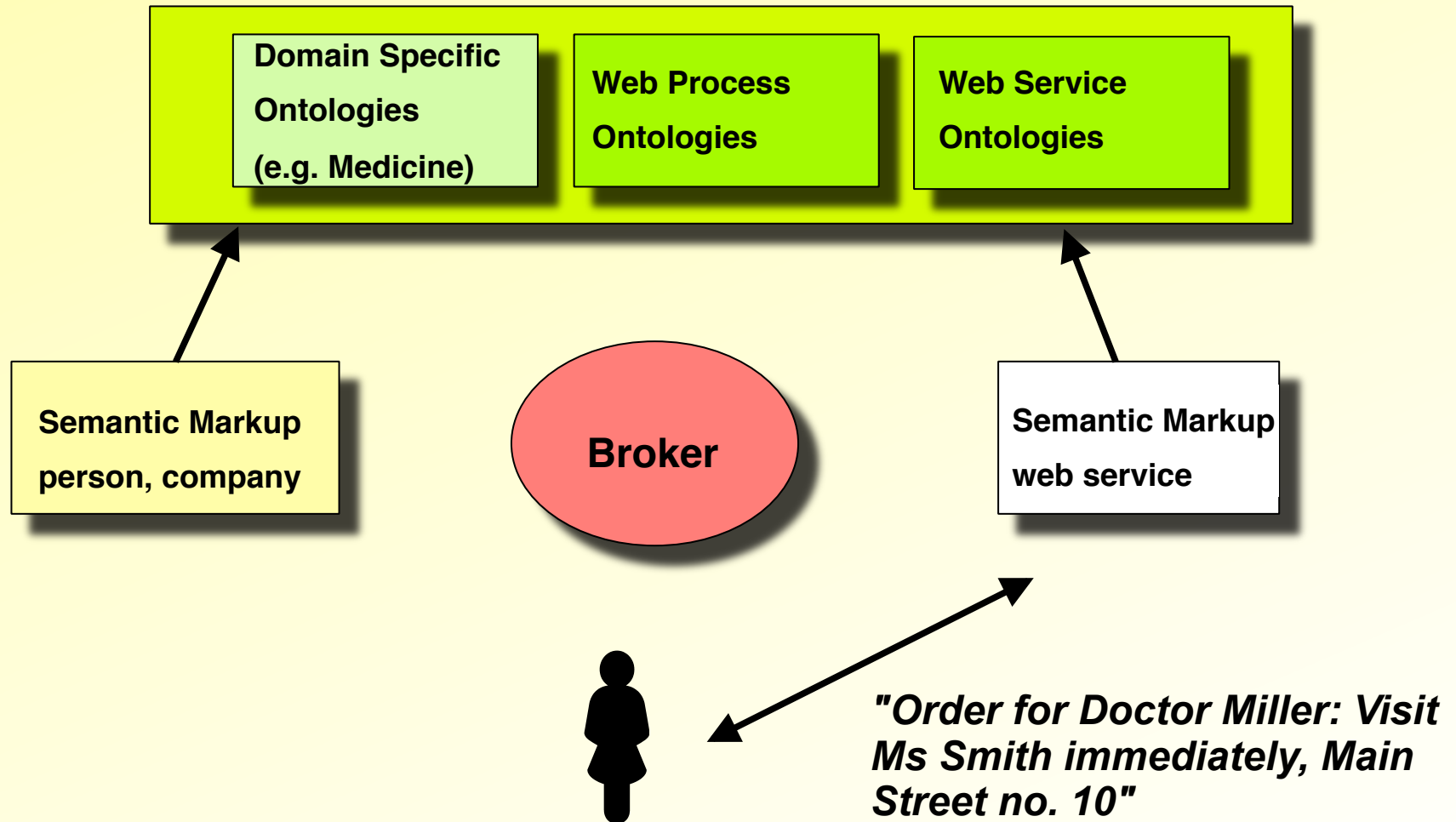
# Match-Making Web Services



# Match-Making Services by Evaluating Constraints



# Match-Making Services



# Stakeholder Problem

---

- **I want to be found, but not to be compared...**
  - Shopping Agents are the enemies of every business
  - They allow for comparison of prices
  - Companies invent dirty tricks not to be comparable
    - Format of outputs in irregular forms
  - No solution...

# Stakeholder Problem

---

- **I want to control who knows about me (information self-determination)**
  - Abuse of information must be prevented (totalitarian governments, economic competitors)
  - The web is one-way: no notification if somebody observed you

# Outlook

---

The most profound technologies are those that disappear.

They weave themselves into the fabric of everyday life until they are indistinguishable from it.

**M. Weiser**

# Will the Semantic Web Be a Profound Technology?

---

- The "Semantic Web" extends the "running horse" XML
- and promises better end-user services by
  - Standardized document processing architecture
  - Standard vocabularies
  - Standard context constraint languages
- However:
  - The stakeholder, technical and security problems should not be underestimated
  - It will take a long time to make the technology "invisible"

# Resources

---

- [www.daml.org](http://www.daml.org) The DAML+OIL committee
- [www.w3c.org/2001/sw](http://www.w3c.org/2001/sw) The Semantic Web activity of the W3C
- [www.semanticweb.org](http://www.semanticweb.org) A nice portal
- [www.ontology.org](http://www.ontology.org) A website for ontologies
- [www.dublincore.org](http://www.dublincore.org) The Dublin Core Ontology
- [www.ontoweb.org](http://www.ontoweb.org) The OntoWeb European Network
- [www.easycomp.org](http://www.easycomp.org) (UKA and LIU's project on component composition for the Web)
- [www.ibrow.org](http://www.ibrow.org) IBROW Project
- [www.ontoknowledge.org](http://www.ontoknowledge.org) (OIL), [www.ontobroker.org](http://www.ontobroker.org), [www.wonderweb.org](http://www.wonderweb.org)
- [www.ida.liu.se/sweb](http://www.ida.liu.se/sweb) The Swedish Semantic Web Initiative (SWEB)



# Bibliography

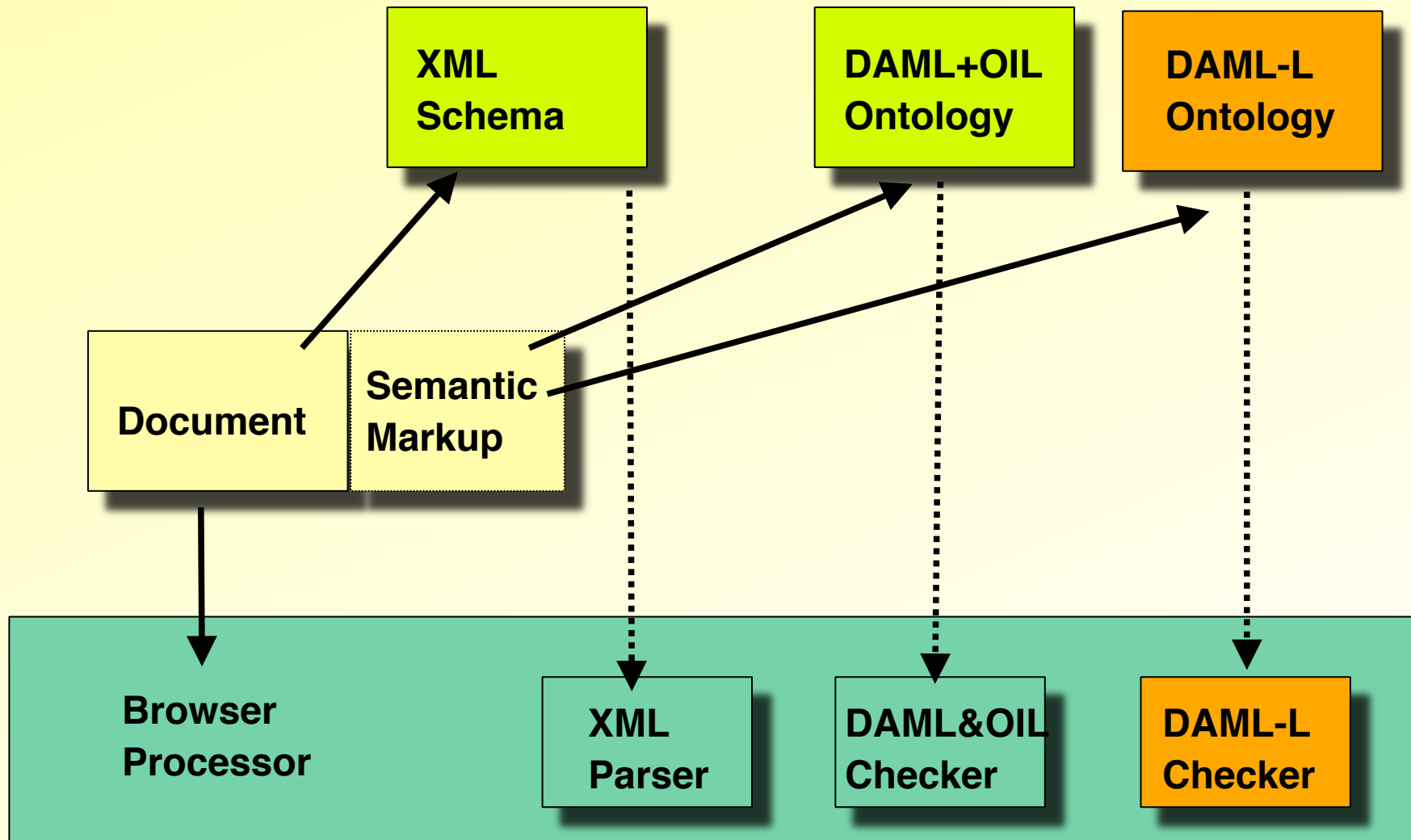
---

- T. Berners-Lee. Semantic Web Roadmap. Sept. 1998. See also <http://www.w3.org/2000/Talks/1206-xml2k-tbl/Overview.html>
- D. Fensel: Ontologies - a Silver Bullet for Electronic Commerce. Springer, 2000
- S. A. McIlraith et. al. Semantic Web Services. IEEE Intelligent Systems, March 2001
- N. Sadeh. The Semantic Web - Challenges, Opportunities, and Challenges. Talk OntoWeb Kickoff, Crete, June 2001

**The End**



# Several Markup Languages can be Referenced



# Web Services and Standardization

---

- Requirement:
  - Uniform document processing architecture
  - Vocabularies for Yellow Pages are standardized
    - domain-independent and domain-specific Vocabularies
  - Constraint languages are standardized
- Goes beyond CORBA services

# Standardized Context Constraint Languages for Web Services

---

- Markup of
  - User and group preferences
  - Web services (advertisements)
    - Prerequisites, consequences
  - Broker processes, partial compositions of web services
- Evaluation combines all markups
- And infers which services are executed when
- Example: DAML-S, a set of ontologies for Web Services
  - [www.daml.org/services](http://www.daml.org/services)

# Stakeholder Problem

---

- **I want web services, but do not want to be traced...**
  - I want anonymous money
  - I don't want to be traced to my location
  - I want anonymous web services

# One of the Languages: DAML&OIL

---

- Language Features
  - Class hierarchy for terms
  - Inheritance on relations
  - Simple inference with subproperties and operators  
Conjunction, Disjunction, Difference
  - Cardinality constraints on domains and ranges of relations (similar to UML)
  - Disjointness specifications for classes and relations
  - Transitive relations
- Based on decidable description logic
- DAML&OIL can be evaluated by checker tools